

**Mathematical
Surveys
and
Monographs**

Volume 64

Stochastic Partial Differential Equations: Six Perspectives

**Rene A. Carmona
Boris Rozovskii
Editors**



American Mathematical Society

Selected Titles in This Series

- 64 **Rene A. Carmona and Boris Rozovskii**, Stochastic partial differential equations: Six perspectives, 1999
- 63 **Mark Hovey**, Model categories, 1999
- 62 **Vladimir I. Bogachev**, Gaussian measures, 1998
- 61 **W. Norrie Everitt and Lawrence Markus**, Boundary value problems and symplectic algebra for ordinary differential and quasi-differential operators, 1999
- 60 **Iain Raeburn and Dana P. Williams**, Morita equivalence and continuous-trace C^* -algebras, 1998
- 59 **Paul Howard and Jean E. Rubin**, Consequences of the axiom of choice, 1998
- 58 **Pavel I. Etingof, Igor B. Frenkel, and Alexander A. Kirillov, Jr.**, Lectures on representation theory and Knizhnik-Zamolodchikov equations, 1998
- 57 **Marc Levine**, Mixed motives, 1998
- 56 **Leonid I. Korogodski and Yan S. Soibelman**, Algebras of functions on quantum groups: Part I, 1998
- 55 **J. Scott Carter and Masahico Saito**, Knotted surfaces and their diagrams, 1998
- 54 **Casper Goffman, Togo Nishiura, and Daniel Waterman**, Homeomorphisms in analysis, 1997
- 53 **Andreas Kriegel and Peter W. Michor**, The convenient setting of global analysis, 1997
- 52 **V. A. Kozlov, V. G. Maz'ya, and J. Rossmann**, Elliptic boundary value problems in domains with point singularities, 1997
- 51 **Jan Malý and William P. Ziemer**, Fine regularity of solutions of elliptic partial differential equations, 1997
- 50 **Jon Aaronson**, An introduction to infinite ergodic theory, 1997
- 49 **R. E. Showalter**, Monotone operators in Banach space and nonlinear partial differential equations, 1997
- 48 **Paul-Jean Cahen and Jean-Luc Chabert**, Integer-valued polynomials, 1997
- 47 **A. D. Elmendorf, I. Kriz, M. A. Mandell, and J. P. May (with an appendix by M. Cole)**, Rings, modules, and algebras in stable homotopy theory, 1997
- 46 **Stephen Lipscomb**, Symmetric inverse semigroups, 1996
- 45 **George M. Bergman and Adam O. Hausknecht**, Cogroups and co-rings in categories of associative rings, 1996
- 44 **J. Amorós, M. Burger, K. Corlette, D. Kotschick, and D. Toledo**, Fundamental groups of compact Kähler manifolds, 1996
- 43 **James E. Humphreys**, Conjugacy classes in semisimple algebraic groups, 1995
- 42 **Ralph Freese, Jaroslav Ježek, and J. B. Nation**, Free lattices, 1995
- 41 **Hal L. Smith**, Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems, 1995
- 40.3 **Daniel Gorenstein, Richard Lyons, and Ronald Solomon**, The classification of the finite simple groups, number 3, 1998
- 40.2 **Daniel Gorenstein, Richard Lyons, and Ronald Solomon**, The classification of the finite simple groups, number 2, 1995
- 40.1 **Daniel Gorenstein, Richard Lyons, and Ronald Solomon**, The classification of the finite simple groups, number 1, 1994
- 39 **Sigurdur Helgason**, Geometric analysis on symmetric spaces, 1994
- 38 **Guy David and Stephen Semmes**, Analysis of and on uniformly rectifiable sets, 1993
- 37 **Leonard Lewin, Editor**, Structural properties of polylogarithms, 1991
- 36 **John B. Conway**, The theory of subnormal operators, 1991
- 35 **Shreeram S. Abhyankar**, Algebraic geometry for scientists and engineers, 1990
- 34 **Victor Isakov**, Inverse source problems, 1990

(Continued in the back of this publication)

This page intentionally left blank

Stochastic Partial Differential Equations: Six Perspectives

This page intentionally left blank

**Mathematical
Surveys
and
Monographs**

Volume 64

Stochastic Partial Differential Equations: Six Perspectives

**Rene A. Carmona
Boris Rozovskii
Editors**



**American Mathematical Society
Providence, Rhode Island**

Editorial Board

Georgia M. Benkart Tudor Stefan Ratiu, Chair
Peter Landweber Michael Renardy

1991 *Mathematics Subject Classification.* Primary 60H15; Secondary 35R60.

ABSTRACT. Stochastic Partial Differential Equations is an interdisciplinary area on the crossroads of stochastic processes (random fields) and partial differential equations. This volume presents the topic of SPDE's from different perspectives, as seen by six groups of researchers working in the most active and promising areas of the field. The goal of this book is to indicate what the main topics of interest are in this fascinating field, and where breakthroughs are being made today.

This book will be of interest to graduate students and researchers in various areas of Mathematics, Physics, Engineering, Economics, etc.

Library of Congress Cataloging-in-Publication Data

Stochastic partial differential equations : six perspectives / Rene A. Carmona, Boris Rozovskii, editors.

p. cm. — (Mathematical surveys and monographs, ISSN 0076-5376 ; v. 64)

Includes bibliographical references (p. —) and index.

ISBN 0-8218-0806-0 (alk. paper)

1. Stochastic partial differential equations. I. Carmona, R. (René) II. Rozovskii, B. L. (Boris L'vovich) III. Series: Mathematical surveys and monographs ; no. 64.

QA274.25.S746 1998

519.2—dc21

98-38392

CIP

Copying and reprinting. Material in this book may be reproduced by any means for educational and scientific purposes without fee or permission with the exception of reproduction by services that collect fees for delivery of documents and provided that the customary acknowledgment of the source is given. This consent does not extend to other kinds of copying for general distribution, for advertising or promotional purposes, or for resale. Requests for permission for commercial use of material should be addressed to the Assistant to the Publisher, American Mathematical Society, P. O. Box 6248, Providence, Rhode Island 02940-6248. Requests can also be made by e-mail to reprint-permission@ams.org.

Excluded from these provisions is material in articles for which the author holds copyright. In such cases, requests for permission to use or reprint should be addressed directly to the author(s). (Copyright ownership is indicated in the notice in the lower right-hand corner of the first page of each article.)

© 1999 by the American Mathematical Society. All rights reserved.

The American Mathematical Society retains all rights
except those granted to the United States Government.

Printed in the United States of America.

∞ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
Visit the AMS home page at URL: <http://www.ams.org/>

10 9 8 7 6 5 4 3 2 04 03 02 01 00 99

Contents

Preface	ix
PART 1. Stochastic Models	1
Chapter 1. Stochastic Partial Differential Equations: Selected Applications in Continuum Physics by J. GLIMM AND D. SHARP	3
1. The Physical Basis of Stochastic Analysis	3
2. Mathematical and Computational Tools for Stochastic Analysis	13
3. Multi-phase Flow	21
4. Transport and Dispersion	34
Bibliography	40
Chapter 2. Measure-Valued Processes and Renormalization of Branching Particle Systems by D. A. DAWSON AND E. A. PERKINS	45
1. Branching and Interacting Particle Systems.	45
2. Historical Brownian Motion	58
3. Formulation of a General Class of Measure-Valued Branching Processes	61
4. Small Scale Behavior	64
5. Large Scale Behavior	74
6. A Survey of Interactive Branching Systems	84
Bibliography	102
Chapter 3. Deterministic and Stochastic Hydrodynamic Equations Arising From Simple Microscopic Model Systems by G. GIACOMIN, J. L. LEBOWITZ AND E. PRESUTTI	107
1. Introduction	107
PART I	
Non Reversible Dynamical Systems:	
Asymmetric Models with Shocks	111
2. The Burgers Equation	111
3. The Asymmetric Simple Exclusion and the Independent Particle System	113
4. The Weakly Asymmetric Simple Exclusion Process: Hydrodynamics and Stochastic Corrections	119
5. Driven Surfaces and Fluctuations	126
PART II	
Reversible Dynamical Systems:	
Symmetric Models with Long Range Interactions	128
6. Ising Models with Kac Potentials: Glauber and Kawasaki Dynamics	128

7. Nonlinear Fluctuations: Stochastic Allen–Cahn and Cahn–Hilliard Equations	135
8. Macroscopic Effects of Small Fluctuations: the Origin of Spatial Patterns	141
9. The Dynamics on Very Long Times: a Brief Look at Large Deviations	148
Bibliography	149
 Chapter 4. Transport by Incompressible Random Velocity Fields: Simulations & Mathematical Conjectures by R. A. CARMONA AND F. CEROU	 153
1. Introduction	153
2. Gaussian Velocity Fields with Kolmogorov Spectra	155
3. Abstract Ornstein Uhlenbeck Velocity Fields	157
4. Simulation of the Velocity Field	159
5. Transport Simulations	170
6. Homogenization & Spectral Singularity Renormalization	176
7. Poisson Models	178
Bibliography	179
 PART 2. Mathematical Theory	 183
 Chapter 5. An analytic approach to SPDE's by N. V. KRYLOV	 185
1. Introduction	185
2. Generalities	186
3. The Stochastic Banach Spaces	190
4. Model Equations	196
5. Equations with Variable Coefficients	207
6. Proof of Theorem 5.1	214
7. Embedding Theorems for $\mathcal{H}_p^n(\tau)$	220
8. Applications	225
9. Open Problems	240
Bibliography	241
 Chapter 6. Martingale Problems for Stochastic PDE's by R. MIKULEVICIUS AND B.L. ROZOVSKII	 243
1. Introduction	243
2. Stochastic Integrals for Cylindrical Martingales in Topological Vector Spaces	246
3. Martingale Problems	261
4. Equations of Stochastic Quantization	295
5. Appendix	317
Bibliography	323
 Indexes	 327
Notation Index	329
Subject Index	331

Preface

This volume is an attempt to present the topic of Stochastic Partial Differential Equations (SPDE's) from different perspectives, as seen by six groups of researchers working in the most active and promising areas of the field.

As the name suggests, Stochastic Partial Differential Equations is an interdisciplinary area at the crossroads of stochastic processes (random fields) and partial differential equations. Interacting particle systems, nonlinear filtering, super processes, continuum physics, . . . have heavily influenced the development of SPDE's. It is safe to say that in the last two decades SPDE's has been one of the most dynamic areas of stochastic processes. Stochastic effects are of central importance for the development of mathematical models of many phenomena in physics, biology, economics, However, most often the resulting (limiting) models end up being deterministic. (Many classes of particle systems and their hydrodynamic limits provide good examples of this effect.) These models usually result from the law of large number type of averaging and they represent the large scale *mean* dynamics of the modeled phenomena. The predictive power of such models might be limited especially for very complex phenomena such as turbulence, phase transition, dynamic instability, chaos, In these and some other situations, stochastic corrections to the deterministic large scale models are very much in order. These corrections are designed to account for the small scales effects neglected in the large scale models. One heuristic way to incorporate these corrections is to perturb the large scale equation by random noise of some kind. Sometimes the stochastic corrections can be derived rigorously by a type of renormalization/central limit theorem of fluctuations around stable large scale models. In fact, in many applications the derivation of the *right* stochastic corrections to the large scale dynamics and their practical utility present one of the fundamental challenges.

Generally speaking, any partial differential equation should be classified as an SPDE if its coefficients, forcing terms, initial and boundary conditions, or at least some of the above are random. Needless to say, this constitutes an extremely diverse area. For example, the analysis of equations with random coefficients has little in common with the analysis of deterministic equations with random free forces. As typical of research fields in their early stages of development, the paradigm of SPDE's is still fairly *soft*: the name SPDE covers different topics for different people. As of now, the subject of SPDE's with its numerous important applications is an exciting mosaic of interconnected topics revolving around stochastics and partial differential equations. It is arguable whether a single book could or should treat the enormously complex field of SPDE's from a unified point of view and this

is not the goal of this book. The goal of this book is to indicate what the main topics of interest in this fascinating field are, and where breakthroughs are being made today. As much as emphasizing breakthroughs however, we tried to explore the role of SPDE's in stochastic modeling, how SPDE's arise and how their theory is applied in different disciplines.

The book is divided into two parts. Part I, *SPDE's and Stochastic Modeling*, deals with fundamental problems in continuum physics, fluid dynamics, statistical mechanics, and branching particle systems and geneses of various types of SPDE's in these fields. Part II, *Mathematical Theory of SPDE's*, is concerned with methods of solutions and related stochastic analysis. Every chapter of the book presents a comparatively self-contained review of a particular sub-field of SPDE's or a related area. Many new previously unpublished results are included as well. Part I consists of four chapters. Chapter 1 emphasizes the conceptual basis for stochastic modeling and more specifically, modeling with SPDE's. This chapter also investigates the relations between stochastic and deterministic models in continuum physics. It should provide the reader with a considerable insight of the role of SPDE's in modeling of physical phenomena. Chapter 2 deals with the asymptotic renormalization of systems of particles undergoing branching and spatial motion and some classes of measure-valued processes, which arise in the limit. The small scale and large scale behavior of these systems are described and various generalizations and extensions of these models which include interactions both between particles and between particles and the medium in which they live are surveyed. Chapter 3 addresses deterministic and stochastic hydrodynamic equations arising from microscopic model systems. Physical and mathematical aspects of stochastic corrections to the hydrodynamic scaling limit for various important types of interacting particle systems are discussed in this chapter. Chapter 4 is devoted to transport of passive scalars by 2D incompressible random velocity fields. The chapter pays special attention to stochastic numerics. Numerical simulations are used both to illustrate the relevance of the theoretical results and to formulate new conjectures, some of them proved later on. Chapter 5 and 6 constitute the second part of the book. Chapter 5 is devoted to linear and some quasi-linear parabolic SPDE's. Equations of this type arise for example in nonlinear filtering of diffusion processes, the problem that was one of the most important original impetuses for the development of SPDE's. This chapter presents a detailed exposition of recent advances in solvability of these equations in Sobolev spaces and spaces of Bessel potentials with exponent of summability greater than or equal to 2. Chapter 6 is concerned with nonlinear stochastic PDE's with nonsmooth (in some cases singular) coefficients. The examples include stochastic Navier-Stokes equation, Langevin (stochastic quantization) equation in Euclidean quantum field theory, SPDE's for the super-Brownian motion and some related super processes. The emphasis in this chapter is on existence and uniqueness of weak solutions, absolute continuity and singularity of distributions, and ergodicity problems for these equations.

Of course, this book could not possibly cover all or even most of the important developments and problems of SPDE's. However, we believe that it will provide the interested reader an informative snapshot of this rapidly developing area. The idea of the book was conceived at the Workshop on Stochastic Partial Differential Equations held at the University of Southern California, Los Angeles, in January

of 1996. In more than one way the discussions and the lectures at the Workshop prompted this book.

Acknowledgments

As the editors of this volume and the organizers of the Workshop, we would like to take this opportunity to acknowledge the support of the Army Research Office, the Office of Naval Research, and the Institute for Mathematics and Its Applications. Our thanks are due to Patricia Shapiro for the help in editing this volume. Finally, we would like to thank the AMS staff for their cooperation and patience. The contribution of J. Glimm was partially supported by the Applied Mathematics Subprogram of the U.S. Department of Energy DE-FG02-90ER25084, the Army Research Office, grant DAAL03-92-G-0185 and the National Science Foundation, grant DMS-9500568 while D. Sharp was supported by the U.S. Department of Energy. Both D. Dawson and E. Perkins would like to acknowledge the support of NSERC of Canada Research Grants and of a joint NSERC of Canada Collaborative Grant. During the preparation of their manuscript, G. Giacomin was partially supported by the Swiss National Foundation project 20-41'925.94, the IHES and Rutgers University, while J. L. Lebowitz was partially supported by AFOSR grant 92-J0115, NSF grant DMR-95-23266 and the IHES and E. Presutti was partially supported by CEE grant CHRX. CT93-0411, the Courant Institute and Rutgers University. R. A. Carmona's contribution was supported in part by ONR grant # ONR N00014-91-1010. N.V. Krylov's work was partially supported by the NSF Grant DMS-9625483. Finally, B. Rozovskii and M. Mikulevicius were supported by ONR Grant N00014-95-1-0229 and ARO Grant DAAH04-95-1-0164.

RENE A. CARMONA

Statistics & Operations Research, C.E.O.R.
Princeton University
Princeton, N.J. 08544

BORIS ROZOVSII

Center for Applied Mathematics Sciences
University of Southern California
Los Angeles, CA 90089-1113

March, 1998

This page intentionally left blank

Part 1

STOCHASTIC MODELS

This page intentionally left blank

CHAPTER 1

Stochastic Partial Differential Equations: Selected Applications in Continuum Physics

JAMES GLIMM
SUNY at Stony Brook
Stony Brook, NY 11794-3600
glimm@ams.sunysb.edu

DAVID SHARP
The Los Alamos National Laboratory
Los Alamos, NM 87545
dhs@t13.lanl.gov

1. The Physical Basis of Stochastic Analysis

Stochastic modeling occupies a central place in the derivation of the equations of continuum physics. This fact is obscured by another fact: the desired outcome of this modeling is often a deterministic partial differential equation. Thus the stochastic model may be important either as an intermediate step towards a deterministic theory, or as a final theory in its own right. As we move toward a more realistic acceptance of probability as an essential component in the analysis of complex phenomena, the role of a stochastic model as a final outcome of the analysis will increase. This trend is illustrated, for example, by probabilistic weather forecasts [23, 58], probabilistic simulation studies to support petroleum reservoir management decisions [81, 31] and probabilistic financial decision making [102, 103]. As we demand increased accuracy and reliability from our deterministic models, the need to examine their stochastic components and derivations will also increase. For both reasons, we see a need to emphasize the importance of stochastic models of physical systems.

Stochasticity in models of physics arises from several sources. A principal factor leading to stochasticity is dynamical instability and sensitive dependence on initial conditions, in a microscopic dynamics which can be deterministic in principle. Closely related to this factor is the absence of sufficient deterministic data to specify a solution of the governing equations. In a probabilistic approach, such input data is supplied statistically, which leads to a statistical version of the final answer. The deterministic laws of physics cannot be ignored in this procedure, so that the model will be stochastic, and not purely statistical. Moreover, partial information cannot be ignored, so that the probability measures will be conditioned by partial knowledge. Finally, in some cases a complete deterministic model of a system may be lacking, owing to missing science.

This chapter begins with a discussion of the basic ideas underlying stochastic modeling in continuum physics, followed by a brief overview of selected areas where such modeling is playing an important role. In these examples, the governing equations are usually the partial differential equations of classical physics. We next review methods that are under development for stochastic analysis in this context. The application of these methods is illustrated in discussions of multi-phase flow in chaotic mixing layers and of the transport and dispersion for flow in porous media. The scientific issues raised in this chapter are ones which, in the authors' judgment, deserve deeper examination.

1.1. The Conceptual Framework. Stochastic phenomena, such as "noise", commonly arise in models of physical systems. Models are generally an incomplete description of nature, and noise represents some of the effects present in nature, but missing from the model.

Noise may refer to effects omitted from a model due to lack of data but often, and more fundamentally, noise refers to a loss of predictability. This typically occurs when the physical process being modeled is unstable, in which case a deterministic solution will depend sensitively on the problem data. Even minute variations in the data (such as the initial conditions, the physical parameters which occur in the defining equations, or even the functional form of the equations) can yield large variations in the solutions. In such cases deterministic prediction requires unreasonably large amounts of data, so that in a practical sense, there can never be enough data to predict the solution. Moreover, even if sufficient data were available, numerical solution methods with finite accuracy of time step integration and finite accuracy arithmetic do not generally produce stable, reliable solutions and even if they do, they cannot do so in an efficient, cost effective manner. Thus, viewed deterministically, a problem may be unsolvable, or highly intractable. It is in this sense that while lack of data may be seen as the requirement which forces the use of stochastic modeling, the more fundamental cause is lack of stability or sensitive dependence on data.

Stochastic modeling succeeds when the statistical formulations are predictive. This occurs when the mapping from stochastically specified data to a stochastically specified solution is well conditioned, so that it does not suffer from sensitive dependence on its data. It can very well be the case that the stochastic formulation is predictive, even when the deterministic formulation is not. However, in this case one has to settle for less detailed information, such as stochastic averages or moments, than would be present in the deterministic solution. For example, one might be able to predict the probability of rain tomorrow, rather than its certain occurrence.

Thus "noise" is not just an irritating detail, introduced to compensate for defects in a deterministic model, but is rather an intrinsic phenomenon for which a probabilistic description succeeds while a deterministic one fails. An example is the toss of a coin. A skilled performer can flip coins deterministically, and systematically produce heads or tails as desired [28]. But lacking such skill, concentration, and training, the common probability model is more satisfactory as a description of the event than would be a deterministic description based on forces imparted to the coin, the difference in height between the point where the coin is tossed and the point where it is caught, and its mass and moment of inertia.

For the purposes of the present discussion, we define our model by a partial differential equation, either in conservation form,

$$(1.1) \quad U_t + \nabla F(U) = 0 ,$$

or in conservation-diffusion form,

$$(1.2) \quad U_t + \nabla F(U) = \nabla \nu \nabla U .$$

Here $U = U(x, t) = U(x, t, \omega)$ is the solution. In the deterministic case the dependence on ω is missing, while in the stochastic case, $\omega \in \Omega$, with Ω a measure space. Then U is a random variable or, more properly, a random field, because it depends on multiple independent variables, x, t . U is random either because its data is random, or because of random terms in the equation itself. Thus the initial data, $U(x, t = 0)$ may be random, so that $U(x, t = 0) = U(x, t, \omega)|_{t=0}$, $\omega \in \Omega$, or the boundary data (not shown above) may be random. Randomness in the equation can occur if, for example, the flux function $F = F(U) = F(U, x, t, \omega)$, is random or if the diffusion coefficient ν is random. In view of the above discussion, we regard U as stochastic when it has sensitive dependence on the problem data.

We emphasize the use of random fields to represent noise in partial differential equations. A partial differential equation is an infinite system. Its data belongs to an infinite dimensional function space. The solution correspondingly possesses an infinite number of degrees of freedom. Potentially the noise does also.

There are situations in which only a few of these infinitely many degrees of freedom are unstable. In this case the deterministic system is infinite dimensional but the noise (at each instant of time) is finite dimensional. Such a system displays finite dimensional chaos. The stochastic description, if there is one, is given by a stochastic process, where the random function $\omega(t)$ takes its values in a finite dimensional space, such as the subspace of instability or chaos in the solutions of the partial differential equation. Simple models of atmospheric motion, obtained by truncating the Oberbeck - Boussinesq equations for fluid convection in a heated two-dimensional layer, lead to a set of low order ordinary differential equations, known as the Lorenz equations [71], whose solutions exhibit a strange attractor in the long time asymptotic limit. A measure on this strange attractor would be a probabilistic description of the chaotic, or deterministically unpredictable, solutions of the Lorenz system, or of this portion of the solution of the original partial differential equation.

Such a situation may be relevant to climatology. Climatology attempts to predict long term trends in the weather, using the same equations and computer simulation programs (known as the general circulation model) which are used to predict the weather. Current accuracy for weather predictions from these models is on the order of five days. In order to assess such questions as global warming or ozone depletion, predictions for years or hundreds of years are needed. Obviously the climatology predictions suffer from a sensitive dependence on data, and thus must be stochastic in nature. In fact, stated conclusions from these models, such as average annual temperature, are stochastic in nature, because of the large degree of averaging contained in this form of the answer. To some extent, there are or may be a moderate or small number of degrees of freedom in the atmosphere which largely govern such averaged predictions. For example the El Niño ocean current is believed to influence rainfall patterns over a large portion of the globe. In an atmospheric model (with no ocean coupling), El Niño would be modeled by stochastic boundary

data, and the resulting rainfall statistics would be a stochastic aspect of the solution, with both of these features modeled in finite dimensional subspaces of the data and solutions.

We are interested here in the opposite situation, in which there is a very large, or effectively infinite, number of unstable degrees of freedom which participate actively in determining the solution. Commonly, and as we assume here, the infinite number of degrees of freedom arise from a finite or possibly infinite number of degrees of freedom in each finite spatial domain. In other words, we assume an infinite domain for our partial differential equation, and spatially distributed and localized noise over this infinite domain. In this situation, both the noise and the solutions must be modeled as random fields. In some examples, such as infinite Reynolds number turbulence, or as an approximation to finite but large Reynolds number turbulence, we encounter an infinite number of unstable modes in a finite volume. Physically this simply means that more, or all, of the available problem degrees of freedom have become unstable. Fully developed turbulence is such a problem.

According to this discussion, we now assume that the measure space Ω is built on a function space, i.e. a space of functions of space or space and time, and that a measure $d\omega$ is defined on this space. In accordance with the probabilistic interpretation of the measure, we assume that the total measure of Ω is 1, so that $\int_{\Omega} d\omega = 1$. For each $\omega \in \Omega$, the corresponding (deterministic) functions, e.g., $x, t \rightarrow U(x, t, \omega)$ or $U, x, t \rightarrow F(U, x, t, \omega)$, are called a *realization* of the random fields U or F . It is logically correct, and conceptually clearer to suppress ω entirely, and just say that $U \in \Omega$, where Ω is a suitable function space having a distinguished probability measure. Having done this, we will replace Ω with an alternate notation also.

Using the measure $d\omega$, we define averaged field quantities $\bar{U} = \int_{\Omega} U d\omega$, and in a mathematically exact manner, derive an averaged equation, such as

$$(1.3) \quad \bar{U}_t + \nabla \bar{F}(U) = 0.$$

Because of the nonlinearity, $\bar{F}(U) \neq F(\bar{U})$. The averaged equation does not close, because it contains the unknown quantity $\bar{F}(U)$. Closure is the step in which \bar{F} is replaced by some known quantity. It is thus the process of defining the otherwise undefined averages of products or other nonlinear expressions in terms of known quantities, such as products (or some other function) of the averages. These relations are not derived mathematically as a deterministic identity, but rather reflect a statement concerning typical, or average, configurations of the stochastic degrees of freedom. In this sense, closure is very much akin to determination of the equation of state from the statistics of likely particle configurations in a gas. The equations which result from closure are called effective equations.

Effective equations are the final step in the stochastic modeling process. With these equations, we have gone from a deterministic but unstable equation, to a stochastic model and stochastic partial differential equation, to the solution of the stochastic aspects of this equation, and back to a new deterministic equation. The final equation, even if it has the same form as the original, has a very different interpretation, and generally very different physical parameters. Small length scales have been integrated out in the solution, and the equation now refers to quantities defined on a larger (macro) scale.

The simplest way to obtain a stochastic partial differential equation is to add a Wiener measure term $d\mu$ or white noise to a deterministic partial differential

equation. Starting with 1.1, we thus obtain

$$(1.4) \quad U_t + \nabla F(U) = d\mu .$$

Assuming that U approaches a constant at infinity, so that there is no net flux coming in from infinity, 1.1 implies that $\int_{R^n} U dx$ is time independent, i.e., conserved. Components of U represent quantities such as mass and momenta, which must be conserved in an isolated system, according to basic laws of physics. Additive noise, as in 1.4, models the incomplete isolation of a subsystem from some larger universe, and the coupling between the two typically represents the transfer of mass, momentum, etc. into the subsystem. While this type of random additive coupling is of interest, we emphasize the randomness introduced above from the problem data ($U(x, t = 0)$ and F), involving random fields rather than random processes, as being more typical of applications. Moreover, the correlation structure of the noise or data, which has a significant effect on the stochastic nature of the solution U , is not arbitrary. It is dictated by the application, and is generally not given by Wiener measure or white noise. Finally, we note that the random field is usually not Gaussian, or equivalent to Gaussian, for nonlinear problems. For these reasons, we do not consider 1.4 further.

1.2. Examples. We next discuss examples which illustrate, to varying degrees, the conceptual framework presented above. The examples are drawn from the physics of continuum systems, and are also selected for their current interest.

Brownian Motion and the Diffusion Equation. This example is familiar and highly developed. For this reason, our discussion will be brief. A small particle is suspended in a fluid and subject to random collisions with fluid particles. Each collision results in a random displacement of the particle, which thus executes the (mathematically defined model of) Brownian motion. In a continuum limit of infinitesimal waiting times between collisions and infinitesimal collision displacement steps, this process is the Wiener process. Upon averaging, or through consideration of expected values, the Wiener process yields the solution of the diffusion equation, a deterministic partial differential equation, which is the effective equation, in the terminology of §1.1.

The same modeling process applies to a number of other physical situations, such as heat conduction, and the dispersion of fluid flow through a heterogeneous medium.

Statistical Physics and the Equation of State. This is also a mature example. Molecules move according to Newton's laws, in response to intermolecular forces. Since the number of molecules, on the order of Avogadro's number, 10^{23} , is very large, the state space for this system is approximately infinite dimensional. (We leave it as an exercise to the mathematically minded to determine which finite numbers are approximately infinite and which are not.) Individual values of all the molecular coordinates and momenta are not available as initial data in a practical sense, and their evolution under Newton's equations depends sensitively on this data. Thus we need a stochastic description, which is given by the theories of statistical physics [59]. Various equilibrium ensembles correspond to different methods of conditioning the system, i.e., of specifying the allowed system states and the measure on the space of states. Averaging the equations of motion relative to suitable measures over this state space defines the moment equations of kinetic theory, such as the Boltzmann equation. A further asymptotic limit, in which the

probability distribution for the velocity is replaced by its low order moments, leads to the Euler and Navier-Stokes equations for fluid flow. The fluid equations must be supplemented with a statistically defined equation of state which, for example, expresses the total system energy as a function of other averaged variables, such as density and entropy. The determination of the functional form of the system energy is an example of the closure problem, to be discussed in more detail below. The solution to this problem is also determined in principle by a suitable averaging procedure, which leads to equations of state whose study is the subject of thermodynamics. This well developed example illustrates the statistical issues associated with closure, and provides guidance for the analysis of closure problems in other contexts.

Computational evaluation of the equation of state comes from a computational solution of the statistical physics problem. There are a number of open problems, both theoretical and computational, in connecting the statistical physics model to the thermodynamics, *i.e.*, the equation of state.

We have seen that stochastic modeling connects unstable microscopic deterministic equations (Newton's equations for a large number of particles) with macroscopic deterministic equations (e.g., the Euler equations). As discussed below, the macroscopic flow equations themselves may or may not be unstable and thus may, depending on the flow regime, require another level of stochastic modeling, to yield a new deterministic averaged equation.

It is a characteristic feature of the modeling of natural phenomena that distinct models are needed on distinct length scales, and that phenomena which are deterministic on one scale may be stochastic on another, to be averaged out. The equations, as one passes from smaller to larger length scales, may change their form, and even when they do not, the physical parameters within the equations may change their values. The integration of small length scales to obtain new equations which govern only the large length scale phenomena is called *renormalization*. In some cases, the integration steps to larger length scales can be carried out in a continuous, rather than discrete, fashion. In this case, the equations change continuously, and perhaps even differentiably, with the change of length scale l , leading to a differential equation for equations (or the parameters in them) of the symbolic form

$$(1.5) \quad \frac{d(\text{Equation})}{dl} + \beta(\text{Equation}) = 0 .$$

for some functional β . See Section 2.3 for further discussion.

Dislocations, Burger's Vectors and Constitutive Laws. Most solids are not perfect crystals. For example, common metals are polycrystals, made up of microscopic crystalline domains, joined along walls which represent defects in the crystalline structure. Additional crystal defects occur along dislocation lines, and as point defects, such as impurities, voids and inclusions. The statistical distribution of the defects and impurities is very important, as it determines the properties of the polycrystalline material. These properties are very different from those of the pure crystal. The line defects occur at the edge of an extra half plane of lattice sites inserted into the lattice. Their motion through the lattice is associated with plastic deformation. A continuum model of the line defects in a lattice is a curve, with a tangent known as a Burger's vector. When these are considered statistically and averaged, a density of Burger's vector fields (line defects) [64] results. A

further integration of the continuum level defect models leads to stress-strain laws, *i.e.*, constitutive laws for the solid, including plastic yield. These laws are just the equation of state (and possibly its non-equilibrium extension) encountered above for fluids, but now in the context of solid materials. Thus we see that material properties are defined by phenomena at several different length scales. In principle, behavior on one length scale is determined from a stochastic model of defects on a smaller one. But the practical solution of these stochastic models to give their coupling across length scales still presents many difficulties. In summary, we view the equations for the flow of solid materials, such as those describing elasto-plastic deformations, as resulting from the integration of stochastic models describing phenomena on smaller length scales, starting from the atomic scale. To fully implement this vision is a major challenge to contemporary science.

Material Damage Models, Shear Bands, Voids, and Fracture. Here we start with the continuum equations of, for example, an elasto-plastic material. These materials support several types of failure modes. Shear bands result from a thermal instability: Laminar shearing is unstable to highly localized shearing, since shearing induces heat, which softens the material, allowing a further (and localized) concentration of the shear stress. Voids and fracture result from a breaking of lattice bonds. The one sided damage to the material at a fracture tip results in a concentration of stress at that location, again giving rise to an instability. In some regimes, there is a profusion of micro-fractures, or of shear bands, or of voids. In such cases, these features must be treated stochastically. The deterministic continuum equations have effectively stochastic solutions. An averaging procedure will lump the combined effects of many micro-voids, bands, or fractures into a type of average material property for the partially damaged material. These averaged properties appear in averaged equations, and describe in a deterministic fashion the stochastic aspects of the ensemble of micro-voids, bands and fractures. However, there are again many difficulties – theoretical, practical, and computational – in implementing this program, and it is far from a completed body of science.

Multi-phase Flow. When two fluids mix at a molecular level, they behave as a single fluid, but with a modified equation of state which describes the single fluid resulting from the mixture. Multi-phase flow refers to any other type of mixing, *i.e.*, mixing which does not occur at the molecular level, but rather in terms of droplets, slugs, etc. of one fluid mixed into the other. At a microscopic level the flow is modeled by the Euler or Navier-Stokes equations applied to each fluid individually with an interface, moving with the local fluid velocity, to separate them. There are many distinct flow regimes in which this mixing can occur, each having distinct behavior. Spray or mist consists of small droplets of heavy fluid mixed into an ambient lighter fluid. A bubbly fluid describes the opposite situation of small bubbles of light fluid interspersed in an ambient heavy fluid. Slug flow describes a situation in which the droplets are larger, on the order of the length scale characterizing the flow itself; for example the pipe diameter, for flow in a pipe. Further complications arise from mass transfer between fluid components, as in the case of partial solubility of one component in another, or in the case of phase transitions. These flow configurations are unstable in the sense of §1.1, and the microscopic equations are not suitable for prediction. The microscopic equations are regarded as random, and after averaging [29, 30], a new set of deterministic equations are introduced. These equations describe average quantities such as the local average bubble or droplet concentration, considered as new system dependent

variables, and represented as functions of space and time. Multi-phase flow is important in such applications as fuel injection in diesel and automotive engines, pipeline flow, and the design of chemical reactors.

Turbulence. Turbulence is the phenomenon of instability and loss of predictability for flows described by the Navier-Stokes equation. It is characterized by flows with a large Reynolds number R , where $R = vL/\nu$ with v the velocity, L a typical system length scale, and ν the viscosity. At small length scales on the order of $l_{\text{visc}} = \nu/v$ the effective Reynolds number is 1, and the flow is laminar (*i.e.*, stable, and with a smooth dependence on flow data). This length scale is called the dissipation length. At this length scale, dissipation effects dominate, and energy is rapidly converted into heat. At length scales comparable to L , the flow is dominated by system dependent features, *e.g.*, large vortices in the flow pattern. Thus neither the small nor the large length scales should be considered to be stochastic.

Kolmogorov proposed that for length scales in the range $l_{\text{visc}} \ll l \ll L$, called the inertial range, turbulence is universal, *i.e.*, independent of the Reynolds number R , and isotropic. This is the range of length scales to which stochastic modeling applies. This fundamental hypothesis – that inertial range turbulence is independent of viscosity as $\nu \rightarrow 0$ – has been the accepted basis for all modeling of inertial range turbulence since its introduction in 1941. (See also §4.3.) Under this assumption, a dimensional analysis shows that the energy stored in turbulent vortices is a simple (power law) function of their size. Expressed, as is customary, in Fourier space, with the wave number $k = 2\pi/l$, Kolmogorov's law is

$$(1.6) \quad E(k) = c_K \epsilon^{2/3} k^{-5/3},$$

where ϵ is the energy dissipation rate and c_K is a dimensionless universal constant, known as the Kolmogorov constant. In this theory, c_K is independent of the Reynolds number of the flow and of its large scale driving forces. In the inertial range, one speaks of a cascade, in which energy is transferred from larger vortices to smaller ones, but never dissipated. Kolmogorov's theory, including the independence of c_K from the Reynolds number, is supported by experimental results [93] and numerical simulations [61]. (See also the discussion in §4.3.)

To achieve a stationary, or time independent, theory energy must be added at the same rate that it is removed. Thus turbulence theories often include stirring forces, to maintain a constant level of turbulent flow. According to the cascade picture, the stirring should occur at the large scales. Other types of stirring forces, which stir at wavelengths in the inertial range, may influence the answer, *i.e.*, the solution behavior within the inertial range.

Intermittency refers to a class of proposed corrections to the Kolmogorov theory [73, 74, 21]. One of these theories asserts that the energy cascade is modified so that a larger vortex does not drive a turbulent flow in smaller vortices within all of its spatial extent, but only in a certain part of it. In this sense, the observed turbulence will be patchy, or intermittent, on each length scale. If this idea is followed, and the Reynolds number R is taken to infinity, then the turbulence becomes concentrated on a Cantor set of lower Hausdorff dimension $d_{\text{turb}} \leq 3$. With less turbulence spatially at any given length scale, the amount of energy flowing down the energy cascade $E(k)$ must be reduced, and formulas relating corrections of the exponent 5/3 to the turbulence deficiency dimension $3 - d_{\text{turb}}$ have been proposed [34]. Recent very accurate direct numerical simulations [17] show a small correction to the Kolmogorov exponent.

An alternate view of intermittency is that the corrections to the Kolmogorov theory do not change the exponents associated with the energy cascade, but do affect the exponents of higher order moments. In this view, intermittency is a departure from Gaussian statistics which includes more frequent large deviations than Gaussian statistics would predict. Such theories are discussed in [74, 21]; recent developments and a survey of the literature can be found in [89, 88]. The computations of [17] confirm this non Gaussian view, and show larger corrections to the higher moments than would be inferred within a Gaussian theory from the corrections to the second moments.

Thus it appears that both types of corrections to the Kolmogorov theory are valid. Since observational and numerical simulation data require at most only a very small modification to the Kolmogorov exponent, and since this data does show definite evidence of non Gaussian behavior, the second of these two classes of theories of intermittency appears to be more significant, and to be better supported by the data.

An alternate approach to turbulence, and one more directly connected to the practical requirements of predictive engineering simulations, is to average the Navier-Stokes equations relative to a statistical ensemble. The resulting equations do not close, *i.e.*, they contain undefined expressions, as well as new terms, and perhaps new dependent variables. The choice of closure relations is a difficult aspect of this approach, because this choice turns out to depend on the particular flow regime being modeled. Since the equation of state has discontinuities across phase transition boundaries (*e.g.*, across the sublimation boundary, separating the gas from the solid phase), there is no logical problem with discontinuous behavior of a statistical ensemble. The problem is not the existence of discontinuous dependence, but our lack of understanding of the discontinuities, or knowledge of which discontinuities are fundamental and which can be attributed to an imperfect theory. Closure relations seem to change even between fairly similar flow situations, and they may contain too many parameters, which are then not effectively determined by the data. In summary, the determination of suitable closures is still an open problem. Since closure is in principle the substitution in the flow equations of an approximate statistical identity between mathematically distinct quantities, the validation of closure is the statistical test of the proposed identity and the determination of the flow regime for which the identity is approximately correct. A more frequently used validation of closure is to test whether the closed, averaged equations are themselves valid, for example by having solutions which agree with some class of data.

Porous Media. Flow in porous media has two rather different, and competing, stochastic aspects. In their simplest form, the flow equations are a coupled system. Darcy's law and incompressibility,

$$(1.7) \quad \vec{v} = -\frac{K}{\mu} \nabla P, \quad \nabla \cdot \vec{v} = 0,$$

where v is the total fluid velocity, μ is the fluid viscosity, and K is the total permeability of the porous media, form the first equation of the system. For two phase flow, such as oil and water, characterized by a relative volume fraction s of one of the phases, the Buckley-Leverett equation

$$(1.8) \quad \frac{\partial s}{\partial t} + \nabla \cdot \vec{v} f(s) = 0,$$

describes the transport of the substance or concentration with partial volume fraction s of one of the phases. These two equations are coupled nonlinearly through a dependence of

$$(1.9) \quad K/\mu = K_{\text{total}}/\mu_{\text{effective}} = K_{\text{absolute}} \left(\frac{k_{r,w}(s)}{\mu_w} + \frac{k_{r,o}(s)}{\mu_o} \right)$$

on the saturation s . Here K_{absolute} is the permeability defined in terms of rock properties alone, and is independent of the fluid properties, according to Darcy's law 1.7. The quantities $k_{r,w}(s)$ and $k_{r,o}(s)$ are corrections to the absolute permeability due to two phase interference effects, such as surface tension, and depend on the relative saturation s of the two fluids. The denominators μ on the right hand side are the pure phase viscosities, while $\mu_{\text{effective}}$ on the left is defined by 1.9. Each term, $k_{r,j}$, $j = w, o$, gives a relative permeability for flow in one phase only. The media are generally heterogeneous, so that K_{absolute} and k_r may each depend on the space variable \vec{x} . The nonlinear coupling and the heterogeneity are the two competing sources of stochasticity. The nonlinear coupling gives rise to a flow instability (viscous fingering) [56], and thus to a sensitive dependence on problem data, and a need for stochastic modeling. The heterogeneity is typically fairly strong, occurring on many length scales, from the rock pore spacing of a few microns to the inter-well spacing, which can be fractions of a kilometer. Only the variation on the largest length scales is reliably reconstructed from the limited data available, so that the very important effects of heterogeneity on the smaller length scales have to be treated stochastically, in terms of a probability model. As a result, the flow equations have random field coefficients, and must also have random field solutions. Thus predictions made with this model are inherently stochastic. Based on a probability model for the data, one can state that the solution has a certain probability of behaving in one manner or another. An early expression of this stochastic solution concept is given in [45].

Composite Materials. Requirements for high performance materials are often met by composites. As a simple example, consider the strength to weight problem, of obvious importance to aircraft and satellites, but important in other applications as well. The simple device of removing material (holes or voids in the structure) in lower stress regions can greatly improve the performance. Introduction of reinforcements, as with steel reinforced concrete, is another good example of a composite material. A similar, and more recently introduced, composite is the fiber reinforced polymer plastic composite. The fiber is first twisted into threads, which are woven into cloth, and then placed in layers in a mold (giving rise to a porous medium) into which the liquid plastic is injected in a manufacturing process known as Resin Transfer Molding (RTM). Methods to predict the mechanical strength, fluid flow properties, or electrical response of composites suffer from sensitive dependence on data. The microscopic equations are deterministic but cannot be integrated in a practical sense from the microscopic data, even if it could be obtained. Thus an averaging procedure is needed, to express the macroscopic response of the media to the average effects of small scale variations within it. As a typical problem of this type, consider the flow of liquid plastic resin into the fiber mat. This flow problem, which is an example of flow in a porous medium, is modeled by Darcy's law, 1.7. The derivation of Darcy's law, and the determination of the permeability K from microscopic data is an averaging problem, with the microscopic data specified in a statistical fashion only. The Navier-Stokes equations are

the microscopic equations. The averaging method used is known as homogenization [57]. It is totally deterministic, and assumes an infinite separation of length scales between the microscopic variation and the scale of measurement, on which the Darcy constant K is defined. Although Darcy's equation can be derived by these methods, the actual value predicted for K is not correct. Thus this approach is qualitatively but not quantitatively accurate. The deficiency lies in the deterministic and infinite scale separation assumptions of the homogenization method. Note also that the form of the averaged equation is different from the microscopic equation on which it is based.

2. Mathematical and Computational Tools for Stochastic Analysis

In this section we discuss some of the methods of a mathematical and computational nature which have proved useful for stochastic modeling in the context of continuum physics. We do not claim that the mathematical basis for these methods is well developed. On the contrary, the methods presented here are, for the most part, far less well developed mathematically than such methods as Ito integrals, martingales, or Wiener measure. In general, the lower level of development of the mathematical foundations is due to the greater difficulty presented by random fields as opposed to random processes. From a formal point of view, these two are related as partial differential equations are related to ordinary differential equations; alternately, we can say that the stochastic aspect of the problem converts a (deterministic) partial differential equation into such an equation in an infinite dimensional space.

2.1. Random Fields and Ensemble Averages. The difference between a field and a random field is that the former is a function defined on space-time, whereas the latter implies that the function space to which the function belongs comes with a probability measure defined on it. Thus we are interested in measures on infinite dimensional function spaces.

There is a large difference between measures on finite and on infinite dimensional spaces (such as a function space). There are also significant differences between measures on functions of one, two, three, ... independent variables. According to a general construction due to Kolmogorov, if there are measures, consistently defined, on the finite dimensional subspaces of the function space, then there is a measure space Ω and a measure $d\omega$ defined on it which has these finite dimensional restrictions. This measure space is abstractly defined, and in order to explain where it is actually supported we specialize, and temporarily consider Hilbert spaces and Gaussian measures only. The Kolmogorov measure is Gaussian if and only if each of its finite dimensional restrictions is Gaussian. It is then uniquely determined by its covariance, which is abstractly characterized as a positive definite bilinear form, defined as

$$(2.1) \quad \langle f, g \rangle = \int_{\Omega} f(\omega)g(\omega)d\omega .$$

Here the inner product on the left is taken in the Hilbert space to which f and g belong. Thus $f \in \mathcal{H}$, where \mathcal{H} is the Hilbert space defined by the left side of 2.1. On the right, f is the function (linear functional) defined on \mathcal{H} by the formula $h \rightarrow \langle f, h \rangle$, for $h \in \mathcal{H}$. Thus the covariance defines a Hilbert space \mathcal{H} . One might hope that the integration 2.1 would be identified with an integral over \mathcal{H} . Unfortunately, \mathcal{H} is a set of measure zero in the measure space Ω . But if we

consider larger Hilbert spaces $\mathcal{H}' \supset \mathcal{H}$ with the obvious embedding a continuous mapping, then \mathcal{H}' is a set of measure one in Ω if and only if the embedding is trace class. We assume this, and identify \mathcal{H}' with Ω [36].

A somewhat universal version of this construction is given by the integration theory over the dual of a nuclear space. Consider, for example, the Schwarz space \mathcal{S} of rapidly decreasing smooth functions. \mathcal{S} is the *nuclear space*

$$(2.2) \quad \mathcal{S} = \bigcap \mathcal{H}_j, -\infty < j < \infty$$

where $\mathcal{H}_{j+1} \subset \mathcal{H}_j \subset \mathcal{H}_{j-1}$ with a trace class embedding, and the dual \mathcal{S}' of \mathcal{S} is the dual of a nuclear space,

$$(2.3) \quad \mathcal{S}' = \bigcup \mathcal{H}_j.$$

Here $\mathcal{H}_0 = L_2(\mathbb{R}^n)$ and \mathcal{H}_j and \mathcal{H}_{-j} are in duality via the \mathcal{H}_0 inner product. In this framework, every Gaussian measure whose covariance is continuous in the \mathcal{S} topology is supported on the space \mathcal{S}' [36].

The most serious conclusion to be drawn from these facts is that the measure spaces are non-unique, and are intrinsically related to the problem under consideration. In contrast, for the case of measures over finite dimensional spaces, such as \mathbb{R}^n , the measures used in PDE applications are very often equivalent to Lebesgue measure.

Nonequivalence of Gaussian measures is the beginning of a series of technical complications associated with measures over function spaces. The completion, \mathcal{H}' , of the original Hilbert space \mathcal{H} may be, and commonly is, a space of distributions, i.e., generalized functions, with the ordinary (L_1) functions themselves being a subspace of measure zero. The nonlinear terms in the equations, such as the flux term $F(U)$, will not make sense in general for U a distribution, and physically motivated regularizing terms (which will lead to increased regularity of \mathcal{H} and thus of \mathcal{H}') will probably be needed to overcome this problem. This possibility is related to the idea expressed in §1.1 that only part of the complete function space of solutions need be considered from a stochastic point of view. If the behavior of the solution at short distances is deterministic, as occurs with Navier-Stokes turbulence, then integration over function spaces which are locally smooth will be sufficient to capture the randomness inherent in the problem. If the local regularity of \mathcal{H} is that of a Sobolev space with k derivatives, and the space or space-time functions are defined on \mathbb{R}^n , then \mathcal{H}' can be chosen to have $k - (n/2) - \epsilon$ derivatives locally, but cannot be more regular than this [24, 16].

The natural regularity space for the Navier-Stokes equation is the Sobolev space \mathcal{H}_1 of functions with one L_2 derivative because the dissipation, or decay of energy, leads to an *a priori* bound in this norm. For $n \geq 2$ space dimensions, this degree of regularity will not yield functions in the space \mathcal{H}' . However, local regularity of solutions (existence of strong solutions), while not proven mathematically, is a very reasonable conjecture on physical grounds. Assuming this, the solutions are known to be smooth (C^∞), and higher order Sobolev norms for the solution can be bounded in terms of the data. Thus we can work in function spaces \mathcal{H} for which the associated \mathcal{H}' has the regularity of at least a classical solution, and the elements of the function space are path-wise sufficiently regular to be interpreted as classical solutions of the Navier-Stokes equation. The probability interpretation of the \mathcal{H} covariance which defines this regularized measure requires a physical interpretation, which would be to suppress the stochastic aspects of the short length scale fluctuations. The basis

for such a choice of probability measure would be that the short length fluctuations are not stochastic, as these degrees of freedom are deterministic.

Gaussian measures arise naturally in the solution of linear problems. Nonlinear problems are of greater importance in most applications. Nonlinear interactions result in essential changes in an infinite number of degrees of freedom, and the study of stochastic nonlinear problems requires measure spaces which are inequivalent to the Gaussian measure space. Examples of this phenomenon are provided by the measures developed for the study of interacting quantum field theories, such as the $\lambda\phi^4$ theory, using the Feynman-Kac construction [41, 90]. The discussion of intermittency in terms of non Gaussian corrections to the higher moments suggests that the function space measure associated with turbulence will also not be equivalent to a Gaussian measure.

There are several ways to give a physical meaning to the averages taken with respect to the measure $d\omega$. One is to consider an ensemble of related problems, with statistically specified data. The measure on the data then translates naturally into a measure on the solution. Another is to consider a single solution, which may have rapid oscillations over time or space intervals. Then a time or space average (or both) is taken in a region near a (time or space) point, and the resulting average defines a new function, depending on the point about which the average was taken. Often these three distinct averages (time, space, and ensemble) give the same theory, and the averaging is said to be *ergodic*. If these averages do not agree, then it is necessary to examine the problem formulation, to decide how each of the averages should be interpreted. Here we will use the method of ensemble averages.

The use of function space integrals, defined over a space of initial conditions for the Navier-Stokes equations, as a description of the turbulent state, was proposed by Hopf, and has long been advocated by Foias. Chorin [21] proposes integration over coherent structures (vortex tubes) in the models of turbulence he considers. For two dimensional turbulence, described by solutions of the incompressible Euler equation, such a program has been worked out. A statistical measure on the (time averaged) mean vorticity is characterized by the principle of maximum entropy, subject to constraints associated with conserved integrals of motion [76, 85]. This state describes the large time, averaged flow in terms of a vorticity density ρ , which is a solution of a constrained optimization problem which has been solved numerically [101]. In many cases, the numerical solutions are consistent with the results of direct integration of the microscopic dynamics, but discrepancies are also reported, presumably associated with cases in which the approach to equilibrium is too slow to be resolved by the simulations of the microscopic dynamics.

2.2. Moment Expansions and Effective Equations. A moment of a measure defined on a linear space is the integral of a monomial. The order of the moment is the order of the monomial. Whether a measure is uniquely determined by its moments is a problem which has concerned mathematicians. Here our focus is different, to understand the information contained in the moments, and especially to construct differential equations which they must satisfy.

Elements of \mathcal{H}' can be identified with the linear functions on \mathcal{H}' itself. Thus their integrals are the moments of order one. Integrals of products of elements of \mathcal{H}' thus define the moments. Assuming the solution of the moment problem, a general function on \mathcal{H}' can be represented as a limit of sums of moments, or perhaps as an infinite series of moments. These series are the moment expansions.

We return to the averaged equation 1.3. We suppose that for each component U_i of U , $\bar{U}_i \in \mathcal{H}'$, $i = 1, \dots, n$. While this might not be the case, and while we actually require a stronger hypothesis, such as $\bar{U} \in \mathcal{S}$, which is still less likely, we proceed formally. Then \bar{U} is a first moment. Averaged equations for second moments can be derived from 1.1 or 1.2. From 1.1, we have the equation

$$(2.4) \quad \frac{1}{2} \frac{\partial \langle UU \rangle}{\partial t} + \langle U \nabla F \rangle = 0$$

for the $n(n - 1)/2$ distinct components of the second moment. This same identity gives equations for the second moment evaluated at distinct points, called the two point function. Higher moments and n point functions satisfy similar equations, obtained in the same manner. We suppose that F is a polynomial or limit of polynomials in U . Then the full set of moment equations closes, in the sense that flux terms which appear in the space derivative terms are known expressions of the quantities which enter in the time derivatives. Similarly, the full set of equations for the n -point functions close. These infinite systems are not very useful, and their solutions are more or less equivalent to the direct numerical evaluation of the ensemble averages discussed in §2.4. Usually it is desirable to stop the expansion with the first moments and some or all of the second moments. The closure problem is thus to express \bar{F} or $\langle U \nabla F \rangle$ in terms of the first moments \bar{U} and those components of the second moments $\langle U^2 \rangle$ retained from 2.4. The Reynolds stress, given by

$$(2.5) \quad R = \langle \rho v v \rangle - \frac{\langle \rho v \rangle \langle \rho v \rangle}{\rho}$$

for compressible fluid dynamics, gives the flux of momentum due to the transport of velocity fluctuations in the fluid. Its gradient enters in the averaged momentum equation, and must be modeled, *i.e.*, approximated, by some function of the first and second moments, to close the equations. Once closed, the equations are called *effective equations*.

2.3. Renormalization Group Methods. The *renormalization group* (RNG) is a procedure for integrating small length scales, so that the remaining unknowns only describe the large length scale aspects of the system. The purpose of the integration is to obtain an equation for the large length scales alone. In this sense most of the topics considered here fall under the category of RNG methods. Usually the term renormalization group applies when the system is self-similar, *i.e.*, displays essentially the same phenomenology on all length scales, or over a range of length scales. In this case the basic integration step is the same at each length scale and so, when applied repeatedly, obeys a multiplication law having a semi-group structure, hence the name renormalization group.

Assuming the group law and the system self-similarity it is based upon, we now decompose the elementary step into a two step procedure. The first step is an integration, or coarse-graining, step whose purpose is to eliminate the small scale problem information by averaging. The second step is rescaling, to change length scales in proportion to those eliminated, so that a nominally constant length scale is under study. If the integration step size is infinitesimal in length scale, then a differential equation rather than a group law is obtained for this operation. The differential equation is the infinitesimal generator of the group.

As in §2.2 with the derivation of effective equations, the RNG operations introduced up to now result in complex expressions which are no easier to evaluate

than is the solution of the original problem. We further use the assumption of self-similarity to seek a fixed point for the RNG equation . Since the system has the same behavior on all length scales, and the RNG integration step is the same for all length scales, the solution of this equation may also be the same, at least asymptotically, after many such integrations, so that the influence of the starting point can be neglected. If this is the case, the RNG and its equation have a fixed point, and the solution corresponding to this fixed point has scale invariance, or self-similarity. The fixed point will then have some degree of universality. It is certainly independent of the initial point (the detailed specification of micro-physics at the smallest length scale) for all initial conditions in the basin of attraction of the fixed point. In many cases of interest, the universality domain is large enough to be a significant help in the solution process.

The RNG was first applied by Gell-Mann and Low [38] and by Bogoliubov and collaborators [12] to the solution of problems in quantum field theory. Wilson [104] introduced practical solution methods for RNG problems, using the ideas of irrelevant variables and an ϵ expansion, as explained below. With these advances, he greatly extended the domain of applicability of this method, and obtained useful solutions for a large number of problems in statistical physics.

Applying the basic integration step, the form of the original equation will change in most cases. The modified equation will contain (a) terms having a similar form as before, but with modified numerical values for the system parameters and (b) new terms. As in §2.2 with the closure problem, the new terms in (b) will proliferate with each integration step. Based on the universality hypothesis, we divide terms into three types: relevant, neutral and irrelevant, using a dimensional analysis, to see whether their magnitude grows, is unchanged dimensionally, or decreases (a change of type (a)) with each integration step. Most of the terms are irrelevant, and these can be dropped without change of the answer, due to universality. The retained terms, finite in number, are now stable under the integration process, and the rest of the problem is to determine the modified coefficients, *i.e.*, changes of type (a). Thus we now assume that renormalization only produces changes in the equation parameters. Then 1.5 is the RNG equation for the change in these parameters. The relevant variables are those with dimension, and we assume that they are set phenomenologically, in terms of the macroscopic system properties. In other words, these parameters are data of the problem and of the integration process. Thus they are reset to the values given by problem data after every integration step. It remains to consider the neutral variables. These will be the dimensionless coupling constants of the problem. According to the sign of β in 1.5, they will either grow or decrease with increasing length scale. Denoting such an equation parameter as λ , the fixed point of 1.5 is given by $\beta(\lambda) = 0$.

The single coarse graining integration step should be contrasted to the ideas of homogenization. Homogenization is a two scale theory, and it assumes an infinite separation between these scales. In most versions, it also assumes that the micro-scale structure is periodic. For these two reasons, its successes are more qualitative than quantitative. It has successfully predicted the form of averaged equations, but it has been less successful in determining the numerical values of parameters which appear in these equations. See [57, 5].

It might appear that we have achieved a fundamental improvement over the homogenization method, which requires an infinite separation of length scales for its validity. This is not entirely correct. The formal justification of the RNG

method requires an assumption of weak coupling between adjacent length scales. Normally this is not the case, but for typical scale invariant problems, there is a unique dimension, called the critical dimension, with the required property. Thus the RNG method is in some sense exact in the critical dimension d_c . A further expansion in powers of $\epsilon = d - d_c$ then relates the RNG solution in the critical dimension to that in the dimension d of interest.

Of the efforts to apply this method to partial differential equations, we mention [14, 15, 50] and references cited therein. Related material emphasizing self-similarity in the solutions of partial differential equations can be found in [9].

We propose an alternate derivation of the RNG equation and the evaluation of its fixed points. To do this, we return to a pre-Wilson, Bogoliubov-type RNG formulation. The alternate derivation appears to be applicable to self similar non-equilibrium processes which are dynamically self-coarsening. We explain our approach in terms of the unstable micro-physical processes discussed in more detail in the following sections. For these problems, approximations depending on the distinctions between relevant and irrelevant terms are not needed, but approximations in terms of a closure hypothesis (for averaged equations) or other approximations, including the Boltzmann homogeneous chaos approximation (for a statistical n -body model) are used. Also appeal to critical dimensions and the ϵ expansion is not needed. Because of the assumption of dynamic self-coarsening, the group integration process has the meaning of a discrete time step, and the RNG parameter l in 1.5 is a rescaling of physical time. The coarse graining originates in the dynamic instability of the micro-physics. The self similarity expresses the fact that the coherent structures of the problem interact to merge and form larger structures in such a way that the dynamics acquires an independence of length and time scales. Details of this picture are provided in §3 for the Rayleigh-Taylor instability. It would appear to have application to Taylor-Saffman and Kelvin-Helmhotz instabilities as well.

2.4. Direct Numerical Simulations. Direct numerical simulation plays a very important role in the study of stochastic models. Simulation methods can be applied to the unstable micro-physics, to the evaluation of ensemble averages, and to the solution of averaged equations. This range of topics, and the variety of computational methods which have proven to be useful, is too broad to be fully summarized here. We restrict our discussion to selected subproblems.

Direct simulation of unstable micro-physics. We consider the simulation of the motion of an unstable fluid interface as a representative and important example of unstable micro-physics. For this problem, the interface will become exceedingly complex. The essential difficulty is that the instability is sensitive to details of the numerical algorithm. Typical discretization errors, such as numerical mass diffusion across the interface and numerical orientation effects of the interface with respect to the grid, mimic physical effects to which the problem, due to its inherent instability, is sensitive.

Front tracking is a numerical method which completely eliminates numerical mass diffusion across interfaces, and which greatly minimizes numerical grid orientation errors. See [40, 51, 54, 13] and the references cited therein. In this method solution discontinuities, such as interfaces between distinct fluid phases, are given distinct computational degrees of freedom, as discretized surfaces, moving by their own dynamics. The computed solution between the discontinuity surfaces is thus

conceptually smooth, with the tracked fronts representing discontinuity interfaces of the solution. The smooth part of a piecewise smooth solution is represented by a regular rectangular grid defined in all of space. There are thus two distinct grids, a stationary, volume filling grid, and a moving hyper-surface grid of co-dimension one. The interfaces described by the moving grid may self intersect along designated co-dimension two objects (curves in three dimensions and points in two dimensions), etc., so that the interface describes what is known as a non-manifold geometry, and in its discrete form, as a simplicial complex.

The computation on the regular grid, for the solution of hyperbolic conservation laws such as those for compressible fluid dynamics, uses higher order Godunov methods or other upwind methods [69, 91]. The motion of the interface is, in principle, a solution of a one dimensional problem with a simple jump discontinuity (*i.e.*, a *Riemann problem*). Thus Riemann solvers are used to advance the fronts. Coupling of the regular grid system to the fronts is given by nonlocal corrections to the elementary Riemann solution. The nonlocal data used in these corrections come from the regular volume filling grid. A grid stencil for grid points near the front is used in coupling a piecewise smooth solution on the regular volume filling grid to the front. The Godunov finite difference operator takes five neighboring solution values along a line as input to the determination of the solution values at the next time step. These are required to come from a single side of the front, so that the difference operators are never applied across discontinuities. Missing data, from the wrong side of the front, is supplied from data at the front itself. By this method, the solution state values defined at the location of the front grid system are coupled the the solution state values defined at interior grid cells.

This method introduced a significant improvement in the solutions for two fluid, *i.e.*, multi-phase simulations, improving specifically the validation and agreement with experiment for the Rayleigh-Taylor and Richtmyer-Meshkov instabilities. On physical grounds, the solution to these problems (*e.g.*, the instability growth rate) is known to depend sensitively on mass diffusion. Naturally, it also depends on the artifact of numerical diffusion present in most finite difference methods. The front tracking surface grid, which represents the discontinuity interface between the distinct fluids, is not constrained to be a rectangular grid, but is oriented with the interface itself. Thus the computation of its dynamics is largely free of mesh orientation effects. In contrast, finite difference methods, based on a regular grid, will bias the flow, and exhibit different flow properties for flow which is parallel to a grid direction as opposed to flow which is obliquely oriented relative to the grid. Other solution discontinuities, such as shock waves, are also computed by this method.

Local mesh refinement is the method of adding extra grid resolution non-uniformly, where needed to best achieve solution accuracy. This method has been proven to be extremely useful for a broad class of problems. There are several different methods for achieving locally refined grids [11, 97]. While to some extent front tracking and local mesh refinement are alternate methods for solving the same problems, we view them as complementary. Each method has distinct strengths, and ideally it would be advantageous to use both methods in the same computation.

Lagrangian computational methods are an alternative to tracking of fluid interfaces because, in the Lagrangian representation, the grid moves with the fluid particles and the fluid interfaces are automatically tracked. This method encounters difficulties for complex flow patterns, due to mesh tangling. To combine the

strengths of both Eulerian and Lagrangian computations, arbitrary Lagrangian-Eulerian (ALE) methods have been proposed, in which the grid motion is a local solution option. Since front tracking combines Eulerian grids (for the interior solution) with a Lagrangian grid (for the fronts which represent fluid interfaces), it can be regarded as belonging to the ALE family.

Interface reconstruction, developed as an extension of the SLIC algorithm [80] has been proposed to deal with the difficulties of computing fluid interfaces, as has the method of artificial compression [53, 70]. These methods each have strengths and weaknesses, which we will not attempt to survey here.

Also proposed is the level set method [79], in which the interface is a level set $\phi = 0$ of some auxiliary function ϕ , for which a simple transport equation is introduced, using the local fluid velocities. This method has two drawbacks. The first is that when the two fluids have the same equation of state, the level sets are decoupled from the computation, and can be regarded as a kind of graphics post processor. In this mode, the level sets do not add to the quality of the computation, and as graphics they are misleading, since the location of the $\phi = 0$ contour tends to drift from the optimal location at the center of the fluid mixing zone. The second drawback of this method occurs for fluids with different equations of state. In this case, the level set can be coupled into the computation, and used to determine which equation of state should be used for a mixed fluid cell block. This use of the level set method appears to have problems, in the form of artificial shock waves at the interface [40].

Because the fluids are mixed at the cell level (the level set method uses conventional finite differences, and does not avoid mass diffusion), the use of a discontinuous equation of state coming from the level set is inconsistent with the mass diffused interface in the computational representation of the fluid variables. In this sense, the level set method has a description of physics which is internally inconsistent, and does not actually solve the equation of state problems which it addresses. The fluid instabilities are sensitive to mass diffusion, and also to numerical mass diffusion. For this reason, very fine grids are needed to obtain numerical convergence for finite difference computations of unstable interfaces. Since the level set method does nothing to reduce mass diffusion, this feature of finite difference methods is inherited unchanged by the level set method.

Numerical evaluation of ensemble averages. The numerical analysis issues related to computation of ensembles are important, and not very well explored. For linear problems, whose solution is usually given by a Gaussian, the statistical solution can be obtained in closed form. This solution is also a possible starting place for perturbative solutions of problems which are nearly linear. However, many problems are strongly nonlinear, and require direct computation of the ensemble average.

The problem of constraints arises when it is necessary to evaluate the measure over a subset of the full probability space Ω . Often the space Ω and the measure $d\omega$ on it come from general considerations, while the integration subset Ω_0 defining a conditioned ensemble represents additional information from problem specific data. The easiest case occurs when $d\omega$ is Gaussian and Ω_0 is defined by linear restrictions on the space \mathcal{H}' associated with a single degree of freedom. Then the method of Kriging [22] applies. In the representation of geological data as a random field, the knowledge of point values or regionally averaged values of the field variables is data of this type.

Many constraints are nonlinear. For example, the measure $d\omega$ is usually defined in terms of the problem data. If this data is used to solve a nonlinear stochastic flow problem, then constraints on the solution (even if they are linear) define nonlinear constraints on the problem data. Given partial data of the solution, such as data at wells, for flow in porous media, the reconstruction of geology data which fits this constraint is called history matching. History matching is an inverse problem, to find the problem data in terms of (portions of) the solution. The method of solution is a Newton type iteration optimization, using multiple solutions of the forward, or direct, simulation. History matching is highly non-unique. Given flow problems with inherent instability or sensitive dependence on data, history matching will have only a limited time period of expected validity for predicting the future; *i.e.*, the portions of the solution not used in the history match. A systematic error in both Kriging and history matching is that they are biased to select atypically smooth realizations of the problem data. This selection introduces a systematic error in the forward prediction based on this match. For example, Kriging selects the maximum likelihood geology, subject to constraints. To understand the influence of this criterion, we note that it will select the data with no random variation in the absence of any constraints at all. Such smooth data gives fluid flow solutions which are very far from maximum likelihood. The smooth data suppresses highly conductive flow channels, and predicts an overly high recovery rate, and thus introduces a systematic bias in the statistics of the problem solution. In other words, maximum likelihood geology does not yield maximum likelihood fluid behavior.

A more fundamental approach, and one which is gaining favor, is to use full ensemble averages, and to compute enough solutions of the direct problem to evaluate the statistics of the stochastic partial differential equation directly. In this method, the constraints (other than the linear constraints of Kriging, which can be used to modify the ensemble directly) enter as a filter in the statistics. Only the direct solutions which agree (within error) with the data are used in the statistical prediction. This method is very computationally intensive, and it places a premium on fast algorithms for the solution of the direct problem. Trade offs in accuracy between the quality of the solution of the ensemble in terms of number of realizations, and the quality of the individual solution of each direct integration of the flow equations deserve further study.

3. Multi-phase Flow

Consider two fluids separated by an interface which is unstable relative to the flow pattern. As time progresses, the interface will become highly convoluted, and portions of each fluid will become entrained in the other, giving rise to a two-fluid mixing layer. The flow in the mixing layer has sensitive dependence on data, and thus requires stochastic modeling, even though the microscopic equations are deterministic. Shear layers (for which the tangential component of the velocity is discontinuous) are unstable according to the Kelvin-Helmholtz instability, leading to vortex formation and roll-up of the interface. Jets are subject to breakup and atomization in a variety of regimes, governed by the jet velocity relative to the ambient medium, among other factors [4, 99]. Acceleration driven interfaces may also be unstable, if the fluids have different densities, so that the forces of acceleration on the two fluids are not equal. We consider here two cases of acceleration driven instabilities, the Rayleigh-Taylor (RT) instability associated with a continuous and

perhaps constant acceleration, such as gravity [86], and the Richtmyer-Meshkov (RM) instability caused by an impulsive acceleration, such as by a shock wave.

The quantities of interest are the size of the mixing region, as a function of time, and the statistical composition of the mixture to give, for example, the volume or mass average of each fluid, velocities, and mixing length scales. We use two theoretical methods to extract this information: a statistical analysis of elementary modes and closure of averaged equations with a statistical validation of the closure hypothesis. We also use direct numerical simulation to gather data to test these theories, and we analyze laboratory data for the same reason.

For RT and RM mixing, the elementary fluid instability modes are inter-penetrating fingers, known as bubbles (for the light fluid) and spikes (for the heavy fluid). The bubbles and spikes each move toward the opposite fluid (in the RT case, due to gravitational forces). In an ideal RT flow, with constant acceleration forcing, the bubble and (for $A < 1$) the spike terminal velocity would be constant, due to a balance between buoyancy and form drag forces, with a value proportional to the square root of the radius. In more complex flows typical of chaotic RT and RM mixing, a competition process allows larger (and thus faster) bubbles to displace the slower ones, and thus expand in radius, so that the velocity can be further increased. The RT result is a constant acceleration of the mixing zone size, and a self-similar dynamic coarse graining of solution features, rather than a constant terminal velocity and constant feature size. With h_b (h_s) denoting the height (also called amplitude) of the mixing zone, measured from its initial position to the edge of the bubbles (spikes), we have the RT acceleration law

$$(3.1) \quad h_b = \alpha_b A g t^2, \quad h_s = \alpha_s A g t^2,$$

with $\alpha_b = 0.06$ an approximately universal value for incompressible RT mixing, as was first demonstrated experimentally by Read and Youngs [84, 108].

Our approach to multiscale stochastic problems (such as the RT and RM mixing zone problems) is based on the following steps:

- Identification and analysis of nonlinear elementary modes (vortices, bubbles, dendrites, viscous fingers, lattice dislocations, etc.)
- Analysis of pairwise interactions of elementary nodes
- Statistical analysis of an ensemble of modes governed by pairwise interactions
- Continuum level constitutive laws and macroscopically averaged flow equations.

These steps will be recognized as the program of statistical mechanics. Our objective is to develop a statistical mechanics of coherent structures, which are common features of nonlinear partial differential equations.

We now explain these ideas as they apply to the RT (and to some extent the RM) instability and mixing zone. For this problem, the dimensionless Atwood number $A = (\rho_2 - \rho_1)/(\rho_2 + \rho_1)$, with ρ_k the density in fluid k , plays the role of a buoyancy renormalization of the gravitational acceleration g .

3.1. Elementary Modes. For the RT problem, the elementary modes are the inter-penetrating bubbles of light fluid and spikes of heavy fluid. At small amplitudes, a linear theory (first order approximation in the amplitude of the interface from the unstable planar equilibrium) gives a complete, closed form solution, in

terms of exponentials and Fourier modes [18, 75, 35]. A recent very general treatment [106] of the linear regime allows multiple layers, general equations of state, compressibility, and general temperature or entropy stratification for the unperturbed solution. For a single mode perturbation, theory, simulation and experiment predict a constant terminal velocity v_∞ ,

$$(3.2) \quad v_\infty = C(Agr)^{1/2} ,$$

with r the radius of the bubble tip and C taking the approximate value 0.32 for bubbles in two dimensions and 0.48 in three dimensions. For $A = 1$, an approximate potential theory solution of Layzer [68] gives the velocity of the bubble throughout its entire evolution. Zhang [111] proposed a simple ODE model which extended this solution to spikes, to general values of A and to compressible flow. The model contains four parameters, of which two are set from the linear theory growth rate σ and the terminal velocity v_∞ . The other two, which give the effective gravitational acceleration g_R during the period of free fall and the decay rate b to the terminal velocity v_∞ , are set by careful comparison to front tracking numerical simulations. Zhang's model is the differential equation

$$(3.3) \quad \frac{dv(t)}{dt} = \sigma v \left(\frac{1 - v/v_\infty}{1 + \psi(v/v_\infty) + \phi(v/v_\infty)^2} \right)$$

$$(3.4) \quad \frac{dh(t)}{dt} = v(t)$$

for the bubble or spike velocity $v(t)$ and height $h(t)$. Here ψ and ϕ are expressed in terms of σ , g_R , v_∞ , and b [111]. In the limit $A \rightarrow 1$, $v_\infty \rightarrow \infty$, the expressions for ψ and ϕ , in the case of a spike, reduce to $\psi = \sigma v_\infty / g_R$ and $\phi = -\psi$, so that

$$(3.5) \quad \frac{dv(t)}{dt} = \frac{\sigma v}{1 + \sigma v/g_R} , \quad A = 1 ,$$

and for large t , $dv(t)/dt = g_R$ as required.

The single mode itself depends on two dimensionless parameters, the Atwood number A , and a dimensionless compressibility $M^2 = \lambda g/c^2$, where λ is the single mode wavelength and c^2 is the sound speed in the heavy gas. Here c^2/g is the length over which gravitational forces in hydrostatic equilibrium induce density changes, and M^2 is the ratio of the wavelength to this density variation length scale. Thus M^2 characterizes a dimensionless relation between the equation and its (single mode) initial data. Zhang [111] analyzed the dependence of the model parameters σ , v_∞ , g_R , and b of 3.3 on A and M^2 , for both bubbles and spikes.

For the RM case, the results are nearly as complete. As we will explain, the theory, numerical simulation and experiment are in substantial quantitative agreement. We emphasize especially the simulations [55], which for the first time achieved agreement with the experimentally measured growth rate for a RM instability. Depending on properties of the wave interactions of the idealized one dimensional shock-contact Riemann problem, which are themselves largely determined by the sign of the jump in the acoustic impedance $\gamma\rho$ of the two fluids, the incident shock wave of the RM problem will reflect at the fluid interface either as a shock wave or a rarefaction wave. The original linear (small amplitude) theory of Richtmyer was extended [107] to include the second of these two basic cases. Richtmyer's work only considered the first of these possibilities, and for that case, only three isolated parameter values, each for a γ -law gas, in the case of equal γ 's.

Richtmyer also introduced a very simple *impulse* model, based on the exact formula for the linear solution of the incompressible RT problem, with an impulsive (δ function in t) time dependent acceleration. This model predicts a constant velocity, i.e., a constant growth rate, with a magnitude given by an explicit formula. The model is thus very easy to apply. The impulsive model is widely used as a replacement for the linear theory or as a simple method for evaluating its asymptotic long time dynamics. There are three problems with the use of the impulse model. First, the impulse model may fail to give the long time asymptotics of the linear theory correctly, especially when the shock amplitude is large or the equations of state of the two fluids are dissimilar. The study [107] was the first systematic comparison of the linear theory with the impulse model. The discrepancies between these two models, as reported in [107], were not previously known.

Secondly, and more importantly, the long time asymptotics of the linear theory are unrelated to experimental measurements. In fact, the linear theory is intrinsically a small amplitude theory of a growing instability and its validity is thus limited to small times. The nonlinear growth rate for a single mode is defined as the difference between the fluid velocities at the two extreme tips of the single mode perturbation. It is rapidly rising initially, reaches a peak, and then decays. The linear theory agrees with the nonlinear solutions up to this first peak, for initial amplitudes typical of experiments. In favorable cases, the impulse model gives an approximately correct value for this first peak, before the decay in growth rates begins. Not all cases are favorable in this sense. In some cases, the impulse model gives growth rates lying well below the initial peak. However, experimental periods of measurement are typically later than the first peak, and lie within the decay period. Thus the impulse model has no reliable relation to the experimental numbers, and its widespread use has lead to some confusion, including the opinion, popular until a few years ago, that the experiments were “wrong”.

Thirdly, the long time asymptotics of the nonlinear growth rate is decay, given phenomenologically as a power law. This behavior is qualitatively different from that of both the impulse model and the linear theory, which have constant growth as their long time asymptotics. Thus the impulse model and the linear theory are both qualitatively inaccurate, as far as long term nonlinear asymptotics are concerned. The physical basis for the decay appears to be a nonlinear coupling of the fundamental mode to higher harmonics resulting from the large perturbation amplitude. The RM experiments all have sufficiently large initial amplitude to grow into the nonlinear regime and many of the experiments are initialized at $t = 0$ with an amplitude of sufficient size that the linear theory is inapplicable for all $t \geq 0$. The oscillations about the mean decay are associated with interactions of the interface bubble and spike tips with the transverse acoustic modes. In [2] the decay rates for the velocity of RM bubbles and of $A < 1$ RM spikes is given as $\text{const } t^{-\theta}$ with $\theta = 1$; the $A = 1$ spikes have constant asymptotic velocities. Holmes *et al* [55] report $\theta = 0.6$ for spikes as a fit to simulations of the air-SF₆ experiments of [10] and Holmes (unpublished) reports $\theta = 1$ for bubbles and spikes as a fit to simulations of the Kr-Xe experiments of S. Zaytsev (unpublished). Further study of these decay exponents would be desirable.

Higher order primitive perturbation series for the amplitudes have not been very successful for the RT problem, in extending the domain of validity of the linear theory. In contrast, perturbation theory to third order, re-summed as a Padé approximation, has given identical results to those obtained from direct numerical

simulation in the RM case [116], over a time period large enough to include the experimentally observed interval. The success of this theory can be attributed to two factors. First, Padé approximants often have a larger domain of convergence (and hence validity) than the primitive perturbation approximants upon which they are based. The second factor contributing to this success is that the RM amplitudes of interest, while fully nonlinear, are typically smaller than RT amplitudes, since the RM instability is less strongly forced. To simplify the evaluation of the higher order perturbation terms, the post shock growth rates are modeled as incompressible and the perturbation terms are computed to low order. In view of the Padé approximation, the growth rate is given as a rational function of the lowest order term; the linear theory is substituted for the impulse model velocity in this rational expression [116]. Here we are matching two distinct perturbation expansions. The inner expansion is the linear approximation to the small amplitude compressible problem and the outer expansion is a Padé re-summation of the large amplitude (nonlinear) incompressible expansion. The result is an analytic expression for the amplitude and growth rates of a single mode, valid up to moderately large times. These results have been extended to obtain separate growth rates for bubble and spike, for both the reflected shock and the reflected rarefaction cases, and for three as well as two dimensions, with excellent agreement with full numerical simulations (front tracking and other numerical methods) in cases for which this data is available [116, 117, 121, 120, 119, 118].

A discrepancy remains between the perturbation amplitudes (rather than growth rates) as measured experimentally, at the beginning of the period of experimental observation, and as predicted by simulation and theory (these two agree with each other). The cause of this discrepancy is believed to be a deficiency in the physical model, specifically in the neglect of material strength of the membrane initially placed between the two gases to suppress diffusion of one gas into the other. Partial confirmation of this belief is based on the following experiment [98]: a wire mesh was placed behind the membrane to reinforce it. The result retarded the instability amplitudes, but the same mesh, placed in front of the membrane, so that it had no strengthening effect but an identical effect on the flow, had no influence on the growth rates.

3.2. Mode-Mode Interactions. There are two key manifestations of the RT mode-mode interactions which drive the mixing behavior. The first is that the bubbles are unequal in radius and height; the most advanced (and usually the larger) bubble will acquire an addition to its speed from its advanced position. The second aspect of the mode-mode interactions is an extension of the first. The retarded (usually the smaller) bubbles will fall further and further behind the advancing front of bubble tips at the edge of the mixing layer. The gaps left by these smaller bubbles will be taken up by expansion of the advanced bubbles remaining at the outer envelope of the mixing layer. This phenomenon leads to a self-coarsening of the modes, and to their steady acceleration.

To explain this phenomenology in somewhat more detail, consider the edge of the mixing layer to be defined by an envelope drawn through successive bubble tips located there [42]. The bubbles are not uniform in their position, so the envelope has its own peaks and valleys, which we regard as collective excitations of the ensemble of bubbles. The peaks of this envelope, *i.e.*, of the collective excitation, are

regarded as bubbles of the envelope, while the valleys are spikes in the collective excitation envelope. With this picture, we apply RT instability analysis to the bubble envelope. Thus we understand the extra velocity imparted to the more advanced bubbles: they lie on the bubbles of the envelope, while the retarded bubbles lie on the spikes of the envelope. Thus each bubble receives a velocity contribution from the collective modes as well as one from its own single mode dynamics. The collective mode velocity contribution is positive for the advanced bubbles, and negative for the retarded ones. Although the individual modes (bubble and spike) are highly nonlinear, the above picture can be made quantitative by postulating that the interaction between the single mode and the collective excitation is linear, so that the effects add. Then

$$(3.6) \quad v_{\text{total}} = v_{\text{bubble}} + v_{\text{envelope}},$$

where the left hand side is the velocity of each bubble and the right hand side is the sum of the separate single bubble and envelope contributions [42]. Each of the two terms on the right is given by the single bubble theory, either bubbles proper (first term) or envelope excitations regarded as bubbles of the envelope. To apply this formula, we regard each bubble as being at its terminal velocity, so that $v_{\text{bubble}} = v_{\infty} = C(Agr)^{1/2}$, according to 3.2. To determine v_{envelope} , 3.3 – 3.4 are solved as a system. The envelope amplitude used in this formula is one half of the height difference between the adjacent bubbles and the envelope radius used implicitly to set the coefficients in 3.3 is the sum of the two adjacent bubble radii.

Formula 3.6 has been validated [42] by comparison to front tracking simulations and to experimental data. While approximate, it is sufficiently accurate for use in the statistical models developed below. For use in these models, we say that bubbles whose total velocity v_{total} has changed sign to become negative, and which are thus rapidly removed from the ensemble of bubbles located at the bubble envelope, can be removed from the ensemble. The geometric space which they formerly occupied is used by the advanced neighbors of these bubbles, by expansion. This process of elimination of the slow and expansion of the advanced bubbles is called merger. To avoid confusion, we note that it is the space on the bubble envelope occupied by the bubbles which is merged, not the bubbles themselves.

3.3. Statistics of the Ensemble of Modes. We develop here a statistical model for the envelope formed by the advancing bubble tips at the edge of the mixing layer. We follow the modeling of [48], with some change in the modeling assumptions and also with a more detailed derivation of the fundamental equations and conclusions. We also correct an error in the formula 4 of that reference. Our solution of this model yields a prediction, with no adjustable parameters, of the growth rate of the mixing layer, *i.e.*, the value of the constant α_b . There is an additional need for statistical modeling within the interior of the mixing zone, an issue we address in §3.4.

Following [87], we consider a piecewise constant interface, representing the bubble envelope as one edge of the RT mixing zone. Thus the fluid state is simplified and approximated in two space dimensions by a horizontal line segment drawn through each bubble tip. The length of the segment is the bubble diameter. The velocity of each bubble and conditions for bubble merger define the dynamics of the interface. According to [48], the total velocity of each bubble is determined by the superposition law 3.6 [42] given above, and the merger criteria is that a bubble

is eliminated from the interface when its velocity thus defined changes sign. Upon elimination, the neighbors grow in size, to occupy its space on the interface, while preserving their (more advanced) heights. Thus if bubbles of radii r' and r'' and heights h' and h'' merge, the result is a new bubble of radius r and height h , where

$$(3.7) \quad r = r' + r'' , \quad h = \max\{h', h''\} .$$

A statistical ensemble of such interfaces is then constructed, with the dynamics given by 3.3 – 3.4, 3.6 and 3.7 for each realization. Rather than solve this statistical physical problem directly, as was first done numerically [35], we first simplify it by a series of approximations.

In the Boltzmann approximation, the n -particle dynamics is replaced by a one-particle probability distribution, *i.e.*, by an ensemble of bubbles, characterized by a time dependent probability distribution of heights and radii. In the Boltzmann approximation, the assumption of homogeneous chaos means that the interactions occur with an interaction partner selected at random from the ensemble of bubbles, rather than with the geometrically adjacent neighboring bubble. At this point, the model is defined by a partial differential equation for the probability density of bubbles as a function of radius and height. Next, the single bubble probability distributions are further simplified by giving them in parametric form, so that the dynamics occurs in the finite dimensional space of parameters. Thus the dynamics is now a low order ODE. It is convenient to scale out the constant growth of length scale, and work in scaled coordinates (related to the unscaled coordinates by growth at a constant acceleration rate) in which this part of the dynamics is removed. There remain three equations, for the bubble radius and the mean and the variance of the bubble height.

The resulting dynamics is exactly the RNG dynamics for the bubble envelope of the mixing region as explained in §2.3. Recall that the RNG is defined by a two step process, in which the small lengths of the problem are removed (integrated, or solved using the equations of the problem), and then the problem is rescaled, so that a constant length scale is examined at each step. Each integration and rescaling step defines the elementary group operation. If the steps are infinitesimal in size, then the group generator is a differential operator. In its action on the parameters defining the ensemble, it is the RNG equation.

Note that the RT mixing process involves a steady coarsening of length scales. With the rescaling introduced above, the active length scales are fixed. Also the time parameter coincides with the RNG equation parameter. We believe these features apply more generally to a large class of self coarsening non-equilibrium thermodynamic processes. When the process is self-similar, as appears to be the case with RT mixing, we expect a fixed point for the RNG dynamics, and possible universal behavior.

As in [48], let t_m denote the time to merger for two adjacent bubbles of height h_1 and h_2 , and radii (assumed equal, for simplicity) r . Then t_m is the fundamental quantity which controls the statistics of the merger process. According to 3.2 for the bubble component of 3.6 and 3.3 – 3.4 for the envelope component, t_m is a function of the height difference $h_2 - h_1$, which is the variable h of 3.4, and is defined as the time at which the envelope velocity is equal in magnitude to the (terminal) single bubble velocity, so that the more retarded bubble will have a zero total velocity, and be removed from the ensemble. The argument $h_2 - h_1$ fixes the initial amplitude, and hence the velocity for the solution of 3.3 – 3.4. Thus $t_m(0) = \infty$, and if h_m is

the height separation required for merger, $t_m(h_m) = 0$. It is convenient to define $t_m(h) = \infty$ for $h < 0$.

Let $\bar{h} = \langle h \rangle$ denote the average bubble height in the ensemble. In the scaled variables, $h' = (h - \bar{h})/r$, $h_m' = h_m/r$, $v' = (Agr)^{-1/2}v$, $dt' = (Ag/r)^{1/2}dt$, and $t_m' = (Ag/r)^{1/2}t_m$, the equations lose their r dependence. Note that the bubble radius r is held constant during a bubble merger, and only increases at the end of each individual merger, allowing dt and t_m both to scale as a power of r . Thus the ensemble averaged radius \bar{r} satisfies the equation

$$(3.8) \quad \frac{d\bar{r}}{dt} = \left\langle \frac{1}{t_m'} \right\rangle \bar{r},$$

stating that the average bubble radius increases at a rate equal to the mean merger rate, which is equal to the mean inverse merger time. From the experimental and simulation data, we note that the radii are clustered about a common size at any given time. For this reason, we assume that \bar{r} can be used in place of r in scaling relations, yielding the scaled equation

$$(3.9) \quad \frac{d\bar{r}}{dt} = (Agr)^{1/2} \left\langle \frac{1}{t_m'} \right\rangle$$

for the mean bubble radius. At a RNG fixed point, the scaled rate constant $\langle 1/t_m' \rangle$ is independent of time, so that 3.9 can be integrated to yield the large t asymptotic formula

$$(3.10) \quad \bar{r}(t) = \frac{Agt^2}{4} \left\langle \frac{1}{t_m'} \right\rangle_*^2,$$

where the subscript $*$ on the expectation indicates evaluation at the RNG fixed point.

We next seek an equation for \bar{h} . Using the dt scaling relation and 3.10, we have

$$(3.11) \quad \frac{d\bar{h}}{dt} = \frac{dt'}{dt} \frac{d\bar{h}}{dt'} = (Agr)^{1/2} \frac{1}{\bar{r}} \frac{d\bar{h}}{dt'} = \frac{Agt}{2} \left\langle \frac{1}{t_m'} \right\rangle_* \left(\frac{1}{\bar{r}} \frac{d\bar{h}}{dt'} \right)_*,$$

and so

$$(3.12) \quad \bar{h} = \frac{Agt^2}{4} \left\langle \frac{1}{t_m'} \right\rangle_* \left(\frac{1}{\bar{r}} \frac{d\bar{h}}{dt'} \right)_*.$$

The maximum bubble height, h_b , of 3.1 is the maximum h allowed by the bubble height probability distribution at the RNG fixed point. We will see below that this quantity is given by

$$h_b = \bar{h} + h_m/2 = \bar{h} + \bar{r}h_m'/2,$$

so that

$$(3.13) \quad h_b = \frac{Agt^2}{4} \left(\left(\frac{1}{\bar{r}} \frac{d\bar{h}}{dt'} \right)_* \left\langle \frac{1}{t_m'} \right\rangle_* + \frac{h_m'}{2} \left\langle \frac{1}{t_m'} \right\rangle_*^2 \right).$$

To evaluate this expression, we note that

$$\frac{1}{\bar{r}} \frac{d\bar{h}}{dt'} = (Agr)^{-1/2} \frac{d\bar{h}}{dt}$$

is a scaled version of the mean velocity $d\bar{h}/dt$, according to the velocity scaling specified above. Thus, by 3.2 and 3.6, and with C_b denoting the constant from 3.2

for the bubble case,

$$(3.14) \quad \left(\frac{1}{\bar{r}} \frac{d\bar{h}}{dt'} \right)_* = C_b + \frac{h_m'}{2} \left\langle \frac{1}{t_m'} \right\rangle_* .$$

Here the first term represents the uniform drift in the bubble interface from the single bubble terminal velocity, and the second term is the combined contribution from the envelope velocity and from bubble merger. To understand the second term, consider adjacent bubbles, separated by the height difference δh . According to the definition of h_m as the height separation for merger to occur, we must have $|\delta h| \leq h_m$. Up to the time of merger, the envelope velocity will advance the leading bubble by $(h_m - \delta h)/2$. The retarded bubble will be moved further back by the envelope velocity, but this bubble will be eliminated by merger, and its position on the bubble interface replaced by the advanced bubble. Thus this portion of the bubble interface will move the distance $(h_m + \delta h)/2$. In effect, after merger, both portions of the the interface experience an additional positive envelope velocity contribution, with an average height change of $h_m/2$. Alternately, we state that the envelope velocity has no effect on \bar{h} , as it moves the advanced and retarded bubbles in opposite directions, with velocities of equal magnitude. Here we assume, as an approximation, a symmetry between bubble and spike dynamics. Then the full effect of the dynamics on \bar{h} comes from merger. The interface occupied by the retarded bubble moves by h_m and the other portion, occupied by the advanced bubble, does not move at all during merger. This height change is multiplied by the mean merger rate to obtain the combined envelope velocity and merger contribution. Combining 3.1, 3.13 and 3.14, we obtain

$$(3.15) \quad \alpha_b = \frac{1}{4} \left(C_b \left\langle \frac{1}{t_m'} \right\rangle_* + h_m' \left\langle \frac{1}{t_m'} \right\rangle_*^2 \right) .$$

The formula is not identical with that of [48], due to some changes in the modeling assumptions which appear to be more reasonable from a theoretical point of view.

It remains to establish the existence of the fixed point. For this purpose we examine the variance of the bubble height distribution. Merger and the bubble envelope velocity are the two main phenomena affecting the variance. The bubble envelope velocity is an instability of the $h = 0$ equilibrium RNG fixed point. It causes height variations to increase. Merger has the opposite effect. If the variance is too large, bubble merger will occur, thereby reducing the variance. The competition between these phenomena drives the system into an equilibrium, which is the RNG fixed point. Stated more mathematically, we see that the dynamics, in the scaled variables, induces a flow for the scaled height variance σ' , where $\sigma'^2 = \langle h'^2 \rangle / 2$.

The single bubble velocity makes no contribution to the variance, as it gives a uniform translation to the heights, within the assumption that the distribution of radii is very narrowly peaked. Thus the equation for σ' involves only contributions from the envelope velocity and from bubble merger. We now introduce the modeling assumption that the scaled bubble height distribution is Lebesgue measure on an interval $[-a, a]$. A direct calculation gives $a = \sqrt{3}\sigma'$. For $a < h_m'/2$, there is no merger, and a and σ' both increase in proportion to the envelope velocity. For $a > h_m'/2$, there is instantaneous merger with finite probability, and an infinite negative derivative for da/dt . Thus the fixed point occurs at the upper end of the allowed a interval, $a = h_m'/2$.

Following [110], this fixed point can be evaluated numerically, giving the approximate value $\alpha_b = 0.045$ for incompressible flow. This number is in reasonable agreement with experiment results [84, 108] and numerical simulations [19]. The model also gives results in reasonable agreement with simulations for a significant range of compressible mixing flows.

It should be emphasized that theoretical calculations of RNG fixed points in non-equilibrium thermodynamics with experimental and numerical validation are not at all common. We believe that these methods can be applied to a range of related self similar stochastic instability problems with dynamic growth of length scales.

Both in direct numerical simulations, and in the RNG model described above, the growth rate was observed to depend on the system compressibility M^2 [19]. Moreover, in the direct numerical simulations, the mixing rate and other mixing zone scaling statistics appear to lose their universality in the compressible case. These results, surprising at the time, are now better understood. Compressibility introduces a new length scale, namely the length over which hydrostatic pressure gradients induce significant density change [20]. Such a length scale can break the universality of the fixed point, suggesting that some essential information will be contained in the initial conditions.

For $A = 1$ and incompressible flow, Layzer's formula [68] may replace the single bubble model 3.3 – 3.4. In the context of this formula, which is based on potential theory, the superposition hypothesis 3.6 can also be replaced by the statement that multi-bubbles are modeled by a superposition of Layzer potentials. This step was carried out, leading to a model and results related to those given above and to a new result [1], extending this model to the $A = 1$ RM multi-bubble incompressible growth rate. For this case the amplitude growth rate is given as $\text{const } t^\theta$, with $\theta = 0.4$ (theoretical) and $\theta = 0.35$ (numerical simulations).

3.4. Constitutive Laws and Averaged Equations. Here effective equations governing the flow in the interior of the mixing zone are derived. In the incompressible case, a scaling analysis shows that these equations also determine a RNG fixed point, which has been solved in closed form [46]. The closure of Amsden and Harlow [52] is commonly used to model two phase flow in the RT problem. For example, this closure was used in [33] to give a one parameter fit to the bubble and spike growth rates $\alpha_{b,s}$. We now discuss some features of the equations we propose which differ from those of [52].

The equations proposed here are hyperbolic, and are thus stable under mesh refinement for integration in time. They have independent pressures and temperatures for each fluid phase. The explicit solution in the incompressible RNG fixed point limit has distinct pressures, i.e., the pressures do not equilibrate. This feature of the equations, while unusual, has a simple and intuitive basis in physics. While there will be pressure equilibration between a small droplet of the heavy phase and the local ambient pressure of the less dense phase in which it is entrained, the pressures (in either phase) of a heavy phase-rich or droplet-rich region or column will have different hydrostatic forces and pressures from those of a light phase-rich region, due to the added weight of the heavy phase. Expected pressures \bar{p}_2 in the heavy phase will reflect preferentially the pressure values of the droplet-rich region, which is not equilibrated with the pressure in the droplet-poor region. Thus $\bar{p}_1 - \bar{p}_2 \neq 0$ and non-equilibration of the phase averaged pressures has a simple

physical meaning. Because of the independent pressures and temperatures in each phase, there is no need to introduce an equation of state for mixed phase fluids. Such an equation of state does not refer to any real physical situation. It is thus neither directly measurable nor computable, and its specification in practice is often a source of difficulty.

A characteristic analysis shows that at the boundary of the mixing region, where one of the phases goes to zero volume fraction, there is a missing boundary condition in the equations introduced below. The incoming sound waves for the phase of zero volume fraction cannot be specified by the equations given below. Thus the closure we propose is not complete, and must be supplemented by a boundary condition [47], expressing a balance among inertial, buoyant and form drag forces, on the phase of zero volume fraction, namely

$$(3.16) \quad (\rho_1 + k\rho_2) \frac{dv}{dt} = (\rho_1 - \rho_2)g + C_D \rho v^2 / |z| ,$$

where k is an added mass coefficient, C_D is a drag coefficient, ρ is the density in the fluid with volume fraction 1 at the mixing boundary, and Δ denotes a difference between the values of the same quantity in the two phases. After multiplication by the longitudinal length scale $|z|$, 3.16 equates forces per unit area. The right hand side of this expression is the standard formula for form drag per unit area for flow about an obstacle in the incompressible, large Reynolds number, laminar limit. See for example Fig. 1.2, page 5 of [92] for the drag coefficient for flow about a sphere as a function of Reynolds number.

The starting point for the averaged equations is the two fluid compressible Euler equations. In these equations, there are two time dependent complementary subsets giving the regions occupied by fluid k , $k = 1, 2$. Let X_k , $k = 1, 2$ denote the characteristic function of these sets, so that X_k takes on the value 1 at space time points in fluid 1, and the value 0 otherwise. We introduce the volume fraction $\beta_k = \langle X_k \rangle$ and for any fluid quantity f , we consider volume weighted phase averages $\bar{f}_k = \langle X_k f \rangle / \beta_k$ and with ρ the density, we also consider mass weighted averages $\tilde{f}_k = \langle X_k \rho f \rangle / \bar{\rho}_k$. Let v denote velocity, p pressure, and ϵ internal energy.

The derivation begins by multiplying the Euler equations of compressible fluids by X_k and averaging. Simple mathematical identities allow one to write most of the resulting expressions in terms of phase averaged first moments, but a few terms do not close, *i.e.*, do not have such an expression.

Our analysis leads to the following system of two-phase flow equations, proposed in [20],

$$(3.17) \quad \frac{\partial \beta_k}{\partial t} + v^* \frac{\partial \beta_k}{\partial z} = 0 ,$$

$$(3.18) \quad \frac{\partial \beta_k \bar{\rho}_k}{\partial t} + \frac{\partial \beta_k \bar{\rho}_k \bar{v}_k}{\partial z} = 0 ,$$

$$(3.19) \quad \frac{\partial \beta_k \bar{\rho}_k \bar{v}_k}{\partial t} + \frac{\partial \beta_k \bar{\rho}_k \bar{v}_k \bar{v}_k}{\partial z} = - \frac{\partial \beta_k \bar{\rho}_k}{\partial z} + \beta_k \bar{\rho}_k g + p^* \frac{\partial \beta_k}{\partial z} ,$$

$$(3.20) \quad \frac{\partial \beta_k \bar{\rho}_k \tilde{\epsilon}_k}{\partial t} + \frac{\partial \beta_k \bar{\rho}_k \tilde{\epsilon}_k \bar{v}_k}{\partial z} = - \bar{p}_k \frac{\partial \beta_k \bar{v}_k}{\partial z} + (pv)^* \frac{\partial \beta_k}{\partial z} ,$$

for $k = 1, 2$ and

$$(3.21) \quad \beta_1 + \beta_2 = 1 ,$$

where the quantities

$$(3.22) \quad v^* = \beta_1 \bar{v}_2 + \beta_2 \bar{v}_1 , \quad p^* = \beta_1 \bar{p}_2 + \beta_2 \bar{p}_1 , \quad (pv)^* = \beta_1 \bar{p}_2 \bar{v}_2 + \beta_2 \bar{p}_1 \bar{v}_1 ,$$

result from closure modeling assumptions which have been validated by comparison to front tracking simulations of the RT problem [20]. See [47] for a re-examination of and improvements upon these formulas.

Three approximations are made in the derivation of 3.17-3.22. The first is $\bar{v}_k \approx \tilde{v}_k$, saying that volume and mass weighted averages of the velocity are approximately equal. Although the RT mixing problem contains significant density variations, they are mainly between and not within the distinct phases. The ensemble average, in a single phase, will specify t , z , (the coordinates x and y do not enter due to symmetry of the geometry), and the phase, X_k . After fixing this information, remaining density fluctuations are not significant, as far as the velocity averages are concerned, according to simulation data from [20]. The second approximation is to neglect the single phase second moment Reynolds stress

$$(3.23) \quad R_k = \frac{\langle X_k \rho \vec{v} \vec{v} \rangle}{\beta_k} - \frac{\langle X_k \rho \vec{v} \rangle^2}{\beta_k \langle X_k \rho \rangle}$$

and the related single phase energy velocity correlation

$$(3.24) \quad S_k = \frac{\langle X_k e \vec{v} \rangle}{\beta_k} - \frac{\langle X_k e \rangle \langle X_k \rho \vec{v} \rangle}{\beta_k \langle X_k \rho \rangle} ,$$

where $e = \epsilon \rho$ is the internal energy per unit mass. Analysis of the numerically generated RT mixing data suggests that these quantities are small. In this sense, we are assuming that the turbulence within each phase is of less significance than the mixing caused by the presence of the two phases. If ensemble averages are formed without the presence of the phase function X_k , then the full Reynolds stress R of 2.5 occurs in the averaged equations, and a related energy-density second moment S . These quantities are not negligible. In fact an identity relates R to the R_k and the first moments,

$$(3.25) \quad R = \beta_1 R_1 + \beta_2 R_2 + \beta_1 \beta_2 \frac{\bar{\rho}_1 \bar{\rho}_2}{\bar{\rho}} (\tilde{v}_1 - \tilde{v}_2)^2 ,$$

and a similar identity holds for S . These identities, and the relative vanishing of the first two terms in them show that the RT data set is dominated by mixing phenomena, and its description in terms of turbulence concepts (such as R) is inefficient.

A more general interpretation of 3.25 and related identities for the other second moments is that they provide a smooth blending of turbulence and multi-phase flow models, with the turbulence modeled strictly as a single phase phenomena, occurring within a single phase, while the two phase modeling joins the two phases. In this view, each R_k , etc. can be further modeled by a limited second order closure, such as the $k - \epsilon$ model or by a full second order closure. Our observation that the R_k etc. can be modeled as being zero in the data we have examined is clearly not universal. There will clearly be data for which this assumption is not valid and the R_k etc. should be retained in the flow equations. Late time mixing should be an example of such a flow regime.

The third modeling approximation is contained in 3.22. The exact quantities these formulas approximate are inter-facial quantities. Thus 3.17 replaces the exact quantity $\langle v_{\text{int}} \cdot \nabla X_k \rangle$, where v_{int} is the velocity at the interface, and ∇X_k is a delta function at the interface, by the approximation $v^* \partial \beta_k / \partial z$. In other words, it replaces v_{int} by v^* . Now v^* equals v_1 identically when $\beta_1 = 0$, that is, at the edge of the mixing zone, where the phase $k = 1$ vanishes. At this location, the $k = 1$ fluid is necessarily at the interface, so this formula is exact. Likewise the expression for v^* is exact at the other edge of the mixing zone. Between these two locations where the formula is exact, it provides a reasonable and simple interpolation, which is justified by the simulation data.

We next introduce scaled variables, $z' = z/Agt^2$, $v' = v/2Agt$ and $t' = \ln t$, in accordance with the RNG fixed point scaling from §3.3. Rewritten in these variables, the dynamical equations become the RNG equations for reasons explained in §3.3: The problem coarse-grains itself dynamically in time. Ordinary time, combined with a dynamic rescaling of lengths, is thus the RNG parameter. To obtain a RNG fixed point, we look for time independent solutions of rescaled equations. For simplicity, we drop the primes, as well as the over-bars and tildes which indicate volume and mass weighted averages. We write the RNG fixed point equations as [46]

$$(3.26) \quad -2z \frac{\partial \beta_k}{\partial z} + v^* \frac{\partial \beta_k}{\partial z} = 0$$

$$(3.27) \quad -2z \frac{\partial \beta_k \rho_k}{\partial z} + \frac{\partial \beta_k \rho_k v_k}{\partial z} = 0$$

$$(3.28) \quad -2z \frac{\partial \beta_k \rho_k v_k}{\partial z} + \frac{\partial \beta_k \rho_k v_k v_k}{\partial z} + \beta_k \rho_k v_k = -\frac{\partial \beta_k p_k}{\partial z} + p^* \frac{\partial \beta_k}{\partial z} + \frac{\beta_k \rho_k}{A}$$

$$(3.29) \quad -2z \frac{\partial \beta_k \rho_k \epsilon_k}{\partial z} + \frac{\partial \beta_k \rho_k \epsilon_k v_k}{\partial z} + 2\beta_k \rho_k \epsilon_k = -p_k \frac{\partial \beta_k v_k}{\partial z} + (pv)^* \frac{\partial \beta_k}{\partial z}$$

for $k = 1, 2$ and also the unchanged equations 3.21-3.22.

The closed form solution to these equations in the incompressible limit is easily described [46]. We impose boundary conditions at $z = \alpha_s$ and at $z = -\alpha_b$. The volume fractions β_k vary linearly between 0 and 1 over these limits. The phase velocities v_k are also linear over these limits,

$$(3.30) \quad v_1 = z - \alpha_b, \quad v_2 = z + \alpha_s.$$

The two phase momentum equations can also be integrated explicitly, to yield the two phase pressures, p_1 and p_2 . Integration of these equations introduces two new boundary conditions, which, however, do not influence the solutions of the volume fraction and velocity equations.

If the drag coefficients $C_{D,b}$ and $C_{D,s}$, at the lower and upper boundaries of the mixing zone, as in 3.16, are given, then the equations 3.17-3.21 as well as 3.26-3.29 close. Thus the boundary relations can be written as

$$(3.31) \quad \alpha_b = \alpha_b(C_{D,b}, C_{D,s}),$$

with a similar relation for α_s .

The existence of a fixed point solution of the RNG equations is strong evidence for our claim that the incompressible RT mixing is described by a RNG fixed point.

4. Transport and Dispersion

4.1. Linear Transport and Dispersion for Flow in Porous Media. The simplest case for flow in porous media is that of two miscible fluids with the same viscosity. This case includes transport of a dissolved solute, which includes many environmental problems, and for oil reservoir simulation, which includes the study of tracers, used as a diagnostic to aid in determining flow patterns. This case captures one essential difficulty of the general problem while ignoring others.

For miscible fluids, the flow of each phase is proportional to its relative volume,

$$(4.1) \quad k_{r,w}(s) = s, \quad k_{r,o}(s) = 1 - s,$$

so that with equal fluid viscosities, the relative permeabilities drop out of Darcy's law 1.7 and Darcy's law decouples from the Buckley-Leverett equation 1.8. Moreover, the velocity v in 1.8 is the total velocity of the two phases, and the flux function f in 1.8 is the ratio of these two velocities, which now becomes

$$(4.2) \quad f(s) = \frac{k_{r,w}(s)}{k_{r,w}(s) + k_{r,o}(s)} = s,$$

since the viscosities of the two phases are assumed to be equal, and drop out of the ratio. Thus the problem simplifies to the linear transport equation

$$(4.3) \quad \frac{\partial s}{\partial t} + \vec{v} \cdot \nabla s = 0,$$

with v given as an external velocity field, coming from Darcy's law and depending only on the geology. With these assumptions, the fluid flow is elementary, and we can concentrate on difficulties introduced by the geology. The available information concerning the geology is normally very limited, but what is available is extremely important. Information comes from diverse sources, such as seismic images, laboratory examination of samples of rock removed in the drilling process, local measurements made at the well bore, pressure transient tests made at a single well, and the use of production data in an inverse problem (history matching). Current thinking in the petroleum reservoir modeling community is to construct a probability model, based on general geological principles and data, specialized to the rock types under study, and then conditioned by the problem specific data. See [66, 67, 60, 22] as background references. A commonly used model is a combination of Ising type variables, or indicator functions, to describe the discrete change between distinct rock types, such as sandstone, shale, carbonate etc., and a log-normal distribution for the variation within a single depositional type. Both of these distributions can be described as random fields. The spatial correlation structure of these fields is very important. In the log-normal case, the covariance of the log permeability, $\ln K$, will determine the statistics of K completely. Exponentials, power laws (fractals) and stationary increment power laws are sometimes used as simple models, in the absence of better data.

We now turn to the solution of the linear transport equation with an external velocity field v defined by given statistics [25]. This identical problem arises in a number of other contexts; see §4.3. The main influence of the random variations in the velocity is to introduce dispersion into the flow. With the velocity given as a random field, the flow is stochastic, and the ensemble averaged equations acquire a diffusion term. It is exactly this process, at the level of atomic interactions, which gives rise to diffusion as a continuum concept. But here we are concerned

with geology and velocity variation correlations on all length scales, especially those with correlation lengths extending from a meter to a kilometer. Examination of field data shows that the dispersion coefficient resulting from velocity dispersion is not a constant, but grows systematically, and by several orders of magnitude, as the travel distance increases [37]. For this reason, the dispersion process is called *anomalous*.

In the numerical simulations discussed below, the permeability field in Darcy's law is assumed to be log-normal; in the theoretical work, the velocity field itself is specified directly, and is further simplified and approximated, to be normal. For the purposes of comparison of large distance asymptotic behavior, these two specifications of the velocity field show good agreement.

In a series of papers, both numerical [43, 44], and theoretical [49, 112, 113, 114] (see also references cited there), the connection between multi-length scale heterogeneity and anomalous diffusion (dispersion coefficients which increase with travel distance) has been clearly understood. Basically, as the transport encounters new parts of the reservoir, it not only encounters new instances of heterogeneity, but it also encounters heterogeneity on new length scales. Thus if the transport occurs over a distance of, say, 10 m, then heterogeneities of length 10 m or more will not contribute to the dispersion, but will affect instead the mean flow. However, if the same transport is continued for 100 m or 1000 m, then the 10 m heterogeneities would be removed from the average flow properties, and would be seen as contributing to the dispersion coefficient. This process is now called "evolving heterogeneities" in the ground water community, and is accepted as the explanation of anomalous diffusion in porous media. The analysis and simulations cited above were based on a fractal model of heterogeneity.

The fractal model was chosen because it was the simplest which contains variation on every length scale. Power law (fractal) correlations are guaranteed to be singular at either long or short distances, and should not be extrapolated into their singular limits. Thus stationary processes will have a power law correlation which vanishes at infinity, but is singular at the origin, and thus should not be extrapolated to short distances. Fractal stationary increment processes, which are not themselves stationary, will have undefined correlation functions, which if defined formally are singular at spatial infinity. Thus their covariance is defined in terms of differences, or test functions of vanishing mean, and is then well defined. Their fractal powers, called Hurst exponents, allow locally continuous realizations with probability one. These processes can be extrapolated to short distances, but are qualitatively unsatisfactory at large distances. For these reasons, the fractal model is not more than a model. The stationary increment model can be desingularized by conditioning. If data is given at large distances, for example at well spacings on a regular grid, then the fractal stationary increments model conditioned to this data (by Kriging) will not be fractal, but will be regular at both short and long distances.

To analyze dispersion quantitatively, we write the random velocity field v in 4.3 as a sum of its mean and fluctuating components, $v = \bar{v} + \delta v$, and then expand the solution of 4.3 in powers of δv . To lowest order, the solution moves with the mean velocity field \bar{v} . It is convenient to subtract out this motion through use of

moving coordinates $\xi = x - \bar{v}t$. In this coordinate system, 4.3 becomes

$$(4.4) \quad \frac{\partial s}{\partial t} + \delta v \cdot \nabla s = 0 .$$

In these coordinates, the zeroth order solution, $s_0(\xi, t) = s_0(\xi, t=0)$ is given by the initial conditions, and we define iteratively $s_n = s_{n-1} + \int_0^t \delta v \nabla s_{n-1} dt$, so that $s = \lim_{n \rightarrow \infty} s_n$. Dispersion is a property of the averaged field, $\bar{s} = \langle s \rangle$. The exact equation for \bar{s} is very complicated (a version of the closure problem), so we proceed perturbatively. Only even powers of δv contribute to the averaged solution \bar{s} , and the lowest nontrivial term is the second. At this order, and using the fact that the velocity field is incompressible, we derive the equation

$$(4.5) \quad \frac{\partial \bar{s}}{\partial t} + \bar{v} \cdot \nabla \bar{s} = \nu \frac{\partial^2 \bar{s}}{\partial x^2}$$

for \bar{s} in the original coordinates. Here the viscosity coefficient ν is time dependent, and given by

$$(4.6) \quad \nu(t) = \int_0^t \langle \delta v(x - st) \delta v(x) \rangle ds .$$

The time asymptotic equation is the normal diffusion equation with the diffusion constant $\nu_\infty = \lim_{t \rightarrow \infty} \nu(t)$, provided that ν_∞ is finite. Thus we see that nonintegrability of the velocity correlation function along mean flow streamlines is directly associated with the opposite case, of anomalous dispersion, in which ν_∞ is infinite and the time dependent dispersivity $\nu(t)$ increases with travel distance. Detailed asymptotic analyses of the growth rate of $\nu(t)$ as a function of the decay rates of the velocity correlation $\langle \delta v \delta v \rangle$ can be found in [112].

We next consider a more refined theory based on Corrsin's hypothesis. This analysis can be regarded as the leading order contribution to a renormalized perturbation theory. While its formal level of validity for small amplitudes remains second order, we believe it gives a more accurate solution for larger amplitudes of the perturbation parameter δv . In the Lagrangian picture, we let $x(t) = \bar{x} + \delta x$ be the particle displacement, divided into mean and fluctuating parts, where $dx(t)/dt = v(x(t), t)$ and, for example, $x(0) = 0$. Also let $X = \langle \delta x(t) \delta x(t) \rangle$. Taking Fourier transforms,

$$(4.7) \quad \frac{d^2 X}{dt^2} = \frac{2}{(2\pi)^m} \int \int \langle \delta \tilde{v}(k', 0) \delta \tilde{v}^*(k, t) e^{ik \cdot x} \rangle dk' dk ,$$

where m is the spatial dimension. The triple product in the angle brackets presents a closure problem. Corrsin's hypothesis declares, as an approximation, that the exponent is independent of the velocities, leading to a closed equation if the displacement process $x(t)$ is assumed to be Gaussian.

Normally Corrsin's hypothesis is justified asymptotically as $t \rightarrow \infty$, as an assumption that x and v are independent in that limit. In [114] it is argued that Corrsin's hypothesis results from neglect of non Gaussian correlations between v and x , for finite as well as infinite times, and does not require full independence of v and x . A careful asymptotic analysis of the (closed) Corrsin's equation, for fractally correlated velocity fields, then yields a new derivation of the anomalous dispersion behavior of the transport, with explicit determination of the coefficients in the leading order behavior [113].

Ergodicity is the statement that time or space averages coincide with ensemble averages. When ergodicity fails, these averages will have different meanings and give different answers. In this case, a proper interpretation of the meaning of each average is needed, as is discussed in [26, 27] (see also later papers in this sequence) and [44, 3]. For the case of a plume of a contaminant or passive tracer in the porous media flow, with a bounded (small) source, or initial distribution, ergodicity will be influenced by the amount of lateral dispersion as the plume moves under the random transport 4.3. When evaluated to lowest (second) order primitive perturbation theory, the plume has no lateral dispersion, and so it encounters less randomness than that of the ensemble average; within this computation the flow is non-ergodic [26, 27]. Because of the sensitivity of anomalous dispersion to ergodicity and related matters, further studies of lateral plume dispersion would be desirable. We further note that the ratio of longitudinal to transverse length scales has been identified as a significant parameter [109]. In general, dispersion determined from ensemble means includes effects of uncertainty due to unknown geology, while spatial averages, which yield a smaller dispersion coefficient, omit this effect and describe spreading of the plume about its (uncertain) sample mean position.

4.2. Nonlinear Transport for Flow in Porous Media. Nonlinear transport in porous media depends sensitively on data, and thus requires stochastic modeling, for two reasons. First, the flow exhibits nonlinear instability and chaos, similar to the two phase mixing flows of §3, and secondly, the equations depend directly on data in a sensitive manner, through rapidly varying and statistically specified geology, as indicated in §4.1. It is convenient first to consider the nonlinear instability alone, in an idealized homogeneous porous media. The full equations for two or three phase flow, generalizing those given above to allow for capillary pressure and mass transfer between phases (solubility and evaporation) are given in [65, 83]. For miscible displacement, and with neglect of mass diffusion, the total fluid velocity is given by Darcy's law, 1.7, 1.9, and 4.1. The evolution equations will propagate an initially given sharp interface, separating regions of the two phases (e.g., oil and solvent) as an interface of the same type. This flow is mathematically singular, in that the interface develops cusps, unless regularized by surface tension. See, for example, [95], and also [94, 96], where the same issue arises for the Rayleigh-Taylor mixing problem. These singularities and their regularization, even though mathematically and perhaps physically correct, and irrelevant, because they pertain to small length scales, still within the stochastic regime. The regularization removes some singularities, and has a role in the dynamics, but it does not remove the stochasticity from the solutions. The stochastic regularization and the ensemble averaged equations with deterministic behavior for their solutions is of greater importance, because with this approach, the chaotic degrees of freedom have been integrated out, and only the large scale, deterministic ones remain.

The nature of the stochastic solutions has been the subject of much study. The ratio of phase mobilities in Darcy's law is known as the mobility ratio, denoted M . For miscible displacements, $M = \mu_o/\mu_w$ is equal to the ratio of the phase viscosities. The case $M > 1$ is the unstable case. The method of Diffusion Limited Aggregation (DLA) yields fractal displacement patterns for $M = \infty$, and extensions of this method to finite mobility ratio $M > 1$ miscible displacement and to Buckley-Leverett flow 1.8 [62, 63] suggest the same behavior. However, Otto

derived rigorous results [82] for Buckley-Leverett flow which suggest compact (non-fractal) displacements for the long time asymptotics, so that the fractal regime, if it occurs, would apply to some intermediate time asymptotics.

Todd-Longstaff and Koval introduced the flux functions [65]

$$(4.8) \quad f_{\text{ren}} = \frac{\bar{s}}{\bar{s} + (1/M_{\text{eff}})(1 - \bar{s})}$$

into the Buckley-Leverett equation

$$(4.9) \quad \frac{\partial \bar{s}}{\partial t} + \nabla v f_{\text{ren}} = 0$$

to model miscible displacement in heterogeneous media. Phenomenological formulas have been proposed for M_{eff} . Morris and Ball approach the problem of determining M_{eff} as an RNG fixed point problem. In the case of heterogeneous media, the stochasticity of the media and that of the nonlinear flow interact, presumably shifting the location of the relevant RNG fixed point. This fixed point can then be determined numerically [78, 77], through successive levels of upscaling. Finally we note that the true long distance, large time asymptotics resulting from successive upscaling may be less relevant than the range of possible transient, or partially upscaled, theories, because the number of length scales available in applications is limited, and because the geology will not necessarily be self-similar, but will introduce new information at each length scale.

4.3. Dispersion for Turbulent Transport Revisited. In a series of papers, Avellaneda and Majda [6, 7, 8] studied diffusion of a random scalar by a turbulent velocity field v , based on 4.3. The velocity field v is the input to the problem. It is assumed to be defined by Gaussian random statistics, with a power law (fractal) covariance. There are distinct powers for the space and time correlations, and accordingly, the statistics is defined by two parameters, namely the time and space exponents ϵ and z of the velocity correlation. Although not identified as significant in these papers, the statistics also depends on a third parameter: the factor which multiplies the power law behavior. This factor is also a function of ϵ and z , when the physics is expressed at the integral scales, in their hypothesis. Their results show a remarkable change in structure (termed a phase transition) in the asymptotic behavior of the transport as the exponents pass through the values which define the Kolmogorov spectrum 1.6.

Wallstrom [100] identified the origin of the Avellaneda-Majda phase transition in the input data to the problem, namely in the hypothesis on the statistics of the velocity field v . Specifically, the multiplying factor depends on ϵ and z in a singular fashion. Thus the assumed multiplying factor in the input statistics to the problem contains a phase transition. The energy spectrum of the velocity field assumed by Avellaneda-Majda depends on the Reynolds number R (*i.e.*, on the viscosity), and for exponents near the Kolmogorov values, the energy spectrum has a phase transition itself, diverging to zero or infinity as $R \rightarrow \infty$, depending on which side of the Kolmogorov values the exponents lie. This phase transition in the input data of the problem is an indirect consequence of the Avellaneda-Majda formulation. These authors transform to the viscous length scale before considering the intermittency corrections as we now explain.

To simplify the discussion, we will only consider the spatial exponent ϵ . The energy spectrum, being a power law, is a linear function when plotted in the log-log variables $\ln E$ vs. $\ln k$. At the Kolmogorov point, this line has slope $-5/3$. Nearby energy spectrum plots, suggesting possible intermittency corrections, will have nearly the same slope, but to retain the fundamental hypothesis of turbulence, they will pass through the common integral scale value of E , and each intermittency curve will then be independent of R .

In contrast, Avellaneda-Majda assume all intermittency correction energy spectra to pass through a common point at the viscous length scale. To restate their results at the integral scale does not change the exponents, but it does change the multiplying factor. This is how the Avellaneda-Majda form of the Kolmogorov coefficient c_K of 1.6 acquires its singular, phase transition character, when expressed at integral scales. An essential aspect of their assumptions is that the Kolmogorov constant c_K of 1.6 depends on both the intermittency corrections to the Kolmogorov exponents and on the Reynolds number R . Moreover, the dependence on R is singular, in the sense that the Kolmogorov constant c_K diverges to infinity as $R \rightarrow \infty$, in the hypotheses to their analysis.

Avellaneda-Majda do not identify the unusual features of their hypotheses (the R dependent singular energy spectrum and the divergence of c_K) in their series of papers; this aspect of their hypotheses was first discovered by Wallstrom. See also the discussion of [39, 72]. Mathematical problems with the Avellaneda-Majda theory, first identified in [115], are not pursued here, in view of the more serious difficulties concerning the physical formulation.

In [72], Majda asserts that his modeling assumptions are hotly debated within the turbulence community. We are not aware of this debate. To debate this point would assume, in the case of nonzero intermittency corrections, that the total system energy must diverge as $R \rightarrow \infty$. Conventional physical arguments assume that the total system energy, for example of the atmosphere, is limited by the energy supplied by external driving sources, such as solar heating, and does not have a divergent dependence on R as $R \rightarrow \infty$. Any debate of the Avellaneda-Majda hypotheses would have to overcome or refute this simple physical limitation on the total system energy; it would also have to explain the apparent constancy and lack of divergence of c_K to infinity in simulation, laboratory and observational data, as referenced in §1.2.

There *are* many debates within the turbulence community. The one with closest relevance to the issues raised above is whether there is to some extent an inverse cascade, or partial flow of information from the small scales to the larger ones. A recent version of this argument is given by Eyink [32]. Eyink shows that many of the terms dropped as irrelevant in the RNG analysis of [105] are in fact neutral, and must be retained at each RNG integration step. On this basis, Eyink speculates that the details of the dissipation range dynamics may influence the choice of inertial range scaling laws and that these scaling laws may be non-universal. Eyink also speculates that the higher moment, or non Gaussian, exponents may be more sensitive to such non-universality than are the exponents of the energy spectrum, as in 1.6.

Eyink also questions the stirring forces used by Yakhot and Orszag. These stirring forces are of power law type, with a typical energy spectrum given by the Kolmogorov law 1.6 and nearby exponents. They are exactly the assumed velocity fields of Avellaneda-Majda.

Bibliography

- [1] U. Alon, J. Hecht, D. Mukamel, and D. Shvarts. Scale invariant mixing rates of hydrodynamically unstable interfaces. *Phys. Rev. Lett.*, 72:2867–2870, 1994.
- [2] U. Alon, J. Hecht, D. Ofer, and D. Shvarts. Power laws and similarity of Rayleigh-Taylor and Richtmyer-Meshkov mixing fronts at all density ratios. *Phys. Rev. Lett.*, 74:534–538, 1995.
- [3] L. An, J. Glimm, D. H. Sharp, and Q. Zhang. Scale up of flow in porous media. In A. Burgeat, C. Carasso, S. Luckhaus, and A. Mikelić, editors, *Mathematical Modeling of Flow Through Porous Media*, pages 26–44. World Scientific, 1995.
- [4] M. J. Andrews. The large-scale fragmentation of the intact liquid core of a spray jet. *Atomization and Sprays*, 3:29–54, 1993.
- [5] T. Arbogast, J. Douglas, and U. Hornung. Modeling of naturally fractured reservoirs by formal homogenization techniques. In *Frontiers in Pure and Applied Mathematics*. North-Holland, Amsterdam, 1991.
- [6] M. Avellaneda and A. Majda. Mathematical models with exact renormalization for turbulent transport. *Comm. Math. Phys.*, 131:381–429, 1990.
- [7] M. Avellaneda and A. Majda. Mathematical models with exact renormalization for turbulent transport II. fractal interfaces, non-Gaussian statistics and the sweeping effect. *Comm. Math. Phys.*, 146:139–204, 1992.
- [8] M. Avellaneda and A. Majda. Renormalization theory for eddy diffusivity in turbulent transport. *Phys. Rev. Lett.*, 68:3028–3031, 1992.
- [9] G. I. Barenblatt. *Dimensional Analysis*. Gordon and Breach, New York, 1987.
- [10] R. Benjamin, D. Besnard, and J. Haas. Shock and re-shock of an unstable interface. LANL report LA-UR 92-1185, Los Alamos National Laboratory, 1992.
- [11] M. Berger and P. Colella. Local adaptive mesh refinement for shock hydrodynamics. *J. Comp. Phys.*, 82:64–84, 1989.
- [12] N. N. Bogoliubov and D. V. Shirkov. *Introduction to the Theory of Quantized Fields*. Interscience Publishers Ltd, New York, 1959.
- [13] B. Boston. *Front Tracking of Complex Wave Interactions*. PhD thesis, State University of New York at Stony Brook, 1995.
- [14] J. Bricmont and A. Kupiainen. Renormalizing partial differential equations. Technical Report Preprint, 1994.
- [15] J. Bricmont, A. Kupiainen, and G. Lin. Renormalization group and asymptotic solutions of nonlinear partial differential equations. *Comm. Pure Appl. Math.*, 47:893–922, 1994.
- [16] J. Cannon. Continuous sample paths in quantum field theory. *Comm. Math. Phys.*, 35:215–234, 1974.
- [17] Nianzheng Cao, Shiyi Chen, and Zhen-Su She. Scalings and relative scalings in the Navier-Stokes turbulence. *Phys. Rev. Lett.*, 76:3711–3714, 1996.
- [18] S. Chandrasekhar. *Hydrodynamic and Hydromagnetic Stability*. Dover Publications, New York, Oxford edition, 1981. original 1961.
- [19] Y. Chen, Y. Deng, J. Glimm, G. Li, D. H. Sharp, and Q. Zhang. A renormalization group scaling analysis for compressible two-phase flow. *Phys. Fluids A*, 5(11):2929–2937, 1993.
- [20] Y. Chen, J. Glimm, D. H. Sharp, and Q. Zhang. A two-phase flow model of the Rayleigh-Taylor mixing zone. *Phys. Fluids*, 8(3):816–825, 1996.
- [21] A. Chorin. *Vorticity and Turbulence*. Springer Verlag, New York–Heidelberg–Berlin, 1994.
- [22] G. Christakos. *Random Field Models in Earth Sciences*. Academic Press, New York, 1992.
- [23] S. E. Cohn. Dynamics of univariate forecast error covariances. *Mon. Wea. Rev.*, pages 3132–3149, 1993.
- [24] P. Colella and O. Lanford. Sample field behavior for the free random Markov field. In G. Velo and A. Wightman, editors, *Constructive Quantum Field Theory*, pages 44–70. Springer-Verlag, New York, 1973.
- [25] G. Dagan. *Flow and Transport in Porous Formations*. Springer-Verlag, New York, 1989.
- [26] G. Dagan. Dispersion of a passive solute in non-ergodic transport by steady velocity fields in heterogeneous formations. *J. Fluid Mech.*, 233:1281–1290, 1991.
- [27] G. Dagan. The significance of heterogeneity of evolving scales to transport in porous formations. *Water Resources Research*, 30:3327–3336, 1994.
- [28] P. Diaconis. Private communication, 1992.

- [29] D. A. Drew. Mathematical modeling of two-phase flow. *Ann. Rev. Fluid Mech.*, 15:261–291, 1983.
- [30] D. A. Drew and R. T. Lahey. Analytical modeling of multi-phase flow. In M. C. Roco, editor, *Particulate Two-Phase Flow*. Butterworth-Heinemann, Stoneham, MA, 1993.
- [31] P. W. Dromgoole and R. G. Speers. Managing uncertainty in oil field reserves. *Schlumberger Middle East Well Evaluation Review*, (12):30–41, 1992.
- [32] G. Eyink. The renormalization group method in statistical hydrodynamics. *Phys. Fluids*, 6:3036–3078, 1994.
- [33] N. Freed, D. Ofer, D. Shvarts, and S. Orszag. Two-phase flow analysis of self-similar turbulent mixing by Rayleigh-Taylor instability. *Physics of Fluids A*, 3(5):912–918, 1991.
- [34] U. Frisch, P.-L. Sulem, and M Nelkin. A simple dynamic model of intermittent fully developed turbulence. *J. Fluid Mech.*, 87:719–736.
- [35] C. L. Gardner, J. Glimm, O. McBryan, R. Menikoff, D. Sharp, and Q. Zhang. The dynamics of bubble growth for Rayleigh-Taylor unstable interfaces. *Phys. of Fluids*, 31:447–465, 1988.
- [36] I. Gelfand and N. Vilenkin. *Generalized Functions, Vol IV (English Translation)*. Academic Press, New York, 1964.
- [37] L. W. Gelhar, A. Mantoglou, C. Welty, and K. R. Rehfeldt. A review of field-scale physical solute transport processes in saturated and unsaturated porous media. EPRI Report EA-4190, 1985.
- [38] M. Gell-Mann and F. Low. Quantum electrodynamics at small distances. *Phys. Rev.*, 95:1300–1312, 1954.
- [39] J. Glimm. Turbulent diffusion revisited. *SIAM News*, 28:6, July 1995.
- [40] J. Glimm, J. W. Grove, X. L. Li, K.-M. Shyue, Q. Zhang, and Y. Zeng. Three dimensional front tracking. *SISC*, To appear, 1996.
- [41] J. Glimm and Arthur Jaffe. *Quantum Physics: A Functional Integral Point of View, Second Edition*. Springer-Verlag, New York, 1987.
- [42] J. Glimm and X. Li. On the validation of the Sharp-Wheeler bubble merger model from experimental and computational data. *Phys. Fluids*, 31:2077–2085, 1988.
- [43] J. Glimm, W. B. Lindquist, F. Pereira, and R. Peierls. The multi-fractal hypothesis and anomalous diffusion. *Mat. Aplic. Comput.*, 11:189–207, 1992.
- [44] J. Glimm, W. B. Lindquist, F. Pereira, and Q. Zhang. A theory of macro-dispersion for the scale up problem. *Transport in Porous Media*, 13:97–122, 1993.
- [45] J. Glimm, D. Marchesin, and O. McBryan. Statistical fluid dynamics: Unstable fingers. *Comm. Math. Phys.*, 74:1–13, 1980.
- [46] J. Glimm, D. Saltz, and D. H. Sharp. Renormalization group solution of two-phase flow equations for Rayleigh-Taylor mixing. *Phys. Lett. A*, 222:171–176, 1996.
- [47] J. Glimm, D. Saltz, and D. H. Sharp. Two-phase modeling of a fluid mixing layer. Technical report, SUNY at Stony Brook, 1996. to appear.
- [48] J. Glimm and D. H. Sharp. Chaotic mixing as a renormalization group fixed point. *Phys. Rev. Lett.*, 64:2137–2139, 1990.
- [49] J. Glimm and D. H. Sharp. A random field model for anomalous diffusion in heterogeneous porous media. *J. Stat. Phys.*, 62:415–424, 1991.
- [50] N. Goldenfeld, O. Martin, and Y. Oono. Asymptotics of partial differential equations and the renormalization group. In H. Segur, editor, *Asymptotics Beyond All Orders*. Plenum Press, New York, 1991.
- [51] J. W. Grove. Applications of front tracking to the simulation of shock refractions and unstable mixing. *J. Appl. Num. Math.*, 14:213–237, 1994.
- [52] F. Harlow and A. Amsden. Flow of interpenetrating material phases. *J. Comp. Phys.*, 18:440–464, 1975.
- [53] A. Harten. High resolution scheme for hyperbolic conservation laws. *J. Comp. Phys.*, 49:357, 1983.
- [54] R. L. Holmes. *A Numerical Investigation of the Richtmyer-Meshkov Instability Using Front Tracking*. PhD thesis, State Univ. of New York at Stony Brook, 1994.
- [55] R. L. Holmes, J. W. Grove, and D. H. Sharp. Numerical investigation of Richtmyer-Meshkov instability using front tracking. *J. Fluid Mech.*, 301:51–64, 1995.
- [56] G. M. Homsy. Viscous fingering in porous media. *Ann. Rev. Fluid Mech.*, 19:271–311, 1987.

- [57] U. Hornung. Applications of the homogenization method to flow and transport in porous media. In Xiao Shutie, editor, *Summer School on Flow and Transport in Porous Media*, pages 167–222. World Scientific, Singapore, 1992.
- [58] P. L. Houtekamer, L. Lefavire, J. Derome, H. Ritchie, and H. Mitchell. A system approach to ensemble predictions. *Mon. Wea. Rev.*, 124:1225–1242, 1996.
- [59] K. Huang. *Statistical Mechanics*. John Wiley & Sons, New York, 1963.
- [60] A. G. Journel and Ch. J. Huijbregts. *Mining Geostatistics*. Academic Press, New York, 1978.
- [61] S. Kida and Y. Murakami. Kolmogorov similarity in freely decaying turbulence. *Phys. Fluids*, 30:2030–2039, 1987.
- [62] M. J. King and H. Scher. Probability approach to multi-phase flow in porous media. *Phys. Rev.*, 35:929–932, 1987.
- [63] M. J. King and H. Scher. Geometric dispersion and unstable flow in porous media. *Phys. Rev.*, 41:874–884, 1990.
- [64] A. Kosevich. Crystal dislocations and the theory of elasticity. In F. Nabarro, editor, *Dislocations in Solids*, pages 33–141. North-Holland, Amsterdam, 1979.
- [65] L. Lake. *Enhanced Oil Recovery*. Prentice-Hall, Englewood Cliffs, 1989.
- [66] L. Lake and H. Carroll, editors. *Reservoir Characterization*. Academic Press, New York, 1986.
- [67] L. Lake, H. Carroll, and T. Wesson, editors. *Reservoir Characterization II*. Academic Press, New York, 1991.
- [68] D. Layzer. On the instability of superimposed fluids in a gravitational field. *Astrophys. J.*, 122:1–12, 1955.
- [69] R. LeVeque. *Numerical Methods for Conservation Laws*. Birkhäuser Verlag, Basel–Boston–Berlin, 1992.
- [70] X. L. Li, B. X. Jin, and J. Glimm. Numerical study for the three dimensional Rayleigh–Taylor instability using the TVD/AC scheme and parallel computation. *J. Comp. Phys.*, 126:343–355, 1996.
- [71] E. N. Lorenz. Deterministic non-periodic flow. *J. Atmos. Sci.*, 20:130–141, 1963.
- [72] A. Majda. Turbulent diffusion revisited. *SIAM News*, 28:6, July 1995.
- [73] Benoit Mandelbrot. Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. *J. Fluid Mech.*, 62:331–358, 1974.
- [74] W. D. McComb. *The Physics of Fluid Turbulence*. Oxford University Press, Oxford, 1990.
- [75] R. Menikoff, R. C. Mjolness, D. H. Sharp, and C. Zemach. Initial value problem for Rayleigh–Taylor instability of viscous fluids. *Phys. Fluids*, 21:1674–1687, 1978.
- [76] J. Miller. Statistical mechanics of Euler equations in two dimensions. *Phys. Rev. Lett.*, 65:2137, 1990.
- [77] M. I. Morris and R. C. Ball. Renormalization of miscible flow in hierarchically correlated heterogeneous media. Cavendish laboratory preprint.
- [78] M. I. Morris and R. C. Ball. Renormalization of miscible flow functions. *J. Phys. A*, 23:4199–4209, 1990.
- [79] R. Mulder, S. Osher, and J. A. Sethian. Computing interface motion in compressible gas dynamics. *J. Comp. Phys.*, 100:209–228, 1992.
- [80] W. Noh and P. Woodward. SLIC (simple line interface calculation). In A. I. van de Vooren and P. J. Zandbergen, editors, *Lecture Notes in Physics, Proceedings of the Fifth International Conference on Numerical Methods for Fluid Dynamics*, volume 59, pages 330–340. Springer Verlag, New York, 1976.
- [81] H. Omre, Yuanchang Qi, and L. Hinderaker. Assessment of uncertainty in the production characteristics of a sandstone reservoir. In W. Linville, editor, *Reservoir Characterization III*, pages 2–22. Penwell Books, 1993.
- [82] F. Otto. Evolution of micro-structure in unstable porous media flow: a relaxational approach. Technical Report Preprint.
- [83] D. Peaceman. *Fundamentals of Numerical Reservoir Simulation*. Elsevier, Amsterdam–New York, 1977.
- [84] K. I. Read. Experimental investigation of turbulent mixing by Rayleigh–Taylor instability. *Physica D*, 12:45, 1984.
- [85] R. Robert. A maximum-entropy principle for two-dimensional perfect fluid dynamics. *J. Stat. Phys.*, 65:531, 1991.
- [86] D. H. Sharp. An overview of Rayleigh–Taylor instability. *Physica D*, 12:3–18, 1984.

- [87] D. H. Sharp and J. A. Wheeler. Late stage of Rayleigh-Taylor instability. In *Institute of Defense Analyses. Unpublished Technical Report*, 1961.
- [88] Z.-S. She. Dynamical mechanisms for intermittency effects in fully developed turbulence. In Moffatt and Tabor, editors, *Proceedings NATO Advanced Research Symposium on Topological Fluid Dynamics*. 1991.
- [89] Z.-S. She. Intermittency and non-Gaussian statistics in turbulence. *Fluid Dynamics Research*, 8:143–158, 1991.
- [90] B. Simon. *Functional Integration and Quantum Physics*. Academic Press, New York, 1979.
- [91] G. Sod. *Numerical Methods in Fluid Dynamics*. Cambridge University Press, Cambridge, 1985.
- [92] S. L. Soo. *Particulates and Continuum: Multiphase Fluid Dynamics*. Hemisphere, New York, 1989.
- [93] K. Sreenivasan. On the universality of the Kolmogorov constant. *Phys. Fluids*, 7:2778–2784, 1995.
- [94] S. Tanveer. Singularities in water waves and Rayleigh-Taylor instability. *Proc. R. Soc. Lond. A*, 435:137–158, 1991.
- [95] S. Tanveer. Evolution of hele-shaw interface for small surface tension. *Phil. Trans. Roy. Soc. Lond. A*, 343:155–204, 1993.
- [96] S. Tanveer. Singularities in the classical Rayleigh-Taylor flow: Formation and subsequent motion. *Proc. Royal Soc. London A*, 441:501–525, 1993.
- [97] J. F. Thompson, Z. U. A. Warsi, and C. Wayne Mastin. *Numerical Grid Generation*. North Holland, New York, 1985.
- [98] M. Vetter and B. Sturtevant. Experiments on the Richtmyer-Meshkov instability of an air / SF₆ interface. *Shock Waves*, 4:247–252, 1995.
- [99] G. Wallis. *One-Dimensional Two-Phase Flow*. McGraw-Hill, New York, 1969.
- [100] T. Wallstrom. Turbulent diffusion phase transition is due to viscosity-dependent energy spectrum. *Proc. NAS*, 92:11005–11008, 1995.
- [101] N. Whitaker and B. Turkington. Maximum entropy states for rotating vortex patches. *Phys. Fluids*, 6:3963–3973, 1994.
- [102] P. Wilmott, J. Dewynne, and S. Howison. *Option Pricing*. Oxford Financial Press, 1993.
- [103] P. Wilmott, S. Howison, and J. Dewynne. *The Mathematics of Financial Derivatives*. Cambridge University Press, 1995.
- [104] K. Wilson. The renormalization group: critical phenomena and the kondo problem. *Rev. Mod. Phys.*, 47:773–840, 1975.
- [105] V. Yakhot and S. A. Orszag. Renormalization group analysis of turbulence. *Phys. Rev. Lett.*, 57:1722–1724, 1986.
- [106] Y. Yang and Q. Zhang. General properties of a multi-layer stratified fluids system. *Phys. Fluids A*, 5:1167–1181, 1993.
- [107] Y. Yang, Q. Zhang, and D. H. Sharp. Small amplitude theory of Richtmyer-Meshkov instability. *Phys. Fluids*, 6(5):1856–1873, 1994.
- [108] D. L. Youngs. Numerical simulation of turbulent mixing by Rayleigh-Taylor instability. *Physica D*, 12:32–44, 1984.
- [109] M. Zeybek and Y. C. Yortsos. Parallel flow in hele-shaw cells. *J. Fluid Mech.*, 241:421–442, 1992.
- [110] Q. Zhang. Validation of the chaotic mixing renormalization group fixed point. *Phys. Lett. A*, 151:18–22, 1990.
- [111] Q. Zhang. The motion of single bubble or spike in Rayleigh-Taylor unstable interfaces. *Impact Comput. Sci. Engrg.*, 3:277–304, 1991.
- [112] Q. Zhang. A multi-length scale theory of the anomalous mixing length growth for tracer flow in heterogeneous porous media. *J. Stat. Phys.*, 66:485–501, 1991.
- [113] Q. Zhang. The asymptotic scaling behavior of mixing induced by a random velocity field. *Adv. Appl. Math.*, 16:23–58, 1995.
- [114] Q. Zhang. The transient behavior of mixing induced by a random velocity field. *Water Research Resources*, 31:577–591, 1995.
- [115] Q. Zhang and J. Glimm. Inertial range scaling of laminar shear flow as a model of turbulent transport. *Comm. Math. Phys.*, 146:217–229, 1992.
- [116] Q. Zhang and S. Sohn. An analytical nonlinear theory of Richtmyer-Meshkov instability. *Phys. Lett. A*, 212:149–155, 1996.

- [117] Q. Zhang and S. Sohn. Nonlinear solutions of unstable fluid mixing driven by impulsive force. Report No. SUNYSB-AMS-96-09, State Univ. of New York at Stony Brook, 1996.
- [118] Q. Zhang and S. Sohn. Nonlinear solutions of unstable fluid mixing driven by shock waves. *Phys. Fluids* **9** (1997) 1106-1124.
- [119] Q. Zhang and S. Sohn. Padé approximation to an inter-facial fluid mixing problem. *Applied Mathematics Letters*, 1996.
- [120] Q. Zhang and S. Sohn. Quantitative theory of Richtmyer-Meshkov instability in three dimensions. *ZAMP* (accepted for publication)
- [121] Q. Zhang and S. Sohn. Spike and bubble dynamics in Richtmyer-Meshkov instability. Report No. SUNYSB-AMS-96-10, State Univ. of New York at Stony Brook, 1996.

CHAPTER 2

Measure-Valued Processes and Renormalization of Branching Particle Systems

DONALD A. DAWSON

The Fields Institute
222 College Street
Toronto, Ontario
M5T 3J1 CANADA

ddawson@fields.utoronto.ca

EDWIN A. PERKINS
University of British Columbia
Department of Mathematics
Vancouver, BC
V6T 1Z2 CANADA
perkins@math.ubc.ca

1. Branching and Interacting Particle Systems.

1.1. Introduction. A *branching particle system (BPS)* consists of a finite or possibly infinite number of types of particles distributed in a space E in which the particles branch independently of each other, that is, die and produce a random number of offspring. Furthermore, in the time intervals between branching events the particles alive in this interval move according to an exchangeable Markov process and the motion of any subset of particles is not influenced by the number or locations of other particles present. An important aspect of BPS is the “family structure” based on the “particle-offspring” relationship and the complete description of such a BPS must include this family-structure information. A *branching and interacting particle system (BIPS)* has the same structure except that both the branching of particles and their motion can depend on the number and location of the other particles. Branching and interacting particle systems can model systems which arise in a wide variety of applications in the physical, life, social and information sciences but the analysis of such systems is difficult. By contrast (noninteracting) branching particle systems have a rich but tractable mathematical structure which has been the object of study since the original works of Galton and Watson. One line of research in this direction which has evolved rapidly in recent years is the study of super Brownian motion (SBM). SBM is a measure-valued Markov process which is obtained as the renormalized limit of the empirical

measure of branching Brownian motions in \mathbb{R}^d . Among the main results of this theory is the identification of both small scale and large scale spatial and space-time structures which are a result of the tendency of the particle motions to produce uniform distributions in space on the one hand and the tendency of the branching mechanism to produce local clusters of mass on the other. An important direction of recent research is to develop tools to study the small scale and large scale structures of the more complex and difficult classes of BIPS. The first challenge is to give a precise mathematical characterization to the limiting measure-valued processes which arise as renormalized limits of BIPS. A second is to obtain information concerning the small and large scale spatial and space-time structures which occur in these systems.

One direction of research is to explore to what extent the detailed information which has been obtained for the tractable classes such as SBM is robust and may give information and insight into these more complex models. The purpose of these notes is to give a brief introduction to some of the basic tools and to give an overview of some recent directions on research in this program.

1.2. Non-interacting Branching Particle Systems.

- Let E be a Polish space.
- If S is a topological space (b) (p) $\mathcal{B}(S)$ denotes the (bounded) (nonnegative) Borel measurable functions on S . We also write $\mathcal{E} = \mathcal{B}(E)$.
- $M_F(E)$ denotes the space of finite measures on (E, \mathcal{E}) with topology of weak convergence.
- $M(E)$ is the space of measures on (E, \mathcal{E}) .
- $M_1(E)$ is the space of probabilities on (E, \mathcal{E}) .
- $\mu(f) = \int f d\mu$, $f \in p\mathcal{E}$, $\mu \in M_F(E)$.
- $M^{(\eta)}(E)$ are measures on E consisting of a finite number of atoms of mass η .
- $\Omega_X = C(\mathbb{R}^+, M_F(E))$, the space of continuous $M_F(E)$ -valued functions equipped with the compact-open topology.
- $\Omega_X^d = D(\mathbb{R}^+, M_F(E))$ is the corresponding space of right-continuous $M_F(E)$ -valued paths with left limits, with the Skorokhod J_1 -topology.
- $X_t(\omega) = \omega(t)$ are the coordinate maps on Ω_X^d (or Ω_X).
- The canonical filtration on Ω_X is denoted by (\mathcal{F}_t^X) and \mathcal{P}^X is the σ -field of (\mathcal{F}_t^X) -predictable sets.
- $C_b^2(\mathbb{R}^d)$ is the space of bounded continuous \mathbb{R} -valued functions on \mathbb{R}^d with bounded continuous second partial derivatives.

The space E can model both the “type” and “spatial location” of individuals (particles). However we will begin with the description of a general class of BPS with one type of particle (i.e. identical particles).

Consider a system of particles of mass η which move, die and produce offspring in the space E . Since we assume that the particles are identical, their dynamics must be invariant under permutations of their labels (i.e. they must form an exchangeable particle system). We assume that in time intervals between branching events, the particle system is a non-interacting exchangeable system. That is, the motion of *any* subset $\{\xi_{\alpha_i}^n(t)\}_{i=1,\dots,n}$ is a given n -exchangeable (possibly time-inhomogeneous) Markov process on E^n , or equivalently, that between branching times the empirical

measure process

$$\Xi_t^\eta := \eta \sum_{\alpha \sim t} \delta_{\xi_\alpha(t)}$$

(sum over particles α alive at time t) is an $M_F^{(\eta)}(E)$ -valued time inhomogeneous Markov process. More precisely, we assume that the “space-time” process $(\Xi_t^\eta, t)_{t \geq 0}$ is a Hunt process with weak generator $(\mathcal{A}^\eta, D(\mathcal{A}^\eta))$. That is, $\mathcal{A}^\eta : D(\mathcal{A}^\eta) \subset \mathfrak{B} \rightarrow \mathfrak{B}$ (\mathfrak{B} is the class of bounded, measurable and finely continuous functions (c.f. Chapter I, Section 10 of Sharpe [112]) on $M_F^{(\eta)}(E) \times [0, \infty)$) and as ε decreases to 0

$$\frac{E(F(\Xi_{t+\varepsilon}^\eta, t + \varepsilon) | \Xi_t^\eta = \mu) - F(\mu, t)}{\varepsilon} \rightarrow \mathcal{A}^\eta F(\mu, t) =: (\mathcal{A}_t^\eta + \frac{\partial}{\partial t})F(\mu, t)$$

in the bounded pointwise sense (c.f. for example, Section 4 of Fitzsimmons [55]). In particular, we assume that the one particle motion is a Hunt process in $E \times [0, \infty)$ with weak generator $(A_t + \frac{\partial}{\partial t})$.

EXAMPLE 1.1. $\{\xi_{\alpha_i}^n(t)\}_{i=1,\dots,n}$ is a family of $n = \eta^{-1}$ independent Brownian motions in \mathbb{R}^d and no branching. Then the one particle motions are Brownian motions with generator $A_t \equiv A = \frac{1}{2}\Delta$, and $D(A) = C_b^2(\mathbb{R}^d)$. In this case $\Xi_t^\eta \xrightarrow{\eta \rightarrow 0} \Xi_t$, where Ξ_t is heat flow.

EXAMPLE 1.2. $\{\xi_{\alpha_i}^n(t)\}_{i=1,\dots,n}$ is given by the solution of the system of stochastic integral equations

$$\xi_{\alpha_i}^n(t) - \xi_{\alpha_i}^n(0) = \int_0^t \int_{\mathbb{R}^d} \dot{W}(ds, dy) g(y - \xi_{\alpha_i}^n(s)),$$

where $\dot{W} = (\dot{W}_1, \dots, \dot{W}_d)$ is d-dimensional space time white noise in $\mathbb{R}_+ \times \mathbb{R}^d$ and g is a C^2 function with compact support in \mathbb{R}^d .

Each particle ξ_α^n can die at a rate given $\gamma k(\xi_\alpha^n, dt)$, where $\gamma > 0$ and $k(\xi_\alpha^n, dt)$ is an additive functional of ξ_α^n (which by assumption follows the one particle motion).

EXAMPLE 1.3. In classical continuous time branching

$$k(\xi_\alpha^n, dt) = dt.$$

EXAMPLE 1.4. Catalytic branching Brownian motion with single point catalyst c is given by

$$k(\xi_\alpha^n, [0, t]) = \ell_c(\xi_\alpha^n, [0, t]),$$

where $\ell_c(\xi_\alpha^n, [0, t])$ denotes the local time at c of the one-dimensional Brownian motion ξ_α^n .

When a particle dies at a location x , with probability $p_\ell(x, 1)$ it is replaced by a *random cluster* of ℓ (exchangeable) offspring particles, $\ell = 0, 1, 2, \dots$, with law $\frac{p_\ell(x, dv)}{p_\ell(x, 1)} \in M_1(M_F^{(1)}(E))$. We assume that $\sum_{\ell=0}^{\infty} p_\ell(x, 1) = 1$, $\sum_{\ell} \ell^{1+\delta} p_\ell(x, 1) < \infty$ for some $\delta > 0$. The death times and number and location of offspring of particles are assumed to be independent of each other.

EXAMPLE 1.5. The classical critical binary branching particle system is given by

$$\begin{aligned} p_0(x, \cdot) &= \frac{1}{2}\delta_0 \\ p_2(x, \cdot) &= \frac{1}{2}\delta_{\delta_x + \delta_x} \\ p_k(x, \cdot) &= 0, k \neq 0, 2. \end{aligned}$$

We will consider mass and branching rate renormalization depending on a parameter $0 < \varepsilon \leq 1$. Let η_ε such that $\eta_\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0$ denote the particle mass and let γ_ε denote the particle branching rate. We denote the resulting particle system at time t by

$$X_t^\varepsilon = \eta_\varepsilon \sum_{\alpha \sim t} \delta_{\xi_\alpha(t)}.$$

The process X_t^ε is a $M_F(E)$ -valued Markov process with law $P_{m^*}^\varepsilon$ on $\Omega_X^d = D([0, \infty), M_F(E))$ which satisfies a measure-valued martingale problem:

For $F \in D(\mathcal{A}^{\eta_\varepsilon})$, $t \geq s$,

$$M_F^\varepsilon(t) := F(X_t^\varepsilon) - \gamma_\varepsilon \int_s^t \int \mathcal{B}^\varepsilon F(X_r^\varepsilon, x) K(dx, dr) - \int_s^t \mathcal{A}_r^{\eta_\varepsilon} F(X_r^\varepsilon) dr$$

is a P^ε -martingale, where

$$\mathcal{B}^\varepsilon F(\mu, x) = \sum_{\ell=0}^{\infty} \int_{M^{(\ell)}(E)} p_\ell(x, d\nu_\ell) [F(\mu + \eta_\varepsilon \nu_\ell - \eta_\varepsilon \delta_x) - F(\mu)]$$

and K is the *lifting* of k :

$$K(B \times [a, b]) = \sum_{\alpha} \int_a^b 1(\alpha \sim s) k(\xi_\alpha, ds) 1_B(\xi_\alpha(s)).$$

The process X_t^ε describes the evolution of the number and location of particles but does not contain information on the family structure of the population. In order to keep track of the latter we can identify the relationship between different particles as follows. Initial particles (generation zero) will be labeled by $\alpha_0 \in \mathbb{N}$. An n th generation particle is labeled $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_n) \in \mathbb{N}^{n+1}$ which means that it is the α_n th offspring of the $(n-1)$ st generation particle labeled $(\alpha_0, \alpha_1, \dots, \alpha_{n-1})$ and descends from the initial particle labeled α_0 (c.f. Walsh [122]). We will write $\alpha \sim t$ to denote that $\tau_b(\alpha) \leq t < \tau_d(\alpha)$, where τ_b and τ_d denote the birth and death times of particle α .

The first basic question we will address is the existence of the limit as $\varepsilon \rightarrow 0$ and the identification of the limiting process. Rather than attempting a general result we will illustrate the basic steps and some of the methods involved in a simple but important special case, namely, super-Brownian motion.

1.3. Super-Brownian Motion as the Limit of Branching Brownian Motion. Let us consider a branching particle system which between branching times forms a family of independent Brownian motions on $E = \mathbb{R}^d$ as in Example 2.1. Let $\mu(\phi) = \int \phi d\mu$. We can take $D(\mathcal{A}^\eta) = \{F(\mu) = f(\mu(\phi)) : \phi \in D(A), D(A) = C_b^2(\mathbb{R}^d), f \in C_b^2(\mathbb{R})\}$ and set $\mathcal{A}F(\mu) = f'(\mu(\phi))\mu(A\phi)$. Also we take $p_k(x, \cdot)$ as in Example 1.5 and, $k(dr) = dr$, $\eta_\varepsilon = \varepsilon$, $\gamma_\varepsilon = \frac{\gamma}{\varepsilon}$.

Then for $F \in D(\mathcal{A}^\varepsilon)$,

$$\begin{aligned} G_\varepsilon F(\mu) &= f'(\mu(\phi))\mu\left(\frac{1}{2}\Delta\phi\right) + \frac{\varepsilon}{2}f''(\mu(\phi))\mu(\nabla\phi \cdot \nabla\phi) \\ &\quad + \frac{\gamma}{2\varepsilon^2} \int [f(\mu(\phi) + \varepsilon\phi(x)) + f(\mu(\phi) - \varepsilon\phi(x)) - 2f(\mu(\phi))] \mu(dx). \end{aligned}$$

THEOREM 1.1. Let $P_{m_\varepsilon}^\varepsilon \in M_1(D([0, \infty), M_F(E))$ denote the law of this process. Assuming $X^\varepsilon(0) = m_\varepsilon \Rightarrow m$, then $P_{m_\varepsilon}^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} \mathbb{P}_m \in M_1(C([0, \infty), M_F(E)))$ and \mathbb{P}_m is the unique solution to the martingale problem (1.1): for all $\phi \in D(A)$,

$$(1.1) \quad M_t(\phi) := X_t(\phi) - m(\phi) - \int_0^t X_s(A\phi) ds$$

is an (\mathcal{F}_t^X) -martingale starting at zero with increasing process

$$\langle M(\phi) \rangle_t = \gamma \int_0^t X_s(\phi^2) ds.$$

Proof:

Step 1. Tightness on Ω_X^d and continuity. In order to prove tightness it suffices to prove

for $\delta > 0 \exists$ compact K and $L > 0$ such that

$$\begin{aligned} P_{m_\varepsilon}^\varepsilon(\sup_{0 \leq t \leq T} X_t(K^c) > \delta) &< \delta, \quad P_{m_\varepsilon}^\varepsilon(\sup_{0 \leq t \leq T} X_t(1) > L) < \delta \\ P_{m_\varepsilon}^\varepsilon \circ (X_t(\phi))^{-1} \text{ is tight in } D([0, \infty), \mathbb{R}) \text{ for smooth } \phi \text{ with compact support.} \end{aligned}$$

This can be checked by standard moment and martingale inequality arguments. If \mathbb{P}_m is a limit point, it is also easy to check that for ϕ in $D(A)$, $M_t(\phi)$ is a \mathbb{P}_m -martingale and $(F_1(\mu) = \mu(\phi), F_2(\mu) = \mu(\phi)^2)$

$$\begin{aligned} \langle M(\phi) \rangle_t &= \lim_{\varepsilon \rightarrow 0} \int_0^t (G_\varepsilon F_2(X_s) - 2F_1(X_s)G_\varepsilon F_1(X_s)) ds \\ &= \gamma \int_0^t X_s(\phi^2) ds. \end{aligned}$$

Using (1.1) and standard Ito formula arguments one sees that an equivalent formulation of the martingale problem is: for $f \in C_b^2(\mathbb{R})$, $\phi \in D(A)$, and $F(\mu) = f(\mu(\phi))$,

$$(1.2) \quad F(X_t) - \int_0^t GF(X_s) ds \text{ is a } \mathbb{P}_m \text{-martingale,}$$

where

$$GF(\mu) = f'(\mu(\phi))\mu(A\phi) + \frac{\gamma}{2}f''(\mu(\phi))\mu(\phi^2).$$

From this it is easy to prove a.s. continuity. (e.g. one can check the Bakry-Emery [3] conditions: $R(F_1 F_2, F_3) = F_1 R(F_2, F_3) + F_2 R(F_1, F_3)$, where $R(F_1, F_2) := G(F_1 F_2) - F_1 GF_2 - F_2 GF_1$).

Step 2. (Uniqueness) Applying (1.2) to $\exp(-\mu(\phi))$ with $\phi \in D(A)$, we obtain the exponential martingale problem (1.3)

$$(1.3) \quad \mathcal{E}_t(\phi) = \exp(-X_t(\phi)) + \int_0^t \exp(-X_s(\phi)) X_s(A\phi) ds - \frac{\gamma}{2} \int_0^t \exp(-X_s(\phi)) X_s(\phi^2) ds$$

which is a \mathbb{P}_m -martingale.

From this and following the argument in Leduc [77] we will next verify \mathbb{P}_m also solves the following time dependent martingale problem.

Step 2a. If $\psi : [0, \infty) \times E \rightarrow \mathbb{R}$ is smooth for A , then the integral martingale problem (1.4)

$$(1.4) \quad \begin{aligned} & \exp(-X_t(\psi_t)) + \int_0^t \exp(-X_s(\psi_s)) X_s((A + \frac{\partial}{\partial s})\psi_s) ds \\ & - \frac{\gamma}{2} \int_0^t \exp(-X_s(\psi_s)) X_s(\psi_s^2) ds \end{aligned}$$

is a \mathbb{P}_m -martingale. (The function $\psi : \mathbb{R}_+ \times E \rightarrow \mathbb{R}_+$ is *smooth for A* if $\psi(s, \cdot) \in D(A)$, $\frac{\partial}{\partial s}\psi(s, x)$ exists,

$$\left\| \frac{\psi(s+h, \cdot) - \psi(s, \cdot)}{h} - \frac{\partial}{\partial s}\psi(s, \cdot) \right\|_{\infty} \xrightarrow{h \downarrow 0} 0$$

and ψ , $\frac{\partial}{\partial s}\psi$ and $A\psi$ are bounded and strongly continuous.)

Proof:

Let $u(s, X_t) = \exp(-X_t(\psi_s))$, $v(s, X_t) = \exp(-(X_t(\psi_s))X_t(\frac{\partial}{\partial s}\psi_s))$, $w(s, X_t) = \exp(-X_t(\phi))(X_t(A\psi_s))$. Let $\mathbb{P}_m^{\mathcal{F}_t}$ denote the conditional expectation with respect to \mathcal{F}_t under \mathbb{P}_m . First note that for $t_2 > t_1$ using (1.3) we obtain

$$(1.5) \quad u(t_2, X_{t_2}) - u(t_1, X_{t_2}) = - \int_{t_1}^{t_2} v(s, X_{t_2}) ds$$

$$(1.6) \quad \begin{aligned} \mathbb{P}_m^{\mathcal{F}_{t_1}}[u(t_1, X_{t_2}) - u(t_1, X_{t_1})] &= -\mathbb{P}_m^{\mathcal{F}_{t_1}} \left[\int_{t_1}^{t_2} w(t_1, X_s) ds \right] \\ &+ \frac{\gamma}{2} \mathbb{P}_m^{\mathcal{F}_{t_1}} \left[\int_{t_1}^{t_2} u(t_1, X_s) X_s(\psi_s^2) ds \right]. \end{aligned}$$

Let Λ^n be a partition of $[t_1, t_2]$ with $\text{mesh}(\Lambda^n) \rightarrow 0$ and

$$\begin{aligned} \psi^n(s, x) &:= \sum_{i=1}^n \psi(t_i^n, x) 1_{[t_i^n, t_{i+1}^n)}(s) \\ X^n(s) &:= \sum_{i=1}^n X_{t_{i+1}^n} 1_{[t_i^n, t_{i+1}^n)}(s) \end{aligned}$$

Let $u^n(t, X_t) := \exp(-X_t(\psi_t^n))$. Then by (1.5) and (1.6), we have

$$\begin{aligned} \mathbb{P}_m^{\mathcal{F}_{t_1}}[u^n(t_2, X_{t_2}) - u^n(t_1, X_{t_1})] &= -\mathbb{P}_m^{\mathcal{F}_{t_1}}\left[\int_{t_1}^{t_2} \exp(-X_s^n(\psi_s)) X_s^n \left(\frac{\partial}{\partial s}\psi_s\right) ds\right] \\ &\quad -\mathbb{P}_m^{\mathcal{F}_{t_1}}\left[\int_{t_1}^{t_2} \exp(-X_s(\psi_s^n)) X_s(A\psi_s^n) ds\right] \\ &\quad + \frac{\gamma}{2} \mathbb{P}_m^{\mathcal{F}_{t_1}}\left[\int_{t_1}^{t_2} \exp(-X_s(\psi_s^n)) X_s((\psi_s^n)^2) ds\right]. \end{aligned}$$

Standard arguments show that this converges to

$$\begin{aligned} \mathbb{P}_m^{\mathcal{F}_{t_1}}[u(t_2, X_{t_2}) - u(t_1, X_{t_1})] &= -\mathbb{P}_m^{\mathcal{F}_{t_1}}\left[\int_{t_1}^{t_2} \exp(-X_s(\psi_s)) X_s\left(\frac{\partial}{\partial s}\psi_s\right) ds\right] \\ &\quad -\mathbb{P}_m^{\mathcal{F}_{t_1}}\left[\int_{t_1}^{t_2} \exp(-X_s(\psi_s)) X_s(A\psi_s) ds\right] \\ &\quad + \frac{\gamma}{2} \mathbb{P}_m^{\mathcal{F}_{t_1}}\left[\int_{t_1}^{t_2} \exp(-X_s(\psi_s)) X_s((\psi_s)^2) ds\right] \end{aligned}$$

which completes the proof of (1.4).

Step 2b. Let $v_t = V_t\phi$ be the unique solution (c.f. Pazy [97]) of the log Laplace equation (1.7)

$$(1.7) \quad \frac{\partial v_t}{\partial t} = Av_t - \frac{\gamma}{2}v_t^2, \quad v_0 = \phi \in pD(A).$$

Then v_t is smooth for A. Applying (1.4) we deduce that $\{\exp(-X_s(v_{t-s}))\}_{0 \leq s \leq t}$ is a martingale. Equating mean values at $s = 0$ and $s = t$ we get the fundamental equality (i.e. Laplace functional)

$$(1.8) \quad \mathbb{P}_m(\exp(-X_t(\phi))) = \exp(-m(V_t\phi)).$$

The extension for $\phi \geq 0$ in $D(A)$ to $\phi \geq 0$ in $b\mathcal{E}$ follows easily by considering the “weak form”

$$V_t\phi = P_t\phi - \frac{\gamma}{2} \int_0^t P_{t-s}(V_s\phi)^2 ds$$

and then taking bounded pointwise limits. (1.8) proves the uniqueness of the law $\mathbb{P}_m(X_t \in \cdot)$ for any solution of (1.1) and hence the uniqueness of \mathbb{P}_m (c.f. Theorem 4.4.2 of El Karoui and Roelly [47]). ■

COROLLARY 1.1. (*The total mass process*) (a) *The total mass process $X_t(1)$ is a Feller branching FB process. That is, a \mathbb{R}^+ -valued martingale and also a Markov diffusion with transition Laplace transform*

$$(1.9) \quad \mathbb{P}_m(\exp(-\lambda X_t(1))) = \exp\left(-\frac{2\lambda m(1)}{(2 + \lambda\gamma t)}\right).$$

(b) $\mathbb{P}_m(X_T(E) = 0) = \exp(-\frac{2m(E)}{\gamma T})$. Thus the process becomes extinct in finite time.

Proof:

If $v_0 \equiv \lambda$, then $v_t(x) = \frac{2\lambda}{(2 + \lambda\gamma t)}$ and so (1.8) implies (a). (Alternatively set $\phi = 1$ in (1.4) to see that $X_t(1)$ is the unique solution of $dX_t = \sqrt{\gamma X_t} dB_t$, where B_t is a

one-dimensional Brownian motion.) Letting $\lambda \rightarrow \infty$ we obtain the extinction law (1.10)

$$(1.10) \quad \mathbb{P}_m(X_t(E) = 0 \forall t \geq T) = \mathbb{P}_m(X_T(E) = 0) = \exp\left(-\frac{2m(E)}{\gamma T}\right). \blacksquare$$

COROLLARY 1.2. (*Weak and Strong Multiplicative Properties*)

$$(1.11) \quad \mathbb{P}_{m_1+m_2}(\exp(-X_t(\phi))) = \mathbb{P}_{m_1}(\exp(-X_t(\phi)))\mathbb{P}_{m_2}(\exp(-X_t(\phi)))$$

$$(1.12) \quad -\log[\mathbb{P}_m(\exp(-X_t(\phi)))] = \int (-\log[\mathbb{P}_{\delta_x}(\exp(-X_t(\phi)))]m(dx)$$

\mathbb{P}_m is infinitely divisible.

1.4. The Poisson Cluster Representation. Let X be an infinitely divisible random measure on a Polish space E with $P[X(E)] < \infty$. Then (c.f. Kallenberg [70]) there exists a measure $X_D \in M_F(E)$ and a measure $R \in M(M_F(E) \setminus \{0\})$ satisfying

$$\int (1 - e^{-\mu(E)})R(d\mu) < \infty$$

and such that

$$-\log(P(e^{-X(\phi)})) = X_D(\phi) + \int (1 - e^{-\nu(\phi)})R(d\nu).$$

X_D is called the deterministic component and R is called the *canonical measure*. For example, for a Poisson random measure with intensity Λ , $R(d\nu) = \int \delta_{\delta_x}(d\nu)\Lambda(dx)$. If we replace each Poisson point, x , with a random measure (*cluster*) with probability law $R(x, d\nu)$ then we obtain

$$R(d\nu) = \int \Lambda(dx)R(x, d\nu).$$

In the case of super-Brownian motion, X_t for $t > 0$ is infinitely divisible with $X_D = 0$ and canonical measure $R_t(d\nu) = \int m(dx)R_t(x, d\nu)$, where

$$V_t\phi(x) = \int (1 - e^{-\nu(\phi)})R_t(x, d\nu), \quad \phi \geq 0$$

(c.f. for example, Dawson and Perkins [30]). Since $(V_t\theta)(x) = \frac{2\theta}{(2+\theta)\gamma t}$, $R_t(x, M_F(\mathbb{R}^d) \setminus \{0\}) = \lim_{\theta \rightarrow \infty} (V_t\theta)(x) = \frac{2}{\gamma t}$ and

$$\int e^{-\theta\nu(1)}R_t(x, d\nu) = \frac{\left(\frac{2}{\gamma t}\right)^2}{\left(\frac{2}{\gamma t}\right) + \theta}.$$

Hence $R_t(x, \nu(1) \in \cdot)$ is $\frac{2}{\gamma t}$ times an exponential law with mean $\frac{\gamma t}{2}$. Then using the above and (1.10) one easily sees that

$$\int e^{-\theta\nu(\phi)}R_t(x, d\nu) = \lim_{\varepsilon \downarrow 0} \varepsilon^{-1} \mathbb{P}_{\varepsilon\delta_x}(e^{-\theta X_t(\phi)} 1(X_t > 0)).$$

Hence $R_t(x, \cdot)$ can be interpreted as the unnormalized distribution of X_t starting with infinitesimal mass at x (when it is nonzero).

1.5. Excursions. $w \in C([0, \infty), M_F(E))$ is an $M_F(E)$ -valued excursion of length σ starting at x if

$$\begin{aligned} w_0 &= 0, \quad \sigma(w) > 0, \quad w_t \neq 0, \quad 0 < t < \sigma(w), \quad w_t = 0, \quad t > \sigma(w) \\ w_t(1)^{-1}w_t &\rightarrow \delta_x \text{ as } t \rightarrow 0+. \end{aligned}$$

Let $C_x^0 = C_x^0([0, \infty), M_F(E) \setminus \{0\})$ denote the set of all excursion paths.

THEOREM 1.2. (*El Karoui and Roelly [47], Li and Shiga [87]*) *There is a unique σ -finite kernel $R(x, \cdot)$ from E to C_x^0 such that*

- (i) $x \rightarrow R(x, \cdot)$ is supported by C_x^0 .
- (ii) $\int w_t(1)R(x, dw) \rightarrow 1$ as $t \rightarrow 0+$.
- (iii) The measure $R(x, \cdot)$ is Markov with the same transition laws as super-Brownian motion and has t -marginal distribution equal to $R_t(x, \cdot)$.
- (iv)

$$E_m(e^{-X_t(\phi)}) = \exp\left(-\int \int 1_{\sigma(w) \geq t}(1 - e^{-w_t(\phi)})R(x, dw)m(dx)\right).$$

REMARK 1.1. $R(x, \{\sigma > t\}) = \frac{2}{\gamma t}$ and we can verify that

$$R(x, B \cap \{\sigma > t\}) = \lim_{\varepsilon \rightarrow 0+} \frac{P_{\varepsilon \delta_x}(B \cap \{X_t > 0\})}{\varepsilon}.$$

1.6. Relations with Stochastic Partial Differential Equations (SPDE).

The martingale $\{M_t(\phi) = X_t(\phi) - m(\phi) - \int_0^t X_s(A\phi)ds : \phi \in D(A)\}$ can be extended to a orthogonal martingale measure $M(ds, dx)$ on $[0, \infty) \times E$ such that $M_t(\phi) = \int_0^t \int_E \phi(x)M(ds, dx)$. Then as in the usual construction of the stochastic integral we may also define $M_t(f) = \int_0^t \int_E f(s, \omega, x)M(ds, dx)$ for f , a $\mathcal{P} \times \mathcal{E}$ -measurable function on $\mathbb{R}_+ \times \Omega \times E$ such that

$$\int_0^t \int_E f(s, \omega, x)^2 X_s(dx)ds < \infty \quad \forall t > 0 \quad \mathbb{P}_m a.s..$$

See Chapter 2 of Walsh [122] for details.

1.6.1. Super-Random Walk on \mathbb{Z}^d . A process closely related to, but in some ways simpler than super-Brownian motion is the ξ -superprocess, where ξ is the continuous time random walk on \mathbb{Z}^d , which after an exponential time of mean 1, chooses one of the $2d$ neighbors of its current state and jumps there. If $\{e_1, \dots, e_d\}$ is the standard basis in \mathbb{R}^d , the generator of ξ is the discrete Laplacian

$$A\phi(x) = \frac{1}{2}\Delta_1\phi(x) = \frac{1}{2d} \sum_{i=1}^d (\phi(x + e_i) + \phi(x - e_i) - 2\phi(x)), \quad x \in \mathbb{Z}^d.$$

The resulting super-random walk with parameter γ is denoted by $\{X_t(x) : x \in \mathbb{Z}^d\}$. Using standard arguments we can then construct (on an enrichment of Ω_X) a set of independent one-dimensional Brownian motions $\{B^x : x \in \mathbb{Z}^d\}$ such that $\{X_t(x)\}$ satisfies the stochastic differential equation (sde)

$$(1.13) \quad X_t(x) = m(x) + \int_0^t \sqrt{\gamma X_s(x)}dB_s^x + \int_0^t \left(\frac{1}{2}\Delta_1 X_s\right)(x)ds, \quad x \in \mathbb{Z}^d,$$

and

$$B_t^x = \int_0^t 1_{\{X_s(x) > 0\}} \frac{1}{\sqrt{\gamma X_s(x)}} dM_s(x) + \int_0^t 1_{\{X_s(x) = 0\}} d\tilde{B}_s^x,$$

where $\{\tilde{B}^x\}_{x \in \mathbb{Z}^d}$ are Brownian motions independent of X . Pathwise uniqueness to (5.3) can be proved as in Yamada and Watanabe [126] (c.f. Shiga and Shimizu [116]).

Let $X_t^\varepsilon(x) = X_t(x\varepsilon^{-1/2})$ for $x \in \varepsilon^{1/2}\mathbb{Z}^d$ and define

$$\frac{1}{2}\Delta_\varepsilon f(x) = \frac{1}{2d\varepsilon} \sum_{i=1}^d (f(x + \varepsilon^{1/2}e_i) + f(x - \varepsilon^{1/2}e_i) - 2f(x)), \quad x \in \varepsilon^{1/2}\mathbb{Z}^d.$$

The rescaled measure-valued process $X_t^\varepsilon(f) = \sum_x X_t^\varepsilon(x)f(x)$ satisfies (1.1) with $E = \varepsilon^{1/2}\mathbb{Z}^d$, $A = \frac{1}{2}\Delta_\varepsilon$ and $D(A)$ equal to the set of bounded measurable functions on $\varepsilon^{1/2}\mathbb{Z}^d$. Since $\lim_{\varepsilon \downarrow 0} \frac{1}{2}\Delta_\varepsilon f(x) = \frac{1}{2}\Delta f(x)$ if f is C^2 , one expects the law of X^ε to converge to the law of super-Brownian motion on \mathbb{R}^d provided that the initial laws converge. This is easy to check (c.f. Section 2 of Dawson [19], Blount [8]).

1.6.2. The SPDE in \mathbb{R} . In the case $E = \mathbb{R}$, independently Reimers [106] and Konno and Shiga [72] proved the following.

THEOREM 1.3. $X(t, dx) = X(t, x)dx \forall t \geq 0$, \mathbb{P}_m -a.s., where $X(t, x)$ is jointly continuous in t and x and one can define a white noise, W on $\mathbb{R}_+ \times \mathbb{R}$ (an enrichment of Ω_X) so that $X(t, x)$ satisfies the stochastic partial differential equation

$$(1.14) \quad dX(t) = \frac{1}{2}\Delta X(t)dt + \sqrt{\gamma X(t)}dW(t)$$

in the sense of tempered distributions. That is, for every $\phi \in \mathcal{S}(\mathbb{R})$,

$$(1.15) \quad \langle X_t, \phi \rangle = \langle X_0, \phi \rangle = \int_0^t \int_{\mathbb{R}} \sqrt{\gamma X_s(x)} \phi(x) \dot{W}_s(x) dx ds + \int_0^t \langle X_s, \frac{1}{2}\Delta \phi \rangle ds$$

and $\langle \cdot, \cdot \rangle$ denotes the canonical bilinear form on $\mathcal{S} \times \mathcal{S}'$.

Proof:

We sketch a heuristic proof. First consider an arbitrary dimension d . Let $\phi_\varepsilon := p_\varepsilon$ (the Brownian transition density) and note that $\lim_{\varepsilon \rightarrow 0} \phi_\varepsilon(\cdot - x) = \delta_x(\cdot)$ (the Dirac delta function). We obtain the density at x as the L^2 limit of $X_t(\phi_\varepsilon(\cdot - x))$. Applying the time dependent extension of (1.1) obtained in the proof of Theorem 1.1, we get

$$\begin{aligned} X_t(\phi_\varepsilon(\cdot - x)) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \phi_\varepsilon(x - z) p_t(z - y) m(dy) \\ &\quad + \int_{\mathbb{R}} \int_0^t \int_{\mathbb{R}} \phi_\varepsilon(x - z) p_{t-r}(y - z) dz M(dr, dy), \end{aligned}$$

where $M(dr, dy)$ is the orthogonal martingale measure obtained from $\{M_t(\phi)\}$. We can then show that $\{X_t(\phi_{\frac{1}{n}}(\cdot - x))\}$ is a Cauchy sequence in L^2 if

$$\mathbb{E}_m \left(\int_0^t \int_{\mathbb{R}} p_{t-s}(y - x)^2 X_s(dy) ds \right) < \infty,$$

where $p_t(y) = (2\pi t)^{-\frac{d}{2}} \exp(-|y|^2/2t)$. By criticality, this is equivalent to

$$(1.16) \quad \int_0^t \int \int p_{t-s}(y-x)^2 p_s(y-z) m(dz) dy ds < \infty.$$

It is easy to see that for any non-zero initial measure m , this is satisfied if and only if $d = 1$. If $d = 1$, then using Kolmogorov's criteria, we can prove that the density $X(t, x) = \lim_{\epsilon \rightarrow 0} X_t(\phi_\epsilon(\cdot - x))$ has a version which is jointly continuous in x and $t > 0$. The construction of W in the case $d = 1$ is obtained by an analogue of the construction of the $\{B_t^x\}$ in the case of super-random walk. ■

REMARK 1.2. A similar quadratic variation calculation fails and strongly suggests that $t \rightarrow X_t(x)$ for fixed x is not a semi-martingale (when it is non-zero).

REMARK 1.3. Although we have shown uniqueness in law of solutions to (1.14), it is not known if any two solutions to (1.14) for a given white noise must coincide pathwise a.s. The large number of unsuccessful attempts at proving pathwise uniqueness leads us to guess it may not be true.

REMARK 1.4. If $d > 1$, the obstruction posed by (1.16) is real because super-Brownian motion is a singular measure (c.f. Theorem 4.1). If a solution to (1.14), $X_t(x)$, did exist for $d > 1$, then $X_t(dx) = X_t(x)dx$ would satisfy (1.1) and hence X_t would be a density for super-Brownian motion. This contradiction shows (1.14) has no continuous integrable solution.

1.6.3. Weak Convergence Approach in \mathbb{R} . An alternative approach is to establish that $X(t)$ satisfies an sde with driving Hilbert space valued martingale M . The construction of M starts with the martingale problem decomposition of the system and establishes the weak convergence of the processes and the representing martingale measures to (X, M) . This has been carried out in slightly different contexts by Blount [8] and Wang [121]. In [8], Blount considers the restriction to $E = [0, 1]$ and $A_{[0,1]} = \frac{1}{2}\Delta$ with periodic boundary conditions. Let $H_0 = L_2[0, 1]$ and consider the trigonometric basis: $\phi_n(r) = 1$ if $n = 0$, $\phi_n(r) = \sqrt{2}\cos(\pi nr)$ if n is even and $\psi_n(r) = \sqrt{2}\sin(\pi nr)$ if n is odd. Consider the scale of Hilbert norms: $\|f\|_\alpha = [\sum(< f, \phi_n >^2 + < f, \psi_n >^2)(1+n^2)^\alpha]^{1/2}$.

THEOREM 1.4. (Blount [8]) Let X be the $A_{[0,1]}$ -superprocess and let M be the martingale measure associated to (1.1). Then there is a version of (X, M) such that a.s.

- (a) $\{M(t)\}_{t \geq 0}(\cdot)$ lives in
 $C([0, T], H_{-\alpha}) \cap C((0, T], H_{-\alpha})$ for any $\alpha > \frac{3}{2}$, and
- (b) $X(\cdot) \in C([0, T], H_0) \cap C((0, T], H_\alpha)$ for any $\alpha < \frac{1}{2}$.

Furthermore for $X(0) \in L_2([0, 1])$, $X(t)$ satisfies the following Hilbert space-valued sde and stochastic evolution equation:

$$\begin{aligned} X(t) &= X(0) + \int_0^t \frac{1}{2} \Delta X(s) ds + M(t) \\ X(t) &= S_t X(0) + \int_0^t S_{t-s} dM(s). \end{aligned}$$

REMARK 1.5. Using Fourier transforms this approach can be extended to the case $E = \mathbb{R}$. Wang [121] has also derived the stochastic partial differential equation in a more general setting using a related weak convergence argument.

1.7. Extension to Non-Square Integrable Branching. Consider a spatially homogeneous branching particle system in which $p_k(x, \cdot) = p_k \delta_{k\delta_x}$ and

$$\Psi(z) := \sum_{k=0}^{\infty} p_k z^k = z + \gamma L(z)(1-z)^{1+\beta},$$

where $0 < \beta \leq 1$ and $L(z)$ is an arbitrary nonnegative slowly varying function as $z \uparrow 1$. Assume that the particle masses $\eta_\varepsilon = \varepsilon$ and the branching rate is $\gamma_\varepsilon = \frac{\gamma L(1-\varepsilon)}{\varepsilon^\beta}$. Then carrying out the scaling limit we obtain the log-Laplace equation

$$\begin{aligned} \frac{\partial v}{\partial t} &= Av - \frac{\gamma}{1+\beta} v^{1+\beta} \\ v(0) &= \phi \in pb\mathcal{E}. \end{aligned}$$

The resulting $M_F(\mathbb{R}^d)$ -valued process has discontinuous trajectories when $\beta < 1$ and is called the $(1+\beta)$ -branching superprocess.

1.8. Extension to Infinite Measures. Let $\phi_p(x) := (1+|x|^2)^{-\frac{p}{2}}$. Then the class of p -tempered measures is defined by

$$M_p(\mathbb{R}^d) := \{\mu : \|\mu\|_p := \mu(\phi_p) < \infty\}$$

furnished with the topology of p -weak convergence (c.f. Iscoe [67]).

The $(1+\beta)$ -super-Brownian motion with initial measure $m \in M_p(\mathbb{R}^d)$ is obtained as a monotone limit $X(t) = \lim_{n \rightarrow \infty} X_n(t)$ with $X_n(0, dx) = \mathbf{1}(\{|x| \leq n\})m(dx)$ and it has Laplace functional

$$\mathbb{P}_m(\exp(-X_t(\phi)) = \exp(-m(V_t \phi)))$$

for $\phi \in p\mathcal{E}$ with $\sup_x |\phi(x)|/(1+|x|^2)^{\frac{p}{2}} < \infty$, where $v(t, x) = V_t \phi(x)$

$$\begin{aligned} \frac{\partial v(t, x)}{\partial t} &= \frac{1}{2} \Delta v(t, x) - \frac{\gamma}{1+\beta} v(t, x)^{1+\beta} \\ v(0, x) &= \phi(x) \in p\mathcal{B}(\mathbb{R}^d). \end{aligned}$$

EXAMPLE 1.6. Lebesgue initial measure: $\lambda \in M_p(\mathbb{R}^d)$ if $p > d$.

1.9. Branching in a Single Point Catalyst. If we replace $k(\xi, dt) = dt$ by $\ell_c(\xi, dt)$, the local time at c of one-dimensional Brownian motion ξ , then we obtain the super-Brownian motion in a single point catalyst as in Example 2.5 (c.f. Dawson and Fleischmann [22]). In this case $X_t(db)$ has a density $x_t(b)$ which is continuous for $b \neq c$. Furthermore the occupation time process defined by $Y_t(db) := \int_0^t X_s(db)ds$ has a density $y_t(b)$ which is jointly continuous in (t, b) . We define the measure λ^c on $[0, \infty)$ by $\lambda^c([t, t+s)) := y_{t+s}(c) - y_t(c)$.

The log Laplace functional for the pair $(x_t(b), \lambda^c)$ is given by

$$\mathbb{P}_\mu[\exp(-\theta x_t(b) - \int_0^t \lambda^c(ds)g(s))] = \exp[-\int \mu(dx)u(0, t, x)],$$

where u satisfies

$$\begin{aligned} -\frac{\partial u(s, t, a)}{\partial s} &= \frac{1}{2}\Delta u(s, t, a) + \delta_c(a)g(s) - \frac{1}{2}\delta_c(a)u^2(s, t, a) \\ u(s, t, \cdot)|_{s=t-} &= \theta\delta_b. \end{aligned}$$

This arises from a branching particle system in which branching occurs according to a clock given by Brownian local time at c .

THEOREM 1.5. (*Dawson, Fleischmann, Li, and Mueller [25], Fleischmann and Le Gall [59]*) Let $X_0 = \delta_c$. Then with probability one the super-Brownian local time measure λ^c at c is singular and has carrying Hausdorff dimension one.

THEOREM 1.6. (*Fleischmann and Le Gall [59]*) The density field $x_t(\cdot)$ has the representation

$$\begin{aligned} x_t(a) &= \int_0^t \lambda^c(ds)q(t-s, |a-c|), \quad t > 0, \quad a \neq c \\ q(r, b) &= \frac{b}{\sqrt{2\pi r^3}} \exp\left(-\frac{b^2}{2r}\right), \end{aligned}$$

where $q(r, b)db$ is the Brownian excursion density.

1.10. Lattice Trees and Integrated Super-Excursion. A d -dimensional lattice tree with N bonds is a connected subset of \mathbb{Z}^d with N bonds and $(N+1)$ nodes, containing 0, and in which there are no loops. Here we are viewing \mathbb{Z}^d as a graph with possible bonds between neighbors a distance of one apart. Let P_N denote uniform measure on the set of such trees. Given a tree ω , we let $X_N(\omega)$ be the probability on \mathbb{R}^d obtained by putting mass $\frac{1}{N+1}$ on each vertex in the rescaled tree, $c(d)N^{-\frac{1}{4}}\omega$. Aldous [1] conjectured that if $d > 8$ then for an appropriate $c(d) > 0$, X_N should converge weakly to a random measure he called *integrated super-excursion ISE* and which we denote by X . In terms of super-Brownian motion (with $\gamma = 1$) the law of X may be described as follows. $Y(A) = \int_0^\infty X_s(A)ds$ defines an infinitely divisible random measure under \mathbb{P}_{δ_0} and has a canonical measure $\bar{R}(d\nu)$. (In fact, one can readily check that $\bar{R}(A) = R(0, \int_0^\infty w_s(\cdot)ds \in A)$ - see Theorem 1.2.) Then

$$\begin{aligned} (1.17) \quad P(X \in A) &= \bar{R}(\nu \in A | \nu(1) = 1) \\ &= w - \lim_{\varepsilon \downarrow 0} \bar{R}(\nu \in A | \nu(1) \in [1-\varepsilon, 1+\varepsilon]), \end{aligned}$$

where $w-\lim$ refers to the weak limit as $\varepsilon \downarrow 0$. This description of ISE follows from its construction for the continuous random tree CRT in Aldous [1], the description of the (p-marginals of the) CRT in terms of the normalized Brownian excursion (c.f. for example, Le Gall [81]), the construction of super-Brownian motion (and \bar{R}) via the path process and Itô's excursion measure, n (Le Gall [79]), and the elementary fact that the normalized Brownian excursion is the weak limit of $n(\cdot | 1-\varepsilon \leq \ell \leq 1+\varepsilon)$ as ε decreases to 0, where ℓ is the length of the excursion. In other words, this is the law of $\int_0^\infty X_s ds$ under the canonical measure starting at 0, conditioned on having total mass one. Recently Derbez and Slade [38] have proved Aldous' conjecture if d is sufficiently large or $d > 8$ and one works with "spread-out trees". In the latter model, one may introduce bonds between x and y in \mathbb{Z}^d if $\|x - y\| \leq L(d)$ for sufficiently large $L(d)$ and the rescaled tree is now $c(d, L)N^{-\frac{1}{4}}\omega$, where $c(d, L) \sim \frac{c(d)}{L}$ as $L \rightarrow \infty$. They use Brydges and Spencer's lace expansion method to establish the convergence of moment measures. This approach was also used by Hara and Slade [64] to establish weak convergence of rescaled self-avoiding walk to Brownian motion if $d > 4$. We return to this setting briefly in Section 4.4 to verify the fact that the support of ISE has Hausdorff dimension 4 if $d \geq 4$.

REMARK 1.6. The fact that the "ranges" of two independent super-Brownian motions just fail to intersect if $d = 8$, and do intersect if $d < 8$ (c.f. Section 4) means the closed support of X will be a tree if and only if $d \geq 8$ and leads to the critical value $d = 8$. (Recall the ranges of two independent Brownian motions just fail to intersect if $d = 4$, whence $d = 4$ is critical for self-avoiding random walk.)

REMARK 1.7. Hara and Slade [64] have also conjectured that at the critical point the scaling limit of large percolation clusters in dimensions $d > 6$ is also ISE (c.f. Derbez and Slade [38]).

2. Historical Brownian Motion

In order to obtain a more complete picture we must introduce the Historical Brownian motion studied independently by Dawson and Perkins, Donnelly and Kurtz, Dynkin, and Le Gall in slightly different contexts.

Given the Brownian motion ξ , the *path-valued process* $\bar{\xi}(t) = \xi(\cdot \wedge t)$ is a time-inhomogeneous continuous strong Markov process taking values in the Polish space $C = C(\mathbb{R}_+, E)$. We let $P_{r,y}$ (here $y(\cdot \wedge r) = y$) denote the law of $\bar{\xi}$ starting at y at time r . Historical Brownian motion is obtained by replacing ξ with the path-valued Brownian motion $\bar{\xi}$ in the general superprocess construction outlined above.

We can consider the corresponding empirical process

$$H_t^\varepsilon = \sum_{\alpha \sim t} \eta_\varepsilon \delta_{\bar{\xi}_\alpha(t)}.$$

By a limiting procedure similar to that in Section 1, we get the ξ -historical process $\{H_t\}$ and law $\mathbb{Q}_{\tau,m}$ on Ω_H . Let $y^t := y(\cdot \wedge t)$, $C^t := \{y : y = y^t\}$, and $M_F(C)^t = \{m \in M_F(C) : y^t = y \text{ m-a.e. } y\}$. Here \mathcal{C} is the Borel σ -field of C and \mathcal{C}_t is the sub- σ -field generated by the paths up to time t . Let

$$\begin{aligned} \Omega_H &= \{H \in C(\mathbb{R}_+, M_F(C)) : H_t \in M_F(C)^t \ \forall t \geq 0\}, \\ \mathcal{H}[s, t] &= \sigma(H_u : s \leq u \leq t), \quad \mathcal{H}[s, \infty) = \sigma(H_u : u \geq s). \end{aligned}$$

Let $C_{0,\infty}^d$ denote the space of continuous functions from $[0, \infty)$ to \mathbb{R}^d and $D(\bar{A})$ denote the subset of $f \in C_{0,\infty}^d$ of the form

$$f(y) = g(y(t_1), \dots, y(t_n)), \quad 0 \leq t_1 \leq t_2 \dots \leq t_n,$$

where g is infinitely differentiable and constant outside a compact set.

For $f \in D(\bar{A})$ and $t > 0$, let

$$\bar{A}_t f(y) := \frac{1}{2} \sum_{i=1}^d \sum_{k=0}^{n-1} \sum_{\ell=0}^{n-1} 1(t < t_{k+1} \wedge t_{\ell+1}) g_{x_{k+d+i}, x_{\ell d+i}}(y^i(t_1), \dots, y^i(t_n)).$$

THEOREM 2.1. (*Perkins [105]*) If $\tau \geq 0$ and $m \in M_F(C)^\tau$, $\mathbb{Q}_{\tau,m}$ is the unique law on $(\Omega_H, \mathcal{H}[\tau, \infty))$ such that

$$(2.1) \quad \text{for each } f \in D(\bar{A}), \quad M_t(f) = H_t(f) - m(f) - \int_\tau^t H_s(\bar{A}_s f) ds, \quad t \geq \tau$$

is an $(\mathcal{H}[\tau, \infty))$ -martingale, starting at zero when $t = \tau$, and with quadratic variation

$$\langle M(f) \rangle_t = \int_\tau^t \gamma H_s(f^2) ds.$$

The Laplace functional characterizing H has the form

$$\begin{aligned} \mathbb{Q}_{r,m}(\exp(-H_t(\phi))) &= \exp(-m(V_{r,t}\phi)) \\ V_{r,t}\phi(y) &= S_{r,t}\phi(y) - \frac{\gamma}{2} \int_r^t S_{r,s}(V_{s,t}\phi)^2 ds, \quad y \in C^s \\ S_{r,t}\phi(y) &= P_{r,y(r)}\phi(y/r/\xi(\cdot \wedge t)), \end{aligned}$$

where if $y, \xi \in C_{0,\infty}^d$, $r \in [0, \infty)$, then $(y/r/\xi) \in C_{0,\infty}^d$ is defined by

$$(y/r/\xi)(u) := \begin{cases} y(u) & \text{if } u < r \\ \xi(u-r) & \text{if } u \geq r \end{cases}.$$

COROLLARY 2.1. Let $\pi_t : C \rightarrow E$, $\pi_t(y) := y(t)$. Then

$\pi_t(H_t)$ is a version of the SBM X_t .

REMARK 2.1. If A_1, \dots, A_n are disjoint sets in \mathcal{C}_τ and $f_i(t, y) = 1_{A_i}(y)$, then clearly $f_i \in D(A)$ and $A_{\tau,m} f_i = 0$. (2.1) shows that under $\mathbb{Q}_{\tau,m}$, $X_t^i = H_{\tau+t}(A_i)$, $i = 1, \dots, n$, are independent FB processes starting at $m(A_i)$, $i = 1, \dots, n$. Therefore (c.f. Corollary 1.1)

$$(2.2) \quad \mathbb{Q}_{\tau,m}(\exp(-\sum_i^n \lambda_i H_{\tau+t}(A_i))) = \exp(-\sum_{i=1}^n 2\lambda_i m(A_i)/(2 + \lambda_i \gamma t)), \quad \lambda_i \geq 0$$

and

$$(2.3) \quad \mathbb{Q}_{\tau,m}(H_{\tau+s}(A_i) = 0 \ \forall s \geq t) = \exp(-2m(A_i)/(\gamma t)).$$

PROPOSITION 2.1. (*Historical Cluster Representation (HCP)*) Let $\tau \leq s < t$. Under $\mathbb{Q}_{\tau,m}$, the conditional distribution of $H_t(\{y : y^s \in \cdot\})$ given $\mathcal{H}[\tau, s]$ is the law of $\sum_{i=1}^M \delta_{y_i} m_i$, where $\{y_1, \dots, y_M\}$ are the points of a Poisson point process with intensity $H_s(\cdot)2\gamma^{-1}(t-s)^{-1}$ and $\{m_1, \dots, m_M\}$ are independent exponential masses with mean $\gamma(t-s)/2$.

Proof:

As above we may take $s = \tau$ and argue unconditionally. An easy calculation shows the Laplace functional of the above Poisson cluster random measure, Ξ , is

$$(2.4) \quad \mathbb{P}(\exp(-\Xi(\phi))) = \exp\left(-\int 2\phi(y)(2 + \phi(y)\gamma(t - \tau))^{-1} m(dy)\right).$$

If $\phi(y) = \sum_1^n \lambda_i 1_{B_i}(y)$ for B_1, \dots, B_n disjoint Borel sets in C , (2.2) implies

$\mathbb{Q}_{\tau,m}\left(\exp(-\int \phi(\bar{y}(\tau)) H_t(d\bar{y}))\right)$ is also given by the right side of (2.4). By taking limits of simple functions we see that the Laplace functionals of the random measures in question are equal and hence so are their laws. ■

2.0.1. Historical Modulus of Continuity. According to Corollary 2.1, the historical process is an enriched version of the super-Brownian motion and for this reason has turned out to be a powerful tool in the study of both the small scale and the large scale structure of the later. In this section we will establish a basic result on the historical modulus of continuity which is the main ingredient in extending fixed time results to a result that holds for all $t > 0$. In particular, this will be used in Section 4 to prove that the closed support of X_t has dimension 2 for every $t > 0$ a.s. if $d \geq 2$. Let

$$h(r) := \left(r \log^+ \frac{1}{r}\right)^{1/2}.$$

THEOREM 2.2. (*Historical modulus of continuity (2.5) of Dawson and Perkins [30]*) If $c > 2$ there are constants $C(d, c) > 0$, $\rho(c) > 0$ and a random variable $\delta(c, \omega) > 0$, $\mathbb{Q}_{0,m}$ -a.s. such that for every $t \geq 0$

$$(2.5) \quad S(H_t) \subset \{y : |y_r - y_s| \leq c h(|r - s|) \text{ whenever } r, s \in [0, t], |r - s| < \delta(c, \omega)\}$$

and

$$\mathbb{Q}_{0,m}(\delta(c) \leq r) \leq C(d, c) m(1) \gamma^{-1} r^{\rho(c)} \quad \forall r \in [0, 1].$$

If $c < 2$, (2.5) is false $\mathbb{Q}_{0,m}$ -a.s. .

REMARK 2.2. An analogous result has been proved in Dawson, Hochberg and Vinogradov [30] for the $(1 + \beta)$ -branching superprocess but with

$$c > \frac{2}{\beta}.$$

REMARK 2.3. For a single Brownian path the critical value of c is $\sqrt{2}$ by Lévy's well-known modulus of continuity.

REMARK 2.4. If $X_t(\cdot) = H_t(y_t \in \cdot)$ and $\pi_t : C \rightarrow \mathbb{R}^d$ is the projection onto the t^{th} coordinate, one would expect $S(X_t) = \pi_t(S(H_t))$ for every $t \geq 0$ a.s., and this is easy to show using the above theorem (c.f. Corollary 3.7 of Mueller and Perkins [91]). If $\delta > 0$, Proposition 2.1 shows that $\mathbb{Q}_{0,m}$ -a.s. for sufficiently large N , $H_\delta(\{y : |y_0| > N\}) = 0$ and therefore $H_t(\{y : |y_0| > N\}) = 0$ for every $t \geq \delta$ because this is a non-negative martingale by the Remark 2.1. This, the above

theorem, and the Arzela-Ascoli theorem show that $S(H_t)$ is compact in C for all $t > 0$ a.s. and hence the same is true of $S(X_t) = \pi_t(S(H_t))$.

REMARK 2.5. Proposition 2.1 implies that for a fixed $s < t$, $S(H_t(y(s) \in \cdot)) \subset S(X_s)$ a.s., and with a little more work one can show this inclusion holds for all $s < t$ outside a single null set (c.f. Proposition 8.11 of Dawson and Perkins [30]). This together with the Theorem 3.2 imply that if $c > 2$ and $0 \leq t - s < \delta(c, \omega)$, then each point of $S(X_t)$ is within $ch(t - s)$ of a point in $S(X_s)$. This finite speed of propagation holds more generally but may fail even when ξ is continuous (c.f. Section 8 of Dawson and Perkins [30]). If ξ is a symmetric stable process of index α and X is the ξ -superprocess, then $X_t = 0$ or $S(X_t) = \mathbb{R}^d$ a.s. for every $t > 0$ (c.f. Perkins [102]) and so propagation is instantaneous.

Proof:

We will not use the last assertion and so will only prove the positive assertion for $c > 2$. If $n \in \mathbb{N}$, (2.3) and the Markov property show that

$$\begin{aligned} & \mathbb{Q}_{0,m}(H_t(\{y : |y(j2^{-n}) - y((j-1)2^{-n})| > ch(2^{-n})\}) > 0 \\ & \quad \text{for some } t \geq (j+1)2^{-n} \text{ and } 1 \leq j \leq n2^{-n} \\ & \leq \sum_{j=1}^{n2^n} \mathbb{Q}_{0,m}(1 - \exp(-2H_{j2^{-n}}(\{y : |y(j2^{-n}) - y((j-1)2^{-n})| > ch(2^{-n})\})) \gamma^{-1} 2^n \\ & \leq \sum_{j=1}^{n2^n} \gamma^{-1} 2^{n+1} \mathbb{Q}_{0,m}(H_{j2^{-n}}(\{y : |y(j2^{-n}) - y((j-1)2^{-n})| > ch(2^{-n})\})) \\ & = \gamma^{-1} 2^{n+1} \sum_{j=1}^{n2^n} \int P_{0,y}(|B(j2^{-n}) - B((j-1)2^{-n})| > ch(2^{-n})) m(dy) \\ & \quad (\text{criticality property}) \\ & \leq C(d, c) \gamma^{-1} m(\mathbb{R}^d) n^{\frac{d}{2}+1} 2^{-n((c^2/2)-2)}, \end{aligned}$$

where a standard bound on the tail of the Gaussian distribution is used in the last line. If $c > 2$, the Borel-Cantelli lemma shows there is an $N(\omega) < \infty$ a.s. such that if $n \geq N(\omega)$, then

$$\begin{aligned} |y(j2^{-n}) - y((j-1)2^{-n})| & \leq ch(2^{-n}) \quad \text{for } 1 \leq j \leq n2^n, (j+1)2^{-n} \leq t \\ & H_t - \text{a.a. } y \quad \forall t \geq 0. \end{aligned}$$

It is now easy (but a bit tedious) to adapt a well-known binary expansion argument of Lévy to derive the result. ■

3. Formulation of a General Class of Measure-Valued Branching Processes

3.1. Dynkin's Characteristics. Super-Brownian motion and historical Brownian motion are examples of measure-valued Markov processes which have the “multiplicative” or “branching” property. The problem of characterizing Markov processes satisfying the branching property was initiated by Lamperti, Silverstein and Watanabe (c.f. [20] for more details) in the finite dimensional case. Recently

this program was extended to the measure-valued setting in the following important result of Dynkin, Kuznetsov and Skorohod [44]. First we give the formal definition.

DEFINITION 3.1. A *critical branching measure-valued process (BMV)* with metrizable Lusin base space E is a time-inhomogeneous $M_F(E)$ -valued Hunt process $X = (\Omega_X, \mathcal{F}, X, \{\mathbb{P}_{r,m}\}_{(r,m) \in \mathbb{R}_+ \times M_F(E)})$ satisfying the strong multiplicative property (1.12) and the criticality condition $\mathbb{P}_{r,m}(X_t(1)) = 1$ for all $r < t$ and $m \in M_F(E)$.

THEOREM 3.1. (Dynkin, Kuznetsov, Skorohod [44]) Let X be a Hunt BMV process and $\xi = (\xi_t, \mathcal{F}, \Pi_{r,x})$ be a (time-inhomogeneous) Hunt process with state space (E, \mathcal{E}) . Assume X and ξ are related by

$$\mathbb{P}_{r,\mu}(X_t(\phi)) = \int \mu(dx) \Pi_{r,x} \phi(\xi_t) \quad \forall \phi \in b\mathcal{E}$$

and that

$$\sup_{r < t} \mathbb{P}_{r,\mu}(X_t(1)^2) < \infty.$$

Define $v_{r,t}(x) := -\log \mathbb{P}_{r,\delta_x}(e^{-X_t(\phi)})$. Then the log-Laplace equation for X has the form

$$v_{r,t}(x) + \Pi_{r,x} \int_r^t v_{s,t}(\xi_s)^2 Q(ds) + \Pi_{r,x} \int_{(r,t]} \int_{M_F} \mathcal{E}(\mu(v_{s,t})) \ell(ds, d\mu) = \Pi_{r,x} \phi(\xi_t),$$

where

$$\mathcal{E}(u) := e^{-u} + u - 1,$$

and the branching characteristics Q and ℓ satisfy the following necessary conditions:

- (i) The additive functional Q is natural and $\sup_{x \in E} \Pi_{r,x} Q(r, t] < \infty$ for all $r < t$.
- (ii) The measure-valued additive functional ℓ (i.e. $t \rightarrow \ell(t, A)$ is an increasing additive functional for all A and $A \rightarrow \ell(t, A)$ is a measure for all t) is natural and satisfies

$$\sup_{x \in E} \Pi_{r,x} \int_r^t \int_{M_F} \mu(E)^2 \ell(ds, d\mu) < \infty.$$

- (iii) A metric d in E (in which ξ is a Hunt process) can be chosen in such a way that, for every $\varepsilon > 0$ and for all $r < t \in \mathbb{R}_+$, $m \in M_F(E)$,

$$\Pi_{r,m} \int_{(r,t]} \int_{M_F(E)} \mu[U_\varepsilon(\xi)] \ell(ds, d\mu) < \infty,$$

where $U_\varepsilon(y)$ is the ε -neighborhood of y .

- (iv) The measure

$$\bar{\ell}(A, B) = \int_{M_F} \mu(B) \ell(A, d\mu)$$

is dominated by the Lévy measure $N_\xi(ds, dy)$ of the process ξ in the sense that

$$\int 1_{\xi_{s-} \neq y} f^s(\xi_{s-}, y) \bar{\ell}(ds, dy) \leq \int f^s(\xi_{s-}, y) N_\xi(ds, dy) \quad a.s.$$

COROLLARY 3.1. *Let*

$$K(I) := \int_I \int_{M_F} \mu(E)^2 \ell(ds, d\mu) + Q(I).$$

Then we can find a measurable function $b^s(x)$ and a kernel $\ell_x^s(d\mu)$ satisfying

$$0 \leq b \leq 1, \quad 0 \leq \int_{M_F} \mu(E)^2 \ell_x^s(d\mu) \leq 1$$

such that

$$\begin{aligned} Q(ds) &= b^s(\xi_s) K(ds) \\ \ell(ds, d\mu) &= \ell_{\xi_s}^s(d\mu) K(ds). \end{aligned}$$

Putting

$$\Phi(s, x, z) := b^s(x)(z(x))^2 + \int_{M_F} \mathcal{E}(\mu(z)) \ell_x^s(d\mu),$$

we get

$$v_{r,t}(x) = \Pi_{r,x} \phi(\xi_t) - \Pi_{r,x} \int_r^t \Phi(s, \xi_s, v_{s,t}) K(ds).$$

We call the resulting process the (ξ, Φ, K) -superprocess.

REMARK 3.1. It is easy to obtain existence when the conditions involving $\mu(E)^2$ are replaced by the same conditions with $\mu(E)^2 \wedge \mu(E)$. (The idea is simply to truncate

$$\ell_\kappa(ds, d\mu) = 1_{[0, \kappa]}(\mu(E)) \ell(ds, d\mu)$$

and then take limits and use the fact that the mean measures are bounded.)

3.2. The Martingale Problem Formulation. In this subsection we assume that ξ is a time homogeneous Hunt process with strongly continuous semigroup $\{S_t\}_{t \geq 0}$ on $L \subset b\mathcal{E}$ with infinitesimal generator $(A, D(A))$. We assume that L is an algebra. In this section we present the characterization of the (ξ, Φ, K) -superprocess as the unique solution to a martingale problem. Following Leduc [77] this martingale problem is formulated in two steps.

DEFINITION 3.2. The process $X = (X_t, \mathcal{F}, P_{r,\mu})$ is a solution to the partial martingale problem for A if for every $\phi \in D(A)$

$$X_t(\phi) - X_r(\phi) - \int_r^t X_s(A\phi) ds$$

is a $P_{r,\mu}$ -martingale for $s \in [r, t]$.

DEFINITION 3.3. The lifting of an additive functional, a , of ξ is a natural right continuous additive functional A of X such that for every $t \geq r$

$$A(r, s] + X_s(\Pi_{s,\cdot} a(s, t])$$

is a $P_{r,\mu}$ -martingale for $s \in [r, t]$.

EXAMPLE 3.1. If $a(r, t] = \int_r^t f(\xi_s)ds$, then the Markov and criticality properties show that $A(r, t] = \int_r^t X_s(f)ds$ is a lifting of a . In general, $A(r, t]$ simply adds up the increments of a over all particles in the population on $(r, t]$.

THEOREM 3.2. (*Leduc [77]*) Assume that

$$h_t^r(x) := \Pi_{r,x} k(r, t) \rightarrow 0 \text{ uniformly in } x \text{ as } t - r \rightarrow 0$$

and that $\ell_x^s(d\mu) = u\delta_x\ell(x, du)$. Then there exists a unique lifting $K^{\Phi(\phi)dK}$ of the additive functional $\Phi(\xi_s, \phi)K(ds)$ and the only solution to the full martingale problem:

$$\exp(-X_t(\phi)) + \int_r^t \exp(-X_s(\phi))X_s(A\phi)ds - \int_r^t \exp(-X_s(\phi))K^{\Phi(\phi)dK}(ds)$$

is a $P_{r,\mu}$ -martingale for every $\phi \in D(A)$,

is the (ξ, Φ, K) -superprocess.

REMARK 3.2. The proof proceeds in two steps. In the first step it is shown that if X is a solution to the partial martingale problem then there exists a unique lifting of the additive functional $\Phi(\xi_s, \phi)K(ds)$. Then in a second step it is established that any solution to the full martingale problem is the (ξ, Φ, K) -superprocess and existence is obtained by verifying that the superprocess satisfies the full martingale problem.

4. Small Scale Behavior

In this section we study small scale qualitative properties of super-Brownian motion $X_t \in M_F(\mathbb{R}^d)$ with law \mathbb{P}_m and its associated historical process $H_t \in M_F(C^d)$ with law $\mathbb{Q}_{\tau,m}$. The required tools include the historical modulus of continuity and cluster representation from Section 2, Hausdorff measures, and capacities.

4.1. Structure of Closed Support of SBM. Dawson and Hochberg [29] proved that if $d \geq 2$, X_t is a singular measure supported by a Borel set of Hausdorff dimension two. To state a more precise result we first introduce some notation. Let

$$\phi_d(r) = \begin{cases} r^2 \log^+ \log^+ \frac{1}{r} & d \geq 3 \\ r^2 (\log^+ \frac{1}{r}) (\log^+ \log^+ \log^+ \frac{1}{r}) & d = 2 \end{cases}$$

and let

$$\phi - m(A) = \liminf_{\delta \downarrow 0} \left\{ \sum \phi(d(B_j)) : d(B_j) \leq \delta, A \subset \bigcup_{j=0}^{\infty} B_j \right\}$$

where $d(B_j)$ is the diameter of B_j .

Here $S(\mu)$ is the closed support of a measure μ on a metric space. We work with super-Brownian motion with branching rate γ .

THEOREM 4.1. (a) If $d \geq 2$, $\exists C_d > 0$ such that for all $m \in M_F(\mathbb{R}^d)$ and all $t > 0$

$$X_t(A) = \gamma C_d \phi_d - m(A \cap S(X_t)) \quad \forall \text{ Borel sets } A, \quad \mathbb{P}_m - \text{a.s.} .$$

(b) If $d \geq 3$, $\exists 0 < C'_d \leq C''_d < \infty$ such that for all $m \in M_F(\mathbb{R}^d)$

$$\begin{aligned} \gamma C'_d \phi_d - m(A \cap S(X_t)) &\leq X_t(A) \leq \gamma C''_d \phi_d - m(A \cap S(X_t)) \\ \forall \text{ Borel sets } A, t > 0, \mathbb{P}_m - \text{a.s.} \end{aligned}$$

(c) If $d = 2$, $X_t \perp dx \forall t > 0$, $\mathbb{P}_m - \text{a.s.} \forall m \in \mathbb{M}_F(\mathbb{R}^d)$.

Items (b) and (c) are proved in Perkins [100], [101] respectively by using a nonstandard model for the approximating branching particle systems. For $d \geq 3$, (a) is then proved in Chapter 5 of Dawson and Perkins [30] through a 0-1 law which shows that the Radon-Nikodym derivative $dX_t/d\phi_d - m(\cdot \cap S(X_t))$ is a.s. constant. The $d = 2$ case of (a) is proved in Le Gall and Perkins [86] using Le Gall's path-valued process (c.f. Le Gall [81]). Whether or not one can take $C'_d = C''_d$ in (b) remains unresolved. The upper bound on X_t implicit in (a) is proved in Dawson [19] using the machinery of infinitely divisible random measures. We present a self-contained proof of a weakened version of Theorem 4.1 below.

4.1.1. Bounds on Hausdorff Measures of the Closed Support Process. Let $\psi_1(r) = r^2(\log^+ \frac{1}{r})^{-1}$ and let $B(x, r)$ be the open ball of radius r , centered at x .

THEOREM 4.2. (*Upper bound*). *If $d \geq 3$, then*

$$\psi_1 - m(S(X_t)) < \infty \quad \forall t > 0 \quad \mathbb{P}_m - \text{a.s..}$$

Proof:

We may work with a historical Brownian motion H and $X_t = H_t(y_t \in \cdot)$. By the Historical cluster representation (Proposition 2.1),

$S(H_{(j+1)2^{-n}}(y(j2^{-n}) \in \cdot))$ is a finite set $\{y_1^{j,n}, \dots, y_{M(j,n)}^{j,n}\}$, where conditional on $\mathcal{H}[0, j2^{-n}]$, $M(j, n)$ is Poisson with mean $2^{n+1} \gamma^{-1} H_{j2^{-n}}(1)$. The modulus of continuity (c.f. Theorem 3.2 - see also Remark 2.4) implies that on $\{\omega : \delta(3, \omega) > 2^{1-n}\}$.

$$(4.1) \quad S(X_t) \subset \bigcup_{i=1}^{M(j,n)} B(y_i^{j,n}, 3h(2^{1-n})) \quad \forall t \in \left[\frac{j+1}{2^n}, \frac{j+2}{2^n} \right] \quad \forall j \in \mathbb{Z}_+.$$

An elementary application of the Borel-Cantelli lemma shows that for sufficiently large n and all $j \in \mathbb{Z}_+$

$$(4.2) \quad M(j, n) \leq 2^{n+2} \gamma^{-1} (H_{j2^{-n}}(1) + 1).$$

This and the definition of Hausdorff measure complete the proof. ■

To prove a companion lower bound on the Hausdorff measure of $S(X_t)$, we use the following well-known lemma of Frostman (c.f. for example, Theorem 1.4 of Perkins [100]).

Assume that $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a strictly increasing and continuous near 0 with $\phi(0) = 0$.

LEMMA 4.1. *If $\nu \in M_F(\mathbb{R}^d)$ satisfies*

$$\limsup_{r \downarrow 0} \nu(B(x, r)) \phi(r)^{-1} \leq K \quad \nu - \text{a.a. } x,$$

then

$$\nu(A) \leq K\phi - m(A) \quad \text{for all Borel sets } A.$$

Let $\{P_t : t \geq 0\}$ denote the Brownian semigroup and let

$$G(f, t) = \int_0^t \sup_x P_s f(x) ds \quad \text{if } f : \mathbb{R}^d \rightarrow [0, \infty] \text{ is measurable.}$$

LEMMA 4.2. *If f is a non-negative, bounded measurable function on \mathbb{R}^d , and $G(f, t)\gamma/2 < 1$, then*

$$\mathbb{P}_m(\exp(X_t(f))) \leq \exp\{m(P_t f)(1 - \frac{\gamma}{2} G(f, t))^{-1}\} < \infty.$$

Proof:

Let $k(s) = (1 - \frac{\gamma}{2} G(f, t-s))^{-1}$. By Theorem 1.1, for $\phi \in D(A)$,

$$(4.3) \quad M_t(\phi) = X_t(\phi) - X_0(\phi) - \int_0^t X_s(A\phi) ds$$

is an (\mathcal{F}_t^X) -martingale starting at zero with quadratic variation

$$< M(\phi) >_t = \int_0^t X_s(\gamma\phi^2) ds$$

for $\phi \in D(A)$. This and the easily verified fact that $D(A)$ is bounded pointwise dense in the set $b\mathcal{E}$ allows one to take bounded pointwise limits and extend $M_t(\phi)$ to all ϕ in $b\mathcal{E}$ so that $M_t(\phi)$ is still a continuous martingale with quadratic variation $\int_0^t X_s(\gamma\phi^2) ds$. As in the usual construction of the stochastic integral, we may now define

$$M_t(\phi) = \int_0^t \int \phi(s, \omega, x) dM(s, x)$$

for $\mathcal{P}^X \times \mathcal{E}$ -measurable ϕ on $\mathbb{R}_+ \times \Omega \times E$ such that $\int_0^t \int \phi(s, \omega, x)^2 X_s(dx) ds < \infty$ for every $t > 0$ \mathbb{P}_m -a.s.. $M_t(\phi)$ is the a continuous local martingale with square function equal to the above integral. See Chapter 2 of Walsh [122] for details.

Using Itô's lemma, (1.1) and an argument similar to that in Step 2 of the proof of Theorem 1.1, it can be verified that

$$\begin{aligned} Y_s &\equiv X_s(P_{t-s}f)k(s) = X_0(P_tf)k(0) + \int_0^s \int P_{t-r}f(x)k(r) dM(r, x) \\ &\quad - \frac{\gamma}{2} \int_0^s X_r(P_{t-r}f)G'(f, t-r)k(r)^2 dr, \quad s \leq t. \end{aligned}$$

Therefore, by Itô's lemma again,

$$\begin{aligned} \exp(Y_s) &= \exp(Y_0) + \int_0^s \int \exp(Y_r) P_{t-r}f(x)k(r) dM(r, x) \\ &\quad + \int_0^s \exp(Y_r) \frac{\gamma}{2} X_r((P_{t-r}f)^2 - (P_{t-r}f)G'(f, t-r))k(r)^2 dr. \end{aligned}$$

Since $G'(f, t-r) = \|P_{t-r}f\|_\infty \geq P_{t-r}f(x)$, it follows that $\exp(Y_t)$ is a non-negative local supermartingale and hence a supermartingale by Fatou's lemma. The resulting inequality, $\mathbb{P}_m(\exp(Y_t)) \leq \mathbb{P}_m(\exp(Y_0))$, completes the proof. ■

Let $\alpha(d) = \begin{cases} 2 & d=2 \\ 1 & d>2 \end{cases}$ and $\psi_d(r) = r^2(\log^+ \frac{1}{r})^{\alpha(d)}$.

In the following result we assume $d \geq 2$.

THEOREM 4.3. (*Lower bound*) (a) *There is a $C(d) > 0$ such that \mathbb{P}_m -a.s. if $\delta > 0$ there is an $r_0(\delta, \omega) > 0$ for which*

$$\sup_{t \geq \delta} \sup_x X_t(B(x, r)) \leq \gamma C(d) \psi_d(r) \quad \forall r \in (0, r_0(\delta, \omega))$$

$$(b) \gamma C(d) \psi_d - m(A \cap S(X_t)) \geq X_t(A) \quad \forall \text{ Borel sets } A \text{ and } t > 0, \quad \mathbb{P}_m - \text{a.s.}$$

REMARK 4.1. A slightly refined version of (a) is Theorem 4.7 of Barlow, Evans and Perkins [5]. The proof we give here is similar but contains some simplifications.

Proof:

(b) is an immediate consequence of (a) and Lemma 4.1. Turning to (a), we may assume as usual that $X_t(\cdot) = H_t(y_t \in \cdot)$, where H has law $\mathbb{Q}_{0,m}$. Let $\mathcal{C}(n)$ be the set of open balls $B(x, (\sqrt{d} + 4)h(2^{-n}))$ centered at a point x in $2^{-n/2}\mathbb{Z}^d \cap [-n, n]^d$. For each B in $\mathcal{C}(n)$ and j in \mathbb{Z}_+ , Remark 2.1 shows that $M(B, j2^{-n})(t) = H_{j2^{-n}+t}(\{y : y(j2^{-n}) \in B\})$ is an $\mathcal{H}[0, t+j2^{-n}]$ -martingale and hence $M(j2^{-n}, t) = \sup\{M(B, j2^{-n})(t) : B \in \mathcal{C}(n)\}$ is a non-negative submartingale. Note also that if $\lambda < 2^{n+1}/\gamma$, then the Remark 2.1 shows that for $n \geq n_0$,

$$\begin{aligned} (4.4) \quad & \mathbb{Q}_{0,m}(\exp(\lambda M(B, j2^{-n})(2^{-n})) | \mathcal{H}[0, j2^{-n}]) \\ &= \mathbb{P}_{X(j2^{-n})|B}(\exp(\lambda X_{j2^{-n}}(1))) \\ &\leq \exp(2X_{j2^{-n}}(B)\lambda(2 - \gamma\lambda 2^{-n})^{-1}) \quad (\text{by Lemma 4.2}). \end{aligned}$$

Assume $\delta(3, \omega) > 2^{-n}$ (δ as in Theorem 3.2). If $x \in [-n, n]^d$, choose $x_0 \in 2^{-n/2}\mathbb{Z}^d \cap [-n, n]^d$ such that $|x_0 - x| \leq \sqrt{d}2^{-n/2} \leq \sqrt{d}h(2^{-n})$. If $t \in [j2^{-n}, (j+1)2^{-n}]$ ($j \in \mathbb{Z}_+$), then for H_t -a.a. y (recall Theorem 3.2) $y(t) \in B(x, h(2^{-n}))$ implies

$$|y(j2^{-n}) - x_0| \leq |y(j2^{-n}) - y(t)| + |y(t) - x| + |x - x_0| \leq (\sqrt{d} + 4)h(2^{-n})$$

and so $H_t(y(t) \in B(x, h(2^{-n}))) \leq H_t(y(j2^{-n}) \in B(x_0, (\sqrt{d} + 4)h(2^{-n})))$ which gives

$$(4.5) \quad \sup_{j2^{-n} \leq t \leq (j+1)2^{-n}} \sup_{x \in [-n, n]^d} X_t(B(x, h(2^{-n}))) \leq \sup_{t \leq 2^{-n}} M(j2^{-n}, t).$$

Let $\lambda_n = C_1 \gamma^{-1} 2^n n^{-\alpha(d)}$ and $\epsilon_n = C_2 \gamma 2^{-n} n^{\alpha(d)+1}$, where C_1 and C_2 will be chosen below. (4.5) and the weak maximal inequality for submartingales imply that for n sufficiently large

$$\begin{aligned} \mathbb{Q}_{0,m}(\sup\{X_t(B(x, h(2^{-n}))) : x \in [-n, n]^d, j2^{-n} \leq t \leq (j+1)2^{-n}\}) &> \epsilon_n, \\ \delta(3) &> 2^{-n} \end{aligned}$$

$$\begin{aligned}
&\leq \mathbb{Q}_{0,m}(\sup_{t \leq 2^{-n}} \exp(\lambda_n M(j2^{-n}, t)) > e^{\lambda_n \epsilon_n}) \\
&\leq e^{-\lambda_n \epsilon_n} \mathbb{Q}_{0,m}(\exp(\lambda_n M(j2^{-n}, 2^{-n}))) \\
&\leq e^{-\lambda_n \epsilon_n} \sum_{B \in \mathcal{C}(n)} \mathbb{Q}_{0,m}(\exp(\lambda_n M(B, j2^{-n})(2^{-n}))) \\
&\leq e^{-\lambda_n \epsilon_n} C_3 n^d 2^{nd/2} \sup_{B \in \mathcal{C}(n)} \mathbb{Q}_{0,m}(\exp(2\lambda_n (2 - \gamma \lambda_n 2^{-n})^{-1} X_{j2^{-n}}(B))) \\
&\quad (\text{by 4.4 since } \lambda_n < 2^{n+1}/\gamma \text{ for } n \geq n_0) \\
(4.6) \quad &\leq e^{-C_1 C_2 n} C_3 n^d 2^{nd/2} \sup_{B \in \mathcal{C}(n)} \exp\{2\lambda_n P_{0,m}(B(j2^{-n}) \in B) \\
&\quad \cdot (1 - \frac{\gamma}{2}(2\lambda_n) G(1_B, j2^{-n}))^{-1}\}.
\end{aligned}$$

In the last line we require $G(1_B, j2^{-n})2\lambda_n < 2/\gamma$ to apply Lemma 4.2. An easy calculation (c.f. Lemma 2.7 of Perkins [100]) shows that

$$G(1_B, j2^{-n}) \leq C_4 2^{-n} n^{\alpha(d)}$$

and so

$$2\lambda_n G(1_B, j2^{-n}) \leq 2C_1 C_4 \gamma^{-1} = \gamma^{-1},$$

where we have chosen $C_1 = (2C_4)^{-1}$. This justifies (4.6) and allows us to bound (4.6) by

$$\begin{aligned}
&\exp(-C_1 C_2 n) C_3 n^d 2^{nd/2} \sup_{B \in \mathcal{C}(n)} \exp(4\lambda_n P_{0,m}(B(j2^{-n}) \in B)) \\
&\leq \exp(-C_1 C_2 n) C_3 n^d 2^{nd/2} \exp(C_5 \gamma^{-1} 2^n n^{-\alpha(d)} m(1) h(2^{-n})^d (j2^{-n})^{-d/2}) \\
&\leq C_6(\delta_0, m(1)) n^d 2^{nd/2} e^{-C_1 C_2 n},
\end{aligned}$$

where the last bound is valid for $j2^{-n} \geq \delta_0 > 0$. Summing over $j \in [2^n \delta_0, n2^n]$, we get

$$\begin{aligned}
&\mathbb{Q}_{0,m}(\sup\{X_t(B(x, h(2^{-n}))) : x \in [-n, n]^d, \delta_0 + 2^{-n} \leq t \leq n\} > \epsilon_n, \delta(3) > 2^{-n}) \\
&\leq C_6(\delta_0, m(1)) n^{d+1} 2^{n(d+2)/2} e^{-C_1 C_2 n}.
\end{aligned}$$

Choose C_2 large enough so that the above bound is summable. By Borel-Cantelli and Theorem 3.2 we see that $\mathbb{Q}_{0,m}$ -a.s for any $\delta_0 > 0$ there is an $N(\omega, \delta_0)$ such that for $n \geq N(\omega, \delta_0)$

$$\sup\{X_t(B(x, h(2^{-n}))) : \delta_0 \leq t \leq n, x \in [-n, n]^d\} \leq C_2 \gamma n^{\alpha(d)+1} 2^{-n}.$$

An easy interpolation argument together with the facts that a.s. X_t has compact support for all $t > 0$ (recall the Remark 2.4) and a finite lifetime now complete the proof. ■

4.2. Occupation Measures and Local Times. We define the weighted occupation measure

$$Y_{s,t}(\psi) := \int_s^t X_r(\psi(r)) dr, \quad 0 \leq s \leq t, \quad \psi \in p\mathcal{B}([0, t]).$$

Then (c.f. for example, Iscoe [67])

$$\mathbb{P}_{s,m}[\exp(-X_t(\phi) - Y_{s,t}(\psi))] = \exp(-m(v(s, t, \cdot))),$$

where

$$v(s, t, a) = S_{s,t}\phi(a) + \int_s^t S_{s,r}\psi(r)dr - \frac{\gamma}{2} \int S_{s,r}v^2(r, t)dr.$$

THEOREM 4.4. (*Sugitani [118]*) Let $d \leq 3$. Fix $\mu \in M_p$ and $\delta \geq 0$. If $\delta = 0$, assume additionally that

$$[r, z] \rightarrow \mu * q(0, r, z) \text{ is continuous on } \mathbb{R}^+ \times \mathbb{R}^d,$$

where $q(s, t, a, b) = \int_s^t dr p(r, a, b)$. Then with respect to \mathbb{Q}_μ , there is a jointly continuous field

$$y_\delta = y_\delta^{\gamma dr} := \{y_{\delta, \delta+t}(z), t \geq 0, z \in \mathbb{R}^d\}$$

on $\mathbb{R}^+ \times \mathbb{R}^d$ such that for the occupation times $Y_{\delta, \delta+t} = Y_{\delta, \delta+t}^{\gamma dr}$ we have

$$\mathbb{P}_\mu(Y_{\delta, \delta+t}^{\gamma dr}(dz) = y_{\delta, \delta+t}(z)dz \text{ for all } t \geq 0) = 1.$$

Note that y_δ is called the *occupation density field* (super-Brownian local time) of $Y_{\delta, \delta+t}^{\gamma dr}$.

THEOREM 4.5. (*Dawson and Fleischmann [24]*) Let $d \leq 3$, $\xi \in (0, \frac{1}{4})$ and let k denote the smallest natural number satisfying

$$k > \frac{2(d+1)}{1-4\xi}.$$

Fix $\mu \in M_p$ and $\delta > 0$. Then, with respect to \mathbb{P}_μ , there exists a modification of the occupation density field $y^\delta = y_\delta^{\gamma dr} = \{y_{\delta, \delta+t}(z); t \geq 0, z \in \mathbb{R}^d\}$ related to $Y_{\delta, \delta+t}^{\gamma dr}$, such that with

$$\begin{aligned} |y_{\delta, \delta+t_1}(z_1) - y_{\delta, \delta+t_2}(z_2)| &\leq C_{\xi, N, k} |[t_1, z_1] - [t_2, z_2]|^\xi \\ (t_i, z_i) &\in [0, N] \times [-N, N]^d \end{aligned}$$

and $\mathbb{P}_\mu C_{\xi, N, k}^{2k} \leq \text{const}(1 \vee \|\mu\|_p^{2k})$. In particular, y_δ is locally jointly ξ -Hölder continuous.

The proof is based on estimates of the moments of the increments and Kolmogorov's criterion.

4.3. The Rescaled Dead Mass Process. The *dead mass process* $D(t)$, $t \geq 0$ is defined by $D(t) := \sum \delta_{z_i(t)}$, where $\{z_i(t)\}$ are the locations at death of the particles that have died up to time t . Let $D^n(t)$ denote the process with $(1+\beta)$ -branching mechanism as in Section 1.7, branching rate $n^\beta \gamma$, and giving each living particle mass $\frac{1}{n}$ and each dead particle mass $\frac{1}{\gamma n^{(1+\beta)}}$.

THEOREM 4.6. (*Gorostiza and Lopez-Mimbela [62]*) If $X^n(0) \Rightarrow X(0)$ as $n \rightarrow \infty$, then for each $t > 0$

$$\left(X^n(t), \int_0^t X^n(s)ds, D^n(t) \right) \Rightarrow \left(X(t), \int_0^t X(s)ds, D(t) \right)$$

and the Laplace functional of the limit is

$$\begin{aligned} E \left[e^{-\{(X_t(\xi) + \int_0^t X_s(\phi)ds + D_t(\psi))\}} \mid X_0 = \mu \right] &= e^{\mu(u_{\xi, \phi, \psi}(t))} \\ \xi, \phi, \psi &\in C_p(\mathbb{R}^d), \mu \in M_p(\mathbb{R}^d), \end{aligned}$$

where $u_{\xi, \phi, \psi}$ is the unique mild solution of

$$\frac{\partial u(t)}{\partial t} = \Delta u(t) - \frac{\gamma}{1+\beta} u(t)^{1+\beta} + \phi + \psi, \quad u(0) = \xi.$$

4.4. The Range of SBM. The range of X up to time t ($\leq \infty$) is $R(t) = \cup_{0 < s \leq t} S(X_s)$. (4.1) and (4.2) show that

$$\begin{aligned} R_n(\infty) &:= \cup_{2^{-n} \leq s < \infty} S(X_s) \subset \cup_{i=1}^M B(y_i, 3h(2^{1-n})) \\ M &\leq 2^{2n+2} \gamma^{-1} (\sup_t H_t(1) + 1)(\zeta + 1), \end{aligned}$$

where ζ is the lifetime of H . This implies $\chi - m(R(\infty)) = \lim \chi - m(R_n(\infty)) < \infty$, where $\chi(r) = r^4 (\log^+ \frac{1}{r})^{-2}$. This proves that the range has Hausdorff dimension at most 4.

More precisely, (c.f. Dawson et. al. [31], and Le Gall [82]) if $d \geq 4$ and

$$\chi_d(r) = \begin{cases} r^4 \log^{\frac{1}{d}} (\log \log \log \frac{1}{r}) & d = 4 \\ r^4 (\log \log \frac{1}{r}) & d > 4, \end{cases}$$

then there are constants $\bar{C}_d > 0$ such that

$$(4.7) \quad Y_t(A) := \int_0^t X_s(A) ds = \gamma \bar{C}_d \chi_d - m(A \cap R(t)) \quad \forall t \geq 0,$$

for a Borel subset A of \mathbb{R}^d $\mathbb{P}_m - a.s.$.

The result for the critical 4-dimensional case is proved by Le Gall [82] using his path-valued (Brownian snake) process.

It is fairly easy to use scaling properties to transfer (4.7) to Aldous' ISE described in Section 1.10. We provide the details because this result is needed in the recent work of Derbez and Slade [38].

THEOREM 4.7. *If $P(X \in \cdot)$ is the law of ISE, X , then for $d \geq 4$*

$$X(A) = \bar{c}_d \chi_d(A \cap S(X)) \quad \forall A \text{ a Borel subset of } \mathbb{R}^d \text{ } P - a.s..$$

In particular $\dim(S(X)) = 4$, $P - a.s.$.

Proof:

Recall that the canonical measure, \bar{R} of $Y = \int_0^\infty X_s ds$ under \mathbb{P}_{δ_0} satisfies

$$(4.8) \quad \mathbb{P}_{c\delta_0}(e^{-Y(\phi)}) = \exp\{-c \int (1 - e^{-\nu(\phi)}) \bar{R}(d\nu)\}, \quad \phi \geq 0, \text{ measurable.}$$

If $\nu^{(K)}(B) = K^4 \nu(K^{-1}B)$ ($\nu \in M_F(\mathbb{R}^d)$), the scaling property of super-Brownian motion implies

$$\mathbb{P}_{\delta_0}(Y \in \cdot) = \mathbb{P}_{c^2\delta_0}(Y^{(c^{-1})} \in \cdot)$$

and so by (4.8)

$$(4.9) \quad c^2 \bar{R}(\nu^{(c^{-1})} \in \cdot) = \bar{R}(\cdot).$$

Recall that if $I_\varepsilon(x) = [x - \varepsilon, x + \varepsilon]$ then (c.f. (1.17) in Section 1.10)

$$(4.10) \quad P(X \in \cdot) = w - \lim_{\varepsilon \downarrow 0} \bar{R}(\cdot | \nu(1) \in I_\varepsilon(1)).$$

If ϕ is bounded and continuous on $M_F(\mathbb{R}^d) - \{0\}$ and $x > 0$, we may use (4.9) with $c = x^{-1/4}$ to see that

$$(4.11) \quad \begin{aligned} \bar{R}(\phi(\nu)|\nu(1) \in I_\varepsilon(x)) &= \bar{R}(\phi(\nu^{(x^{1/4})})|\nu^{(x^{1/4})}(1) \in I_\varepsilon(x)) \\ &= \bar{R}(\phi(\nu^{(x^{1/4})})|\nu(1) \in I_{\varepsilon x^{-1}}(1)) \\ &\rightarrow P(\phi(X^{(x^{1/4})}) \text{ as } \varepsilon \downarrow 0. \end{aligned}$$

We have used (4.10) in the last line. This shows that the left side of (4.11) converges weakly to a probability P_x on $M_p(\mathbb{R}^d) - \{0\}$ which satisfies

$$(4.12) \quad P_x(\phi(X)) = P(\phi(X^{(x^{1/4})})).$$

It follows from (4.7) that if

$$C = \{\nu \in M_F(\mathbb{R}^d) - \{0\} : \nu(A) = \bar{C}_d \chi_d - m(A \cap S(\nu)) \forall A \in \mathcal{B}(\mathbb{R}^d)\},$$

then

$$(4.13) \quad \bar{R}(C^c) = 0.$$

More precisely this follows from Chapter 11 of Le Gall [82] and the fact that $\bar{R}(\cdot)$ is the law of $\int_0^\infty X_s ds$ under Le Gall's excursion measure for the path-valued process. Standard differentiation theory shows that

$$P_x(X \in \cdot) = \bar{R}(\nu \in \cdot | \nu(1) = x) \quad \bar{R}(\nu(1) \in dx) - a.s.,$$

that is, the left side is a regular conditional probability for the right side. This shows that (4.13) implies

$$P_x(X \in C^c) = 0, \quad \bar{R}(\nu(1) \in dx) - a.s..$$

Fix an $x > 0$ so that the above probability is zero.

We claim that

$$(4.14) \quad \nu \in C \Leftrightarrow \nu^{(K)} \in C \quad \forall K > 0.$$

This is clear from the facts that $S(\nu^{(K)}) = KS(\nu)$ and

$$\chi_d - m(A \cap S(\nu^{(K)})) = K^4 \phi - m(AK^{-1} \cap S(\nu)).$$

Therefore

$$\begin{aligned} P(X \in C^c) &= P(X^{(x^{1/4})} \in C) \quad (\text{by (4.14)}) \\ &= P_x(X \in C^c) \quad (\text{by (4.12)}) \\ &= 0 \quad (\text{choice of } x). \blacksquare \end{aligned}$$

4.5. Hitting Probabilities. Let D be an open set. Then

$$\begin{aligned} \mathbb{Q}_{\delta_x}(Y_T(D^c) = 0) &= \lim_{\theta \rightarrow \infty} \mathbb{Q}_{\delta_x}(e^{-\theta Y_T(D^c)}) \\ &= \lim_{\theta \rightarrow \infty} e^{-v(\theta, 0, T)} \end{aligned}$$

$$v(a, 0, T) = \int_0^T a S_{0,r} 1_{D^c} dr - \frac{\gamma}{2} \int_0^T S_{0,r} v^2(r, T) dr$$

and $v(a) = \lim_{T \rightarrow \infty} v(a, 0, T)$ solves

$$\Delta v = \gamma v^2 \text{ in } D.$$

In fact, Dynkin proved that $v(x) := -\log \mathbb{Q}_{\delta_x}(R(\infty) \subset D)$ is the maximal solution of the equation, $\Delta u(x) = \gamma u(x)^2$, $x \in D$.

If $d \leq 3$, and $D = \{0\}^c$, then the maximal solution is of the form $\frac{c_d}{|x|^2}$ (cf. Dawson, Iscoe, and Perkins [31]) and solving for c_d , we get

$$(4.15) \quad \mathbb{P}_m(x \in R(\infty)) = 1 - \exp\left(-\frac{2(4-d)}{\gamma} \int |y-x|^{-2} dm(y)\right) > 0 \quad (m \neq 0).$$

4.6. Multiple Points. If $R([s, t]) = \cup_{s \leq r \leq t} S(X_r)$, define the set of k-multiple points of X by

$$R_k = \cup(\cap_{j=1}^k R(I_j) : I_1, \dots, I_k \text{ disjoint compact subintervals of } (0, \infty)).$$

If I_1, I_2 are disjoint compact intervals in $(0, \infty)$ and $d \geq 4$, then (4.7) implies $R(I_1)$ and $R(I_2)$ are sets of dimension 4 with zero 4-dimensional measure (if non-empty) and hence we believe should intersect if $d \leq 7$ but not if $d \geq 8$. This intuition is correct and in fact,

THEOREM 4.8. $R_k = \phi$ a.s. if $k \geq d/(d-4)^+$ and $R_k \neq \phi$ a.s. if $k < d/(d-4)^+$ (and $X_0 \equiv m \neq 0$), (c.f. Perkins [102] and Dawson et. al. [31]).

REMARK 4.2. Serlet [108] has shown that for $d \geq 5$, $\phi(r) = r^{(d+k(d-4))} (\log \log(\frac{1}{r}))^k$ is an exact Hausdorff measure function for R_k in the sense that R_k has positive and σ -finite $\phi - m$. He used the path process of Le Gall [80] which has proved to be a powerful tool in these exact Hausdorff measure calculation.

4.7. R and G Polar Sets. The graph of X is $G(X) = \{(t, x) : t > 0, x \in S(X_t)\}$ and the closed graph of X , $\overline{G}(X)$ is the closure of $G(X)$ in $(0, \infty) \times \mathbb{R}^d$. A set $A \subset (0, \infty) \times \mathbb{R}^d$ is G-polar for X if and only if $\mathbb{P}_m(A \cap \overline{G}(X) = \emptyset) = 1$ for all $m \in M_F(\mathbb{R}^d)$ (i.e. if the set of points in space-time traced out by X misses A \mathbb{P}_m -a.s.). A set $A \subset \mathbb{R}^d$ is R-polar for X if and only if $\mathbb{P}_m(A \cap R(\infty) = \emptyset) = 1$ for all $m \in M_F(\mathbb{R}^d)$.

$$\text{Let } g_d(r) = \begin{cases} r^{2-d} & \text{if } d > 2 \\ \log^+ \frac{1}{r} & \text{if } d = 2 \end{cases}.$$

A subset F of \mathbb{R}^d has g_d -capacity zero if and only if for any $\nu \in M_F(\mathbb{R}^d)$,

$$\int_A \int_A g_d(|x - x'|) \nu(dx) \nu(dx') < \infty \quad \text{implies } \nu(A) = 0.$$

A subset A of $(0, \infty) \times \mathbb{R}^d$ has \tilde{g}_d -capacity zero if and only if for any $\nu \in M_F((0, \infty) \times \mathbb{R}^d)$

$$\int_A \int_A g_d(|x - x'| \vee (|t - t'|^{1/2})) \nu(dx, dt) \nu(dx', dt') < \infty \quad \text{implies } \nu(A) = 0.$$

THEOREM 4.9. (Dynkin [42] and Perkins [102]) (a) An analytic set A is G-polar for X if and only if A has \tilde{g}_d -capacity zero ($d > 1$).

(b) An analytic set $F \subset \mathbb{R}^d$ with $d > 3$ is R-polar if and only if F has g_{d-2} -capacity

zero if and only if F is a removable singularity, that is, zero is the only solution of the equation

$$\Delta u = \gamma u^2 \text{ in } D = F^c.$$

The capacity criterion for G-polarity in (a) is different from that in Dynkin [42], but is readily shown to be equivalent to it in our finite variance setting.

The proof of (b) (c.f. Theorem 12.4' of Dynkin [42]) uses the fact that the function $-\log \mathbb{Q}_{\delta_x}(\text{Range} \subset D)$ is the maximal solution of this equation, for D open, and the results of Baras and Pierre [4].

REMARK 4.3. Recently for $d \geq 3$ Le Gall [83] has proved that subsets of the boundary ∂D are polar for SBM stopped on exit from D if and only if it has $\frac{1}{|x|^{d-3}}$ (or $\log |x|$ if $d = 3$)-capacity zero. He uses his Brownian snake process $\{W_s, s \geq 0\}$, where $W_s = (w_s, \zeta_s)$, $w_s \in C([0, \zeta_s], \mathbb{R}^d)$, $\zeta \geq 0$ and ζ is a one-dimensional reflecting Brownian motion. Conditioned on ζ , $w_s(0) = x$, $w_s(t) = w_s(t) \forall t \in [0, m(s, s')]$, $m(s, s') := \inf\{\zeta(u) : u \in [s, s']\}$, $w_{s'}(m(s, s') + t) - w_s(m(s, s'))$ and $w_s(m(s, s') + t) - w_s(m(s, s'))$ are independent Brownian motions stopped at $\zeta_{s'} - m(s, s')$ and $\zeta_s - m(s, s')$, respectively. Let $\alpha := \inf\{s : \ell_s^0(\zeta) \geq 1\}$, where ℓ^a denotes local time of ζ at a . Then $H_t(\phi) = \int_0^\alpha \phi(W_s) d_s \ell_s^t$ has law \mathbb{Q}_{δ_x} (c.f Le Gall [81], [82]). This establishes a very useful representation of historical Brownian motion. He then studies the set of points hit by the paths W_s stopped at their exit times from D . Recently this approach has been extended to infinite variance branching by Le Gall and Le Jan [85].

REMARK 4.4. Recently Peres [98] has introduced a method of characterizing polar sets for the intersections of random sets using percolation on trees. This leads to a capacity characterization for the sets which are polar for the k-multiple points, R_k , of super-Brownian motion (c.f. Peres [99]).

4.8. Collisions of Closed Supports.

THEOREM 4.10. (a) (Barlow, Evans and Perkins [5]) If $X_t, Y_t \in M_F(\mathbb{R}^d)$ are independent super-Brownian motions, then $\mathbb{P}_{m_1} \times \mathbb{P}_{m_2}(\bar{G}(X) \cap \bar{G}(Y) \neq \phi) > 0$ if and only if $d < 6, (m_i \neq 0)$.

(b)(Barlow and Perkins [7]) If B is a Brownian motion independent of the super-Brownian motion X , then $\{(t, B_t) : t > 0\} \cap \bar{G}(X) \neq \phi$ with positive probability if and only if $d < 4$.

Also (c.f. Barlow et. al. [5]) one may directly construct a non-trivial measure $L(X, Y)(dt, dx)$ supported on $\bar{G}(X) \cap \bar{G}(Y)$ if $d \leq 5$ as the limit in probability of

$$L^\epsilon(X, Y)([0, t] \times B) = \int_0^t \int \int p_\epsilon(x - y) 1_B((x + y)/2) X_s(dx) Y_s(dy) ds$$

as $\epsilon \downarrow 0$ (p_ϵ is the Brownian transition density). $L(X, Y)$ is called the *collision local time* of X and Y as it records the space-time locations of the collisions between the two populations.

5. Large Scale Behavior

In order to study the large space-time scale behavior of super-Brownian motion we work in the infinite measure setting of Section 1.8 with p a fixed constant greater than d .

5.1. The Persistence-Extinction Dichotomy. Consider the super-Brownian motion in \mathbb{R}^d with $\beta = 1$ and initial measure $X_0 = m\lambda$, $m > 0$, where λ is Lebesgue measure (Wf. Section 1.8). If $\sup |\phi(x)| \cdot (1 + |x|^2)^{\frac{p}{2}} < \infty$, then the solution, $v_t(x) = V[\phi](t, x)$, to (1.7) with $A = \frac{\Delta}{2}$ is integrable and integrating both sides with respect to Lebesgue measure gives

$$\int v_t(x)dx = \frac{\gamma}{2} \int_0^t \int v_s^2(x)dxds.$$

Therefore the large time limit of the Laplace functional

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{P}_{m\lambda}(\exp(-X_t(\phi))) &= \lim_{t \rightarrow \infty} \exp\left(-m \int v_t(x)dx\right) \\ &= \lim_{t \rightarrow \infty} \exp\left(-\frac{m\gamma}{2} \int_0^t \int v_s^2(x)dxds\right) \end{aligned}$$

exists for every $\phi \in \mathcal{B}_+$ since the right side is monotone in t . Therefore X_t converges in distribution as $t \rightarrow \infty$ to a random measure on \mathbb{R}^d with probability law which we denote by \mathbb{P}_m^{eq} . Replacing ϕ by $\theta\phi$, $\theta > 0$, and evaluating the first and second derivatives with respect to θ at $\theta = 0$, we can verify that the first and second moments are given by

$$\mathbb{P}_{m\lambda}(X_t(\phi)) = m \int \phi(x)dx$$

and

$$\begin{aligned} \mathbb{P}_{m\lambda}(X_t(\phi)^2) &= m^2 \left(\int \phi(x)dx \right)^2 + \gamma m \int_0^t \int \left(\int p_s(y-z)\phi(z)dz \right)^2 dyds \\ &= m^2 \left(\int \phi(x)dx \right)^2 + \gamma m \int_0^t \left(\int p_{2s}(z_1-z_2)\phi(z_1)\phi(z_2)dz_1dz_2 \right) ds. \end{aligned}$$

Recalling that for the Brownian motion transition kernel $\int_0^\infty p_s(z)ds$ diverges if $d = 1, 2$ and is given by $\frac{2c_d}{|z|^{d-2}}$ if $d \geq 3$, so we have

$$\begin{aligned} \mathbb{P}_{m\lambda}(X_t(\phi)^2) &\uparrow \infty \text{ if } d = 1, 2 \\ &\uparrow m^2 \left(\int \phi(x)dx \right)^2 + \gamma mc_d \int \int \frac{\phi(z_1)\phi(z_2)}{|z_1 - z_2|^{d-2}} dz_1 dz_2 \text{ if } d \geq 3. \end{aligned}$$

If $d \geq 3$, the above imply that $\{X_t(\phi)\}_{t \geq 0}$ are uniformly integrable and $\mathbb{P}_{m\lambda}(X_\infty(\phi)) = m\lambda(\phi)$. That is, the limiting equilibrium random measure \mathbb{P}_m^{eq} has the same intensity m , as the initial intensity - this behavior is called *persistence*. Recently, Bramson, Cox and Greven [9] have proved that $\{\mathbb{P}_m^{eq} : m \in [0, \infty)\}$ is in fact the set of all extremal invariant measures.

The divergence of the second moment in the low dimensional case suggests that the behavior is qualitatively different in these dimensions. It was proved in Dawson [17] that in this case the spatially homogeneous super-Brownian motion

with $X_0 = m\lambda$ (or any finite mean spatially invariant initial measure) suffers local extinction (i.e. $X_t(A) \rightarrow 0$ in probability as $t \rightarrow \infty$ for any bounded set A). Iscoe [68] has shown that $X_t(A) \xrightarrow{a.s.} 0$ for any bounded set if $d = 1$ and that this result is false if $d = 2$ (c.f. Section 5.3 below). Recently Bramson, Cox and Greven [9] have established that δ_0 is the only measure which is invariant for the process X_t and that for any locally finite initial measure the system undergoes local extinction or explodes thus ruling out the possibility of an invariant measure with infinite mean.

For the case $0 < \beta < 1$, second moments do not exist so that other methods must be used. In particular using the log-Laplace equation and the cluster representation of H_t (c.f. Section 5.4) it can be shown that persistence occurs if and only if $d > \frac{2}{\beta}$.

5.2. Clumping in Low Dimensions. In order to describe the low dimensional behavior of X_t with $X_0 = \lambda$ (Lebesgue) in more detail we introduce the space-time-mass rescaling

$$\begin{aligned} X_t^{K,\xi}(A) &:= K^{-\xi} X_{Kt}(K^{\frac{\xi}{d}} A) \\ X_0^{K,\xi}(A) &= |A|. \end{aligned}$$

Then

$$\begin{aligned} \mathbb{P}_\lambda(\exp(-X_t^{K,\xi}(\phi))) &= \exp(-\lambda(V_{Kt}\phi_K)) \quad \text{with} \\ \phi_K(x) &:= K^{-\xi}\phi(K^{-\frac{\xi}{d}}x). \end{aligned}$$

Note that

$$\tilde{v}(t, x) := K^\xi V_{Kt}\phi_K(K^{\frac{\xi}{d}}x)$$

satisfies

$$\begin{aligned} \frac{\partial \tilde{v}(t, x)}{\partial t} &= K^{1-\frac{2\xi}{d}} \Delta \tilde{v}(t, x) - \frac{\gamma}{1+\beta} K^{1-\xi\beta} \tilde{v}(t, x)^{1+\beta} \\ \tilde{v}(0, x) &= \phi(x) \end{aligned}$$

and therefore $X^{K,\xi}$ is equivalent to a super-Brownian motion with “diffusion coefficient” $K^{1-\frac{2\xi}{d}}$ and “branching coefficient” $\frac{\gamma}{1+\beta} K^{1-\xi\beta}$. The branching term dominates in the $K \rightarrow \infty$ limit and the diffusion term dominates in the $K \rightarrow 0$ limit if $d < \frac{2}{\beta}$ and the opposite occurs if $d > \frac{2}{\beta}$.

THEOREM 5.1. (Dawson and Fleischmann [21]) (a) Let $d < \frac{2}{\beta}$. Then

$X^{K,\xi} \xrightarrow{K \rightarrow \infty} 0$ if $\xi < \frac{1}{\beta}$ and $X^{K,\xi} \xrightarrow{K \rightarrow \infty} \lambda$ if $\xi > \frac{1}{\beta}$.

(b) If $d = 1$ and $\xi = \frac{1}{\beta}$, then X_K converges in distribution as $K \rightarrow \infty$ to the pure atomic process $\{X_t^0\}_{t \geq 0}$ in which X_t^0 is Poisson with intensity $(\frac{\gamma}{1+\beta}\beta t)^{-\frac{1}{\beta}}X(0)$ and the mass of each atom evolves according to a $(1 + \beta)$ -continuous state branching.

(c) If $d = 2$, $\beta = 1$, then $X^{K,1} \xrightarrow{D} X$, that is X is self-similar.

REMARK 5.1. (b) suggests that for $d = 1$ at time K there are clumps of size $K^{\frac{1}{\beta}}$ with interclump distance $K^{\frac{1}{\beta}}$.

In the case $d = 2$, the phenomenon of *diffusive clustering* arises. This is made precise in the following result of Klenke [71].

THEOREM 5.2. (*Theorem 2 of Klenke [71]*) Let $d = 2$, and $I = (-\infty, 1]$. For $\alpha \in I$, let

$$X_t^\alpha(B) := t^{-\alpha} X_t(t^{\alpha/2}B).$$

Then in the sense of finite dimensional distributions

$$\mathcal{L}^{\frac{(\log t)\lambda}{8\pi}}[\{X_t^\alpha(B)\}_{\alpha \in I}] \xrightarrow{t \rightarrow \infty} \mathcal{L}^1[\{Z_{1-\alpha}\}_{\alpha \in I} \cdot \lambda(B)],$$

where Z is a FB process with $Z_0 = 1$.

In the case $\beta = 1$, $d = 2$ (c) provides a link between the small scale and large scale behaviors. In particular it implies that

$$X_{Kt}(B(0, 1)) \xrightarrow{\mathcal{D}} \frac{X_t(B(0, K^{-1/2}))}{K^{-1}}.$$

For $t > 0$, the left side goes to zero in probability as $K \rightarrow \infty$ because of the local extinction result. Then this shows that the local density at time t is 0 which implies that it does not have a non-trivial absolutely continuous component.

5.3. Ergodic Behavior. The extinction-persistence result implies that if ϕ has compact support, then $X_t(\phi)$ converges to zero in probability if $d \leq \frac{2}{\beta}$ and converges in distribution to a non-degenerate limit if $d > \frac{2}{\beta}$. This can be extended to an ergodic theorem in the latter case.

THEOREM 5.3. (*Iscoe [68]*), and *Fleischmann and Gärtner [60]*).

- (a) For $d > \frac{2}{\beta}$ with probability one, $\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t X_s ds = \lambda$ (in the vague topology).
- (b) For $d = \frac{2}{\beta}$, as $t \rightarrow \infty$ $\frac{1}{t} \int_0^t X_s ds$ converges a.s. in the vague topology to $\eta \lambda$, where η is a non-degenerate infinitely divisible random variable with mean one.

REMARK 5.2. (b) implies that in critical dimension $d = \frac{2}{\beta}$, $X_t(\phi)$ goes to zero in probability, but there are arbitrarily large values of t for which $X_t(\phi)$ is large. Therefore the large clumps revisit bounded sets at arbitrarily large times, that is, they have a recurrence property.

In the case of $\beta = 1$, a more refined study of the ergodic behavior has been made using large deviation theory. To describe this let $\phi \geq 0$ be Hölder continuous with compact support and $\int \phi(x) dx = 1$. Then Iscoe and Lee [69] and Lee [78] have shown that in dimension d ,

$$-\lim_{t \rightarrow \infty} a(t, d) \log \mathbb{P} \left\{ \frac{1}{t} \int \int_0^t \phi(x) X_s(dx) > c \right\} > 0$$

when c is bigger than but close to 1, where $a(t, 3) = t^{\frac{1}{2}}$, $a(t, 4) = \frac{t}{\log t}$ and $a(t, d) = t$ for $d \geq 5$. Deuschel and Wang [37] discussed the corresponding large deviation principle in $d = 3, 4$. This initially surprising large deviation behavior is reminiscent of that of independent random walks or Brownian motions in dimensions $d = 1, 2$ and suggests some type of null recurrence phenomena which will be described below.

5.4. The Equilibrium Clan Decomposition.

5.4.1. An Extension of Historical Brownian Motion. We will consider the historical process with time parameter in $(-\infty, \infty)$. Let $C_{s,t}^d = C([s,t], \mathbb{R}^d)$, $C^d = C_{-\infty,\infty}^d$ and $M_p(C_{s,t}^d)$ denote the space of measures on $C_{s,t}^d$ with marginal distributions in $M_p(\mathbb{R}^d)$. Let μ_λ^t denote the law (on C^d) of $\xi_t := \{\xi(s \wedge t)\}_{s \in (-\infty, \infty)}$ when $\{\xi(t) : t \in \mathbb{R}\}$ is a Brownian motion with 0-marginal given by Lebesgue measure. The transition Laplace functional is given by: for $\phi \in p\mathcal{B}(C^d)$

$$\mathbb{Q}_{s,m}(\exp(-H_t(\phi))) = \exp\left(-\int V_{s,t}\phi(y)m(dy)\right), \quad m \in M_p(C^d)^*,$$

where

$$V_{s,t}\phi(y) = S_{s,t}\phi(y) - \frac{\gamma}{2} \int_s^t S_{s,r}((V_{r,t}\phi)^2)dr.$$

If $H_s = \mu_\lambda^s$, then H_t is an infinitely divisible random measure with canonical measure $\int R_{s,t}(y, A)\mu_\lambda^s(dy)$, where R is characterized by $V_{s,t}\phi(y) = \int(1 - e^{-\nu(\phi)})R_{s,t}(y, d\nu)$.

Let $s < u < t$. By the Historical cluster representation, $H_t(\{y : y^u \in \cdot\})$ (under $\mathbb{Q}_{s,m}$) can be represented as a Poisson random field of clan measures in $M_p(C_{s,t}^d)$ with Poisson intensity $\frac{2}{\gamma(t-u)}H_u$. The typical clan measure $\Xi_t(y')$ can be interpreted as the descendent population from an individual alive at time u with history y' and $r_{s,u}\Xi_t = \Xi_t(C^d)\delta_{y'}$, where for $\mu \in M_F(C_{s,t}^d)$, $s \leq u \leq t$, $r_{s,u}\mu(A) := \mu(\{y : y(\cdot \wedge u) \in A\})$.

5.4.2. The Historical Process Conditioned to Live Forever. For $t > s$, let H_t^I denote a realization of the historical process H_t conditioned to stay alive forever starting from a finite initial measure $H_s = \eta_s \in M_p(C_{-\infty,s}^d)$. The law $\mathbb{Q}^I(H_t^I \in \cdot)$ is defined rigorously as the weak limit as $T \rightarrow \infty$ of $\mathbb{Q}_{s,\eta_s}(H_t \in \cdot | H_T \neq 0)$. Using Remark 2.1, we have

$$\begin{aligned} \mathbb{Q}_{s,\eta_s}^I(H_t^I \in A) &= \lim_{T \rightarrow \infty} \mathbb{Q}_{s,\eta_s}(H_t \in A | H_T \neq 0) \\ &= \lim_{T \rightarrow \infty} \frac{\mathbb{Q}_{s,\eta_s}(1_A(H_t)\mathbb{Q}_{t,H_t}(H_T \neq 0))}{\mathbb{Q}_{s,\eta_s}(H_T \neq 0)} \\ &= \lim_{T \rightarrow \infty} \frac{\mathbb{Q}_{s,\eta_s}(1_A(H_t)(1 - e^{-\frac{2H_t(1)}{\gamma(T-t)}}))}{1 - e^{-\frac{2\eta_s(1)}{\gamma(T-s)}}} \\ &= \eta_s(1)^{-1} \mathbb{Q}_{s,\eta_s}(\mathbf{1}_A(H_t)H_t(\mathbf{1})). \end{aligned}$$

The corresponding *normalized Campbell measure* of H_t is

$$\tilde{\mathbb{Q}}_{s,\eta_s}(H_t^I \in A, \bar{\xi}_t \in B) := (\eta_s(\mathbf{1}))^{-1} \mathbb{Q}_{s,\eta_s}[\mathbf{1}_A(H_t)H_t(B)].$$

Let

$$U_{s,t}(\phi, \psi)(y) := \int \nu(\phi)e^{-\nu(\psi)}R_{s,t}(y, d\nu).$$

LEMMA 5.1. *Let $\phi, \psi \in \mathcal{B}_+(C)$. Then*

$$\begin{aligned} U_{s,t}(\phi, \psi)(y) &= \frac{\partial}{\partial \theta} V_{s,t}(\theta\phi + \psi)|_{\theta=0} \\ &= P_{s,y} \left(\phi(\bar{\xi}_t) \exp \left(- \int_s^t \gamma V_{u,t} \psi(\bar{\xi}_u) du \right) \right). \end{aligned}$$

Proof:

Since (recall the connections between $R_{s,t}$ and $V_{s,t}$)

$$\int e^{-\nu(\psi)} \nu(\phi) R_{s,t}(y, d\nu) = \frac{\partial}{\partial \theta} V_{s,t}(\theta \phi + \psi)|_{\theta=0},$$

the first equality is clear. Recall

$$V_{s,t}\phi(y) = S_{s,t}\phi(y) - \frac{\gamma}{2} \int_s^t S_{s,r}((V_{r,t}\phi)^2) dr,$$

and so, replacing ϕ with $\theta\phi + \psi$ and differentiating with respect to θ we get

$$U_{s,t}(\phi, \psi) = S_{s,t}\phi - \gamma \int_s^t S_{s,r}((V_{r,t}\psi) U_{r,t}) dr.$$

Then by a Feynman-Kac argument for the Brownian path process we get

$$U_{s,t}(\phi, \psi)(y) = P_{s,y} \left(\phi(\bar{\xi}_t) \exp \left(- \int_s^t \gamma V_{u,t} \psi(\bar{\xi}_u) du \right) \right). \quad \blacksquare$$

REMARK 5.3. Let $R_{s,t}^{\bar{\xi}}(y, d\nu)$ denote the Palm measure associated with $R_{s,t}(y, d\nu)$ (i.e. it is the regular conditional probability on $M_p(C_{s,t}^d)$ for $A \times B \rightarrow \int \int 1_A(\xi) d\nu 1_B(\nu) R_{s,t}(y, d\nu)$ given $\bar{\xi}$). Lemma 5.1 implies that

$$(5.1) \quad \int \exp(-\nu(\psi)) R_{s,t}^{\bar{\xi}}(y, d\nu) = \exp(- \int_s^t \gamma V_{u,t} \psi(\bar{\xi}_u) du).$$

Intuitively, $R_{s,t}^{\bar{\xi}}(y, d\nu)$ is the law of a random measure obtained if ξ throws off historical excursions at a constant rate γ on $[s, t]$. We let $\Xi_{s,t}^I(\xi)$ denote such an immortal clan (i.e. a random measure with law $R_{s,t}^{\bar{\xi}}(y, \cdot)$).

PROPOSITION 5.1. Under \tilde{Q}_{s,η_s} ,

- (a) $\{\bar{\xi}_t\}_{t \geq s}$ is a path-valued Brownian motion with initial law $(\eta_s(1)^{-1} \eta_s(dy))$, and
- (b) the regular conditional law, $\tilde{Q}_{s,t}^{\bar{\xi}}$ of H_t^I , given $\bar{\xi}$ has Laplace functional

$$\begin{aligned} & \tilde{Q}_{s,t}^{\bar{\xi}}[\exp - \langle H_t^I, \psi \rangle] \\ &= \int \left(\int \exp(-\nu(\psi)) R_{s,t}^{\bar{\xi}}(y, d\nu) \right) \eta_s(dy) \eta_s(1)^{-1} \tilde{Q}_{s,\eta_s}(\exp(- \langle H_t, \psi \rangle)). \end{aligned}$$

Proof:

$$\begin{aligned}
& \tilde{\mathbb{Q}}_{s,\eta_s} \left(\phi(\bar{\xi}) \exp^{-\langle H_t^I, \psi \rangle} \right) = (\eta_s(1))^{-1} \mathbb{Q}_{s,\eta_s} [\exp^{-\langle H_t, \psi \rangle} H_t(\phi)] \\
&= -(\eta_s(1))^{-1} \frac{\partial}{\partial \theta} \mathbb{Q}_{s,\eta_s} (\exp^{-\langle H_t, \theta \phi + \psi \rangle})|_{\theta=0} \\
&= -(\eta_s(1))^{-1} \frac{\partial}{\partial \theta} \left(\exp^{-\eta_s(V_{s,t}(\theta \phi + \psi))} \right)|_{\theta=0} \\
&= \exp(-\eta_s(V_{s,t}(\psi))) P_{s,\eta_s/\eta_s(1)} \left(\phi(\bar{\xi}_t) \exp\left(-\int_s^t \gamma V_{u,t} \psi(\bar{\xi}_u) du\right) \right) \\
&\quad (\text{use Lemma 5.1}) \\
&= \int P_{(s,y)} \left(\phi(\bar{\xi}_t) \int \exp(-\nu(\psi)) R_{s,t}^{\bar{\xi}}(y, d\nu) \right) \eta_s(dy) / \eta_s(1) \\
&\quad \cdot \mathbb{Q}_{s,\eta_s} (\exp(-\langle H_t, \psi \rangle))
\end{aligned}$$

which proves the result. ■

REMARK 5.4. From the perspective of a typical immortal particle chosen according to H_t^I , H^I is the sum of $\Xi_{s,t}^I(\xi)$ and an independent copy of H_t . The former is the contribution to H_t^I of cousins of ξ .

5.4.3. Convergence to Equilibrium for the Historical Process. Recall that if $H_s = \mu_\lambda^s$, then H_t is an infinitely divisible random measure with canonical measure $R_{s,t} = \int R_{s,t}(y, A) \mu_\lambda^s(dy)$, where R is characterized by $V_{s,t}\phi(y) = \int (1 - e^{-\nu(\phi)}) R_{s,t}(y, d\nu)$.

Integrate both sides of (5.1) with respect to μ_λ^s to see that $R_{s,t}^{\bar{\xi}}$ is still the Palm measure of the canonical measure $\int R_{s,t}(y, d\nu) \mu_\lambda^s(dy)$ of H_t .

As $s \rightarrow -\infty$, the total mass of $\Xi_{s,t}^I(\xi)$ goes to infinity (set $\psi = 1$ in Proposition 5.1(b) and let $s \rightarrow -\infty$). However

PROPOSITION 5.2. (a) As $s \rightarrow -\infty$, $\Xi_{s,t}^I(\bar{\xi})$ converges to a locally finite measure if and only if $d \geq 3$.

(b) As $s \rightarrow -\infty$, $R_{s,t}$ converges vaguely to a measure $R_{-\infty,t}$ such that $\int \mu R_{-\infty,t}(d\mu)$ is locally finite.

(c) $R_{-\infty,t}$ is “an entrance law” for H in the sense that if H_s has the infinite law $R_{-\infty,s}$ and $t > s$, then H_t has law $R_{-\infty,t}$.

Proof:

(a) The Laplace functional in (5.1) converges as $s \rightarrow -\infty$ by monotonicity. In the case $d \geq 3$, we can obtain the mean measure of $\Xi_{s,0}^I(\bar{\xi})$ for a trajectory ξ as follows. Differentiate the expression for the Laplace functional in (5.1) to see that for ϕ continuous with compact support

$$\lim_{s \rightarrow -\infty} E \left[\int_{\mathbb{R}^d} \phi(y_t) \Xi_{s,t}^I(\bar{\xi}, dy) \right] = \gamma \lim_{s \rightarrow -\infty} \int_{s-t}^0 \int p(|r|, y - \xi(t+r)) \phi(y) dy dr < \infty$$

and hence $\Xi_{-\infty,t}^I$ is a.s. locally finite.

(b) Without loss of generality we can take $t = 0$ and $\psi(y) = \tilde{\psi}(y_0)$, $\phi(y) = \tilde{\phi}(y_0)$, when $\tilde{\phi}, \tilde{\psi}$ are continuous with compact support. Then for $d \geq 3$, the above gives

$$\begin{aligned} & \lim_{s \rightarrow -\infty} \int \int \mu(\psi)\mu(\phi)R_{s,0}(y,d\mu)\mu_\lambda^s(dy) \\ &= \lim_{s \rightarrow -\infty} \int \int \psi(\xi)\mu(\phi)R_{s,t}^\xi(d\mu)P_{s,\mu_\lambda^s}(d\xi) \\ &= \gamma \lim_{s \rightarrow -\infty} \int \int \int_s^0 p(2|r|, z-y))\phi(z)\psi(y)drdzdy \\ &< \infty. \end{aligned}$$

(c) See Theorem 6.4 of Dawson and Perkins [30]. ■

THEOREM 5.4. (*Dawson and Perkins [30]*). *The stationary random measure X_∞ of the super-Brownian motion in \mathbb{R}^d is an infinitely divisible random measure with intensity λ and with canonical measure given by $\hat{R}(B) = R_{-\infty,0}(\{\pi_0\mu \in B\})$, where $\pi_0\mu(A) := \mu(\{\xi : \xi_0 \in A\})$, $A \in \mathcal{B}(\mathbb{R}^d)$, $\mu \in M_p(C^d)$.*

REMARK 5.5. This leads to a description of the super-Brownian motion in equilibrium as a countable collection of infinite clan measures.

An infinite clan containing an individual located at 0 at time 0, denoted by $\Xi_{-\infty,0}^I$, has law $\int R_{-\infty,0}^\xi \Pi_0(d\xi)$. It is constructed by running a Brownian trajectory, ξ , backwards to time $-\infty$ and then collecting the mass at time 0 corresponding to $R_{-\infty,0}^\xi$. (Here $R_{-\infty,0}^\xi$ is the limiting law in Proposition 5.2(a) with $t = 0$.)

5.4.4. Clan Dynamics. The clan $\{\Xi_{-\infty,t}^I\}_{t \geq 0}$ then evolves as a historical Brownian motion with initial condition $\Xi_{-\infty,0}^I$. It is an easy consequence of this description that the infinite clan is self-similar (5.2) in the sense that

$$(5.2) \quad K^{-2}\Xi_{-\infty,0}^I(\{y : y_0 \in KA\}) \stackrel{\mathcal{D}}{=} \Xi_{-\infty,0}^I(\{y : y_0 \in A\}) \quad \forall A \in \mathcal{B}(\mathbb{R}^d).$$

There has been considerable recent interest in the dynamics of these infinite clans and this has led to a second dimensional dichotomy - namely clan recurrence or transience. To describe this consider the total weighted occupation time that a given clan spends in the unit ball at the origin $B(0,1)$. In dimensions $d > 2$, (5.2) and spherical symmetry implies that the infinite clan of an individual at the origin and has mean density proportional to

$$\frac{1}{|x|^{d-2}}.$$

Therefore if we exclude the clan mass initially in a ball of radius 2 (which by the extinction property of Feller branching has a.s. finite lifetime), then

$$\begin{aligned} & E[\int_0^\infty \Xi_{-\infty,t}^I(\{y : y_t \in B(0,1), y_0 \notin B(0,2)\})dt] \\ &= \int_0^\infty \int_{B(0,2)^c} \frac{1}{|y|^{d-2}} \left[\int_{B(0,1)} p_t(x-y)dx \right] dydt \\ &= \int_{B(0,2)^c} \left[\int_{B(0,1)} \frac{1}{|x-y|^{d-2}} dx \right] \frac{1}{|y|^{d-2}} dy \\ &< \infty \text{ iff } d \geq 5. \end{aligned}$$

Hence in dimensions $d \geq 5$, $\int_0^\infty \Xi_{-\infty,t}^I(\{y : y_t \in B(0,1)\})dt < \infty$, a.s. - this behavior is called *clan transience*. Analogous results of Stöckl and Wakolbinger [117] showed that the corresponding infinite clan for a branching particle system in dimensions $d \geq 5$ gives positive mass to the unit ball over only a finite time horizon.

Recent results on branching particle systems of Stöckl and Wakolbinger [117] (also see Krone [74]) suggest that in dimensions $d = 3, 4$

$$\int_0^\infty \Xi_{-\infty,t}^I(\{y : y_t \in B(0,1)\})dt = \infty \text{ a.s.},$$

that is, the SBM is clan recurrent.

5.5. Universality of the Extinction-Persistence Dichotomy. Consider the single type branching model which is given by the system of stochastic differential equations

$$(5.3) \quad dx_i(t) = \sum_{j \in S} q_{i,j}[x_j(t) - x_i(t)]dt + \sqrt{2g(x_i(t))}dw_i(t), \\ x_i(0) \equiv \theta > 0,$$

where S is a countable Abelian group, $\{q_{i,j}\}_{i,j \in S}$ are the transition rates of an irreducible random walk on S ($q_{i,j} = q_{0,i-j}$ with symmetrized kernel $\hat{q}_{i,j} = \frac{1}{2}(q_{i,j} + q_{j,i})$) and $\{w_i\}$ are independent Brownian motions. Recall that $g(x) = x$ corresponds to continuous state branching and then the above system is precisely the “super-random walk”.

THEOREM 5.5. *Assume g is locally Lipschitz, $g(x) > 0$ for $x \in (0, \infty)$, and*

$$\limsup_{|r| \rightarrow \infty} \frac{g(r)}{r^2} < \infty.$$

- (a) (Cox and Greven [12], Shiga [114]) *Assume that the symmetrized kernel \hat{q} is transient and for all $i \in S$, $x_i(0) = \theta$. Then $\mathcal{L}(\{x_i(t)\}_{i \in S})$ converges weakly to ν_θ as $t \rightarrow \infty$ and $\langle \nu_\theta, x_0 \rangle = \theta$. Moreover $\{\nu_\theta\}_{\theta \in [0, \infty)}$ is the set of extreme points of the set of shift invariant, locally finite mean stationary measures.*
- (b) (Cox, Fleischmann and Greven [14]). *Assume that the symmetrized kernel \hat{q} is recurrent, $\mathcal{L}(x(0))$ is homogeneous and $\theta = E(x_0(0)) < \infty$. Then*

$$\mathcal{L}(\{x_i(t)\}_{i \in S}) \Rightarrow \delta_0 \text{ as } t \rightarrow \infty$$

in the sense of vague convergence. (Here $\mathbf{0}$ denotes the identically zero configuration.)

REMARK 5.6. Theorem 5.5(b) suggests that the case in which the underlying random walk is recurrent is characterized by growing clusters (c.f. Klenke [71]). The exact form of these clusters depends on the fine structure of the random walk. We will give some partial results on this question in the next section.

5.6. Thermodynamic Limit Approach to Super-Random Walk. In the previous section we analyzed the structure of the equilibrium based on the study of the spatial distribution of all relatives of a given individual at equilibrium. We can think of this as a “bottom-up” approach. Another starting point is the “top-down” approach analogous to the thermodynamic limit of statistical physics. The idea here is to first determine the equilibria of finite subsystems given fixed boundary conditions (in this context this would involve a fixed rate of migration across the boundary of the subsystem) and then to let the size of the subsystem go to infinity. A related question concerns the convergence to equilibrium starting from a class of spatially homogeneous initial conditions. Finally another question is whether the large scale behavior of critical branching systems is a result of their very special structure or whether it is representative of a larger “universality class”. As a first step toward addressing these questions they have been reformulated in Dawson and Greven [35] for a class of super-random walks in which a fairly complete rigorous analysis is possible.

5.6.1. Hierarchical Mean-Field Limit. This phenomenon has been explored in depth in the special case in which S is the hierarchical group $S = \Omega_N = \{\xi = (\xi_1, \xi_2, \dots) : \xi_i \in \{0, \dots, N-1\}, \xi_i = 0 \text{ a.a. } i\}$ (a.a means for all but finitely many) and

$$q_{\xi, \xi'} = \sum_{j \geq d(\xi, \xi')} \frac{c_{j-1}}{N^{2j-1}}, \quad \sum \frac{c_j}{N^j} < \infty.$$

Here $d(\xi, \xi') := \inf\{k : \xi_j = \xi'_j \forall j \geq k\}$ if $\xi = (\xi_1, \xi_2, \dots), \xi' = (\xi'_1, \xi'_2, \dots)$. In this case the random walks are transient for all sufficiently large N if and only if $\sum \frac{1}{c_j} < \infty$. This suggests that the qualitative behavior is similar for at least all large values of N and suggests looking at the “hierarchical mean-field limit”, $N \rightarrow \infty$. The advantage of studying this limiting structure is that in this case the “time scales completely separate”.

We begin by considering the sequence of “block averages” in a natural time scale

$$x_{\xi, k}^N(t) := \frac{1}{N^k} \sum_{\{\xi' : d(\xi', \xi) \leq k\}} x_{\xi'}(N^k t).$$

It is proved in Dawson and Greven [26] that $x_{\xi, k}^N(t)$ converges as $N \rightarrow \infty$ to a diffusion satisfying

$$dY_k(t) = c_k(\theta - Y_k(t))dt + \sqrt{2F_k(Y_k(t))}dw_k(t),$$

where $F_k(\theta) = \int_0^\infty F_{k-1}(x)\Gamma_\theta^{k-1}(dx)$, $\Gamma_\theta^k(\cdot)$ is the unique equilibrium of the diffusion $Y_k(t)$ and $F_0(\theta) = g(\theta)$. In the case $g(\theta) = \frac{\gamma}{2}\theta$, $F_k(\theta) = \frac{\gamma}{2}\theta$ for every k (*fixed point property of Feller branching*).

Consider the k -block average in the time scale $N^j s(N)$, where $\frac{s(N)}{N} \rightarrow 0$, $s(N) \uparrow \infty$ with $j > k$. That is,

$$x_{\xi, k}^{N, j} := \frac{1}{N^k} \sum_{\{\xi' : d(\xi', \xi) \leq k\}} x_{\xi'}(s(N)N^j).$$

The idea is that the higher block averages do not change in time scale $N^k t$ and the k -block subsystem reaches a quasi-equilibrium in time $s(N)N^k$ which evolves in time scale N^{k+1} according to $x_{\xi, k}^{N, k+1}$, etc. Then as $N \rightarrow \infty$, $\{x_{\xi, k}^{N, j}\}_{k=j, \dots, 0}$ converges

to a \mathbb{R}_+ -valued Markov chain, $\{Z_k^j(\theta)\}_{k=j+1,\dots,0}$, $Z_{j+1}^j = \theta$, called the interaction chain.

THEOREM 5.6. (*Dawson and Greven [35]*) Let $g(x) = x$. Then as $N \rightarrow \infty$,

$$\mathcal{L}((x_{\xi,k}^{N,j})_{k=0,\dots,j+1}) \Rightarrow \mathcal{L}((Z_k^j)_{k=0,\dots,j+1}),$$

where $\{Z_k^j\}_{k=0,\dots,j}$ is a reverse time Markov chain (and martingale) with $Z_{j+1}^j = \theta$ and with Gamma transition density

$$(5.4) \quad \begin{aligned} \Gamma^k(z, z') &= \frac{m^\lambda}{\Gamma(\lambda)} (z')^{\lambda-1} \exp(-mz') \\ m &= \frac{c_k}{\gamma}, \quad \lambda = \frac{c_k z}{\gamma}. \end{aligned}$$

REMARK 5.7. $\Gamma^k(z, \cdot)$ has first moment z and variance $\frac{z\gamma}{c_k}$.

REMARK 5.8. The analogue of the persistence-extinction dichotomy is given by the limiting behavior of Z_0^j when $Z_j^j = \theta$ is fixed and we let $j \rightarrow \infty$. If $\sum \frac{1}{c_k} = \infty$, then $Z_0^j \rightarrow 0$ as $j \rightarrow \infty$. On the other hand when $\sum \frac{1}{c_k} < \infty$, then there exists an entrance law for $\{Z_k^j\}$ with constant mean θ . To understand this we will carry out the second moment calculation. By (5.4)

$$\begin{aligned} E[Z_0|Z_j] &= Z_j \\ E[(Z_j - Z_{j+1})^2] &= E[E[(Z_j - Z_{j+1})^2|Z_{j+1}]] \\ &= \frac{\gamma E[Z_{j+1}]}{c_j} = \frac{\gamma\theta}{c_j}. \end{aligned}$$

Therefore for $j \geq j'$,

$$E[(Z_j - Z_{j'})^2] = \sum_{\ell=j'}^j \frac{\gamma\theta}{c_\ell}$$

and so $\{Z_j\}$ converges in L^2 if and only if $\sum \frac{1}{c_k} < \infty$. In the convergent case, we can construct the unique entrance Markov martingale $\{Z_j\}_{j \in \mathbb{Z}_+}$ satisfying $Z_j \rightarrow \theta$ a.s. as $j \rightarrow \infty$.

In order to give an idea of these results we consider the special case $c_j = c^j$ with $c \leq 1$. The case $c = 1$ is analogous to the two dimensional lattice and exhibits “diffusive clustering” which was first discovered by Cox and Griffeath [15]. The case $c < 1$ corresponds to the one dimensional case and exhibits “concentrated clustering”.

THEOREM 5.7. (*Dawson and Greven [35]*) (a) (“Diffusive clustering”) Assume that $c_k \rightarrow c$. Then as $j \rightarrow \infty$,

$$\mathcal{L}\left(\frac{Z_{j+1-[c(j+1)]}^j}{j} \middle| Z_0^j \geq \varepsilon\right) \Rightarrow \mathcal{L}(Z(\alpha) : \alpha \in (0, 1)),$$

where $(Z(\alpha))$ is a one dimensional diffusion with $Z(0) = \theta$ and time inhomogeneous generator

$$\begin{aligned} G_\alpha &= a_\alpha(x) \frac{\partial}{\partial x} + b_\alpha(x) \frac{\partial^2}{\partial x^2} \\ a_\alpha(x) &= 2x \left(\frac{(1-\alpha)^{-1}}{\exp(\frac{x}{1-\alpha}) - 1} \right) \\ b_\alpha(x) &= x(2 + a_\alpha(x)). \end{aligned}$$

(b) (“Low dimensional clumping”) In the case $c_k = c^k$ with $c < 1$, there exists a probability law \hat{F}_∞ on \mathbb{R}_+ such that as $j \rightarrow \infty$

$$\hat{P}^j(\delta) \Rightarrow \hat{F}_\infty(\delta) \quad \forall \delta > 0,$$

where $\hat{P}^j(\delta) := \theta^{-1} P(Z_0^j 1(c^j Z_0^j > \delta))$.

(c) (Equilibrium clan) Assume that $\sum \frac{1}{c_k} < \infty$. Let $M_{0,-k}^N$ denote the family mass of a randomly chosen individual (i.e. under the Palm measure) at a given site for which the last common ancestor lived in $(\tau_{k-1}^N, \tau_k^N]$, where τ_k^N is the exit time from a ball of radius k for the ancestral path. Then as $N \rightarrow \infty$, $M_{0,-k}^N$ converges in distribution to $M_{-k}^* = \sum_{j=k}^{\infty} Z_k^j(m_j)$, where the Z_k^j are independent with random initial condition $Z_j^j = m_j$ and the $\{m_j\}$ are independent exponentially distributed with means $\frac{\gamma}{2c_j}$.

5.7. Universality in the Hierarchical Mean-Field Setting. The results on the large time behavior described in this section depend heavily on the tools of infinitely divisible random measures. However there is reason to believe that this behavior is shared by a large “universality class” including systems of the form (5.3) when $g(x) \neq \gamma x$ and therefore infinitesimal divisibility breaks down. Based on results of Baillon, Clement, Greven and den Hollander [2], it is proved in Dawson and Greven [35] that the results (a), (b) of the above are true for the class of all g ’s satisfying $g(0) = 0$, $g(x) > 0 \quad \forall x \in (0, \infty)$, g is locally Lipschitz and

$$\lim_{x \rightarrow \infty} \frac{g(x)}{x} = \gamma > 0.$$

Moreover, (c.f. Baillon, Clement, Greven and den Hollander [2]) in the recurrent regime (in the hierarchical mean-field setting), there exist d_k so that

$$\left\| \frac{F_k(x)}{d_k x} - 1 \right\|_\infty \xrightarrow{k \rightarrow \infty} 1.$$

6. A Survey of Interactive Branching Systems

In the previous chapters we have explored in some detail both the small scale and large scale behaviors of super-Brownian motion in \mathbb{R}^d . This has been made possible by the use of the special structure of super-Brownian motion, in particular, the fact that we can explicitly compute moments and represent the Laplace functional in terms of the solutions to the log-Laplace equation. Spatially distributed systems arise in many fields of application, for example, reaction-diffusion systems, ecology, population genetics, diffusion of innovations, evolutionary theory, and immunology. An interesting recent example is the notion of *viral quasi-species* in immunology - this is a rapidly evolving “cloud” in a high dimensional phenotype space (c.f. Eigen

[46]). Another class of measure-valued processes arise in modeling populations involving hierarchically structured objects. Examples of this type arise in genetics (internal structures in the genetic material at one level and phenotypic structures at another level) and distributed and hierarchically structured information systems (such as network file systems).

The special independence assumptions which result in the multiplicative property are not satisfied in many of these applications which lead to interactive systems. The first question which we will address is the formulation of measure-valued processes to describe the types of interaction mechanisms which arise in some of these applications and we will survey some of the methods which have been developed to study them. The second basic question concerns the small scale and large scale behavior of these more complex systems and to what extent the structures which we have identified for super-Brownian motion have counterparts in these interactive systems.

The systematic study of these questions is a major area of current research. Some initial steps have been taken but many open problems remain - both in the characterization of the processes and in the study of their behavior. In this chapter we will outline the current state of the subject including a discussion of some of the ideas and methods which have been introduced. However space does not permit more detailed development and proofs will not be included.

We will begin with an overview of some basic classes of interaction at the particle level. This includes

- state dependent mean offspring size
- finite range competition due to locally finite carrying capacity
 - zero range interactive killing
 - the Fisher-Wright and Fleming-Viot models
- interactive particle motions
- particle- medium and particle-medium-particle interactive motion
- interactive branching
 - interactive branching in \mathbb{R}^1
 - catalytic branching
 - mutual-catalytic branching
- interactive Fleming-Viot processes

We have earlier discussed hierarchically structured populations. Interactions can also occur at higher levels and also between levels. We will briefly describe

- multilevel branching
- interactive multilevel branching.

One approach to the questions raised above is to think of super-Brownian motion as a basic building block in the way Brownian motion serves as a basic ingredient in the theory of stochastic differential equations for $x(t) = (x_1(t), \dots, x_n(t))$ given by

$$dx_i(t) = b_i(x(t))dt + \sum_j \sigma_{ij}(x(t))dw_j(t).$$

Under well-known conditions introduced by Itô or Stroock and Varadhan the solution of this system is a Markov process with generator

$$Gf(x) = \sum_{i,j} a_{ij}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum_i b_i(x) \frac{\partial f(x)}{\partial x_i}.$$

This would naturally suggest modeling more complex systems in terms of stochastic partial differential equations. Recall from Section 1 that super-Brownian motion can be viewed as the solution of the SPDE

$$dX_t(x) = \frac{c}{2} \Delta X_t dt + \sqrt{\gamma X_t(x)} \dot{W}(dt, dx),$$

where \dot{W} is space-time white noise. The next step in such a program would be to replace the *constant coefficients* c and γ with state-dependent functions. As we have seen above the SPDE has a strict sense solution in the constant coefficient case only in one dimension but that one can formulate a martingale problem which has a meaning in any dimension. Following this approach leads to the class of martingale problems with pre-generators of the form

$$\begin{aligned} GF(\mu) &= \int (A_0(\mu, x) + A_1(\mu, x) + A_2(\mu, x)) \frac{\delta F(\mu)}{\delta \mu(x)} \mu(dx) \\ &\quad + \int \int \frac{\delta^2 F(\mu)}{\delta \mu(x) \delta \mu(y)} Q(\mu; dx, dy), \end{aligned}$$

where $A_0(\mu, x)$ is a real-valued function of μ and x related to the density dependent mean offspring size at x when the state is μ , $A_1(\mu, x)$ is a first order differential operator describing the spatial drift of a mass element at x when the population state is μ (for example, the attraction or repulsion between particles, or particles driven by a random turbulent motion), and $A_2(\mu, x)$ is a second order differential operator describing the state dependent spatial diffusion of a population element at x when the population state is μ . Finally, $Q(\mu, dx, dy)$ describes the rate of change of the $\text{cov}(X_t(dx), X_t(dy))$. The case of super-Brownian motion corresponds to

$$\begin{aligned} A_0 &= 0 \\ A_1 &= 0 \\ A_2 &= \frac{1}{2} \Delta \\ Q(\mu, dx, dy) &= \mu(dx) \delta_x(dy). \end{aligned}$$

6.1. Particle Motions with Inter-Particle Interactions and Particle Random Medium Interaction. In this section we consider branching particle systems in which there is no interaction in the branching mechanism, but in which the particle motions are interactive.

Consider the system of branching (independent) Brownian motions $\{\xi^\alpha\}_{\alpha \in I}$ with branching rate γ . Let $b : \mathbb{R}_+ \times M_F(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times M_F(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ (the space of $d \times d$ matrices over \mathbb{R}), and $g : \mathbb{R}_+ \times M_F(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ be measurable maps.

For $\alpha \in I$ assume the “ α^{th} individual” in the population, Y^α obeys the Itô equation driven by the trajectory of the α^{th} individual in the system of branching Brownian motions, ξ^α ,

$$(6.1) \quad dY^\alpha(t) = \sigma(t, Z_t^\varepsilon, Y_t^\alpha) d\xi_t^\alpha + b(t, Z_t^\varepsilon, Y_t^\alpha) dt + \int g(t, Z_t^\varepsilon, y - Y_t^\alpha) W(dy, dt),$$

where Y^α jumps to the cemetery state when ξ^α does and

$$(6.2) \quad Z_t^\varepsilon(A) = \eta_\varepsilon \sum_{\alpha \sim t} 1_A(Y_t^\alpha).$$

This corresponds to an exchangeable (but not independent) system of motions. In these equations the first two terms describe the motions of particles which form an exchangeable system of diffusions. In the last term W denotes a \mathbb{R}^d -valued space-time white noise and this contribution to the motion of the α 'th particle can be interpreted as perturbation of the motion by a random (turbulent) force. Under the same scaling of the mass, branching rate and initial measures as in Section 1 one expects that these particle systems will converge to measure-valued processes. We will now consider some special cases in which the limiting dynamics have been characterized.

We first consider the case in which $g = 0$. In this case in the absence of branching ($\gamma = 0$) a law of large numbers is expected and the limit would be given by the weak solution of a McKean-Vlasov equation. If $\gamma \geq 0$ and b and σ satisfy appropriate Lipschitz and growth conditions on b and σ , (6.1) and (6.2) have a pathwise unique solution. If we write $Y_t^\alpha(\omega) = Y_t(\xi^\alpha, \omega)$, let $Y_0^\alpha = \xi_0^\alpha$ and take $\varepsilon \rightarrow 0$, the above equations become

$$(6.3) \quad Y_t(y, \omega) = y(0) + \int_0^t \sigma(s, Z_s, Y_s) dy(s) + \int_0^t b(s, Z_s, Y_s) ds,$$

$H_t - a.a. y \forall t \geq 0$

$$(6.4) \quad Z_t(A) = \int 1(Y_t(y, \omega) \in A) H_t(dy) \quad \forall A \text{ Borel in } \mathbb{R}^d, t \geq 0.$$

Here $(H_t, t \geq 0)$ is a historical Brownian motion with branching rate γ . The precise meaning of the stochastic integral in (6.3) takes some care, but as a typical y chosen according to H_t should be a Brownian motion up to time t , this should be an Itô integral. This is true providing you work with respect to the “Campbell measure”, $H_t(dy)d\mathbb{Q}/\mathbb{Q}(H_t(1))$ (c.f. Perkins [104]).

EXAMPLE 6.1. Suppose a particle at y of mass $\mu(dy)$ exerts a drift $\hat{b}(x - y)\mu(dy)$ on a particle at x and the diffusion rate at x is the approximate local density at x . Then we would set

$$b(t, \mu, x) = \int \hat{b}(x - y)\mu(dy), \quad \sigma(t, \mu, x) = \int p_\varepsilon(x - y)\mu(dy)I_{d \times d},$$

where p_ε is the heat kernel. If \hat{b} is bounded and Lipschitz, the hypothesis of Theorem 6.1 below are satisfied.

We denote $|x|$ to be the Euclidean norm of x . Let $Lip(\mathbb{R}^d) = \{f : \mathbb{R}^d \rightarrow \mathbb{R} : |f(x) - f(y)| \leq |x - y|, |f(x)| \leq 1, \forall x, y \in \mathbb{R}^d\}$. If $\mu, \nu \in M_F(\mathbb{R}^d)$, $d_0(\mu, \nu) = \sup\{|\mu(f) - \nu(f)| : f \in Lip(\mathbb{R}^d)\}$ denotes the Vasserstein metric. Here $(M_F(\mathbb{R}^d), d_0)$ is a Polish space and d_0 induces the usual weak topology on $M_F(\mathbb{R}^d)$.

THEOREM 6.1. (*Theorem 4.10 of Perkins [105]*). *Assume $(H_t, t \geq 0)$ is a historical Brownian motion with law $\mathbb{Q}_m^{\gamma, 0}$. Suppose L is a non-decreasing function such that*

$$\begin{aligned} & |b(t, \mu, x) - b(t, \mu', x')| + |\sigma(t, \mu, x) - \sigma(t, \mu', x')| \\ & \leq L(t \vee \mu(1) \vee \mu'(1))(|x - x'| + d_0(\mu, \mu')). \end{aligned}$$

Then there is a pathwise unique solution (Y, Z) of (6.3), (6.4). Z is a predictable function of H and the law of Z depends only on (m, γ, b, σ) .

The proof proceeds by showing that solutions to (6.3) and (6.4) are fixed points of a contraction on an appropriate Polish space of stochastic processes. By replacing (6.4) with

$$(6.5) \quad K_t(B) = \int 1(Y_{\wedge t}(y, \omega) \in B) H_t(dy), \quad \forall \text{ Borel } B \subset C(\mathbb{R}_+, \mathbb{R}^d)$$

$$H_t - a.a. \text{ } y, \quad \forall t \geq 0,$$

we can obtain a pathwise existence and uniqueness result for the historical process associated with Z_t . A stochastic calculus for the integral in (6.3) (c.f. Perkins [105]) shows that the solutions Z and K satisfy natural martingale problems which extend (1.1) (use only smooth test functions) and (2.1), respectively. Under a non-degeneracy assumption on $a = \sigma\sigma^*$, the martingale problem for K is well-posed (c.f. Theorem 5.13 of Perkins [105]). This allows one to incorporate the interactions and change of measure techniques of Theorem 6.2 below into these models and also leads to a proof of the weak convergence of interactive branching particle described at the beginning of this section (c.f. Lopez [88]). One may also allow the branching rate $\gamma = \gamma(t, x)$ to depend smoothly on space and time, and on the solution in a more restrictive manner (c.f. Perkins [105]). In one dimension (for example) the martingale problem satisfied by these solutions will formally correspond to the parabolic spde:

$$\frac{\partial u(t, x)}{\partial t} = \frac{1}{2}\sigma^2(t, u(t), x)u_{xx}(t, x) + \bar{b}(t, u(t), x)u(t, x)$$

$$+ h(t, u(t), x)u(t, x) + (\gamma(t, x)u(t, x))^{1/2}\dot{W}(t, x),$$

where

$$\bar{b}(t, u(t), x) = \frac{\partial}{\partial x}\sigma^2(t, u(t), x) - b(t, u(t), x) \quad \text{and}$$

$$h(t, u(t), x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\sigma^2(t, u(t), x) - \frac{\partial}{\partial x}b(t, u(t), x)$$

may be handled by introducing an appropriate $\theta(t, u(t), x)$ through an analogue of Theorem 6.2 below. To make the correspondence rigorous we must prove the existence of a continuous density $u(t, x)$ for solutions to (6.3) and (6.4). For σ constant this case has been done by Lopez [88].

Donnelly and Kurtz [39], [40] have incorporated some of these ideas in their “look-down processes” and have introduced spatial interactions into a general class of population models including Fleming-Viot processes and the above measure-valued branching processes.

We next turn to the case in which $g(\cdot, \cdot, x) = g(x) \neq 0$ but $\gamma = 0$ and $\sigma = I$. This has been studied by Kotelenetz [73] and leads to the weak spde.

$$\langle X(t), \phi \rangle = \langle X(0), \phi \rangle + \int_0^t \langle X(s), \frac{1}{2}\Delta\phi + b(\cdot, X(s)) \cdot \nabla\phi(\cdot) \rangle ds$$

$$+ \int_0^t \langle X(s), g(x-y)\nabla\phi(x) \rangle \cdot W(dy, ds).$$

The case in which $g(\cdot, \cdot, x) = g(x) \neq 0$, $\gamma \neq 0$ but in which $b = 0$, $\sigma \equiv 1$ has been studied by Wang [121]. For example in $d = 1$ it leads to a spde of the form

$$dX_t(x) = \frac{1}{2}\Delta X_t(x) + \sqrt{\gamma X_t(x)}W_1(dt, dx) - \nabla \cdot (X_t(x) \int g(y-x)W_2(dt, dy)),$$

where W_1 , W_2 are independent space-time white noises.

6.2. Nonlinear Birth and Death Rates. Consider a version of (1.1) in which the “mean offspring size” is biased by a $\mathcal{P}^X \times \mathcal{E}$ -measurable function: $\theta : \mathbb{R}_+ \times \Omega_X \times E \rightarrow \mathbb{R}_+$. In the notation of Theorem 1.1, (1.1) becomes for every $f \in D(A)$,

$$M_t^\theta(f) = X_t(f) - m(f) - \int_0^t \int (Af(x) + \theta(s, \omega, x)f(x))X_s(dx)ds$$

is an (\mathcal{F}_t^X) -martingale starting at zero and with quadratic variation

$$\langle M^\theta(f) \rangle_t = \int_0^t X_s(\gamma f^2)ds.$$

We refer to this martingale problem as (MP_θ) . Solutions to (MP_θ) may be constructed as limits of branching particle systems in which the mean offspring size of a parent at x is $1 + \varepsilon\theta(s, \omega, x)$. Of course to prove such a convergence result it is very useful to know that solutions to (MP_θ) are unique in law.

THEOREM 6.2. (*Dawson [18], . . . , Evans and Perkins [53]*)

Let $\mathbb{P}_m = P_m^{\gamma, 0}$, assume

$$\int_0^t \int \theta(s, \omega, x)^2 X_s(dx)ds < \infty \quad \forall t > 0, \quad \mathbb{P}_m - a.s.$$

and let

$$R_t^\theta = \exp \left(\int_0^t \int \theta(s, \omega, x)dM(s, x) - \frac{\gamma}{2} \int_0^t \int \theta(s, \omega, x)^2 X_s(dx)ds \right).$$

If $\theta(s, \omega, x) \leq \theta_0$ for some $\theta_0 \in \mathbb{R}$, then R_t^θ is an (\mathcal{F}_t^X) -martingale under \mathbb{P}_m and $\frac{d\mathbb{P}_m^\theta}{d\mathbb{P}_m} \Big|_{\mathcal{F}_t^X} = R_t^\theta$ defines the unique solution, \mathbb{P}_m^θ , of (MP_θ) .

The proof of Theorem 6.2 is an easy application of Itô’s lemma and the uniqueness in (MP_0) , just as in the proof of the more familiar Girsanov theorem for diffusions.

We give several applications of this very useful result.

Application 1. Interacting Particles and SPDE. Interacting particle systems have been studied intensively since they were introduced by Spitzer and Dobrushin in the early 1970’s. Beginning with a conjecture of Rick Durrett in 1988, there has been considerable recent interest in showing that appropriate rescaled limits of these interacting particle systems converge weakly to solutions of stochastic partial differential equations.

If $M \in \mathbb{N}$, the contact process with range M is a set-valued process S_t taking values in the space of finite subsets of $M^{-1}\mathbb{Z}^d$ given by the following dynamics for a fixed value of $\lambda \geq 1$:

1. Particles die with rate 1.
2. A particle at x with rate λ chooses a site uniformly in $x + C_M = x + ([-1, 1]^d \cap M^{-1}\mathbb{Z}^d) / \{0\}$ and produces an offspring there if the site is vacant and suppresses the offspring if it is occupied.
3. Jump and death times at distinct sites fire independently.

For example S_t could represent the set of infected sites in a population, where a “death” corresponds to a recovery. It is known that there is a critical value of λ ,

$\lambda_c(M) \geq 1$, above which the process survives as $t \rightarrow \infty$ with positive probability and below which it dies out a.s. The precise value of $\lambda_c(M)$ is unknown. Bounds on its asymptotic behavior as $M \rightarrow \infty$ are given in Bramson et. al. [11]. Let S_t^N denote the contact process with $M = M_N$ and $\lambda = (1 + \frac{\theta}{N})$, where $\theta \geq 0$ is a fixed parameter and

$$M_N = \begin{cases} N^{\frac{1}{d}} & \text{if } d > 2 \\ (N \log N)^{\frac{1}{2}} & \text{if } d = 2. \\ N^{\frac{3}{2}} & \text{if } d = 1. \end{cases}$$

Consider now the usual ‘‘Brownian rescaling’’ of S_t^N as a measure-valued process, namely,

$$X_t^N(A) = N^{-1} \#(A \cap (\frac{1}{\sqrt{N}} S_{Nt}^N)), \quad A \subset \mathbb{R}^d.$$

If offspring were allowed to jump onto occupied sites, X_t^N is a minor variant of the system of branching Brownian motions introduced in Theorem 1.1 and one easily shows that X_t^N converges weakly to the solution of (MP_θ) with $A = \frac{\Delta}{6}$, $\gamma = 2$, and $\theta(s, \omega, x) \equiv \theta$. The suppression of jumps onto occupied sites in our rescaled model leads to additional killing due to local crowding effects. The greater the value of M_N , the smaller this effect. It is not hard to motivate the above choice of M_N by pretending that X_t^N has the same local properties as super-Brownian motion. This means that it has a density if $d = 1$ (c.f. Theorem 1.3), and a constant density on a random set with respect to the appropriate 2-dimensional Hausdorff measure function if $d > 1$ (c.f. Section 4). This suggests a killing rate proportional to $u(t, x)$ if $d = 1$ and proportional to some constant if $d > 1$. More precisely we have

THEOREM 6.3. *Assume $X_0^N \rightarrow X_0$ as $N \rightarrow \infty$.*

(a) (Mueller and Tribe [93]) If $d = 1$ and $X_0(dx) = u(0, x)dx$ for some continuous integrable $u(0, \cdot)$, then $X^N \Rightarrow X$ on $C(\mathbb{R}^+, M_F(\mathbb{R}^d))$, where $X_t(dx) = u(t, x)dx$ and u is the unique (in law) solution of

$$\frac{\partial u(t, x)}{\partial t} = \frac{1}{6} u_{xx}(t, x) + \sqrt{2u(t, x)} \dot{W}(t, x) + \theta u(t, x) - u(t, x)^2.$$

(b) (Durrett and Perkins (in preparation)) If $d > 1$ and X_0 is non-atomic, then $X^N \Rightarrow X$ on $C(\mathbb{R}^+, M_F(\mathbb{R}^d))$, where X is the unique solution of the following martingale problem:

$$X_t(\phi) = X_0(\phi) + M_t(\phi) + \int_0^t X_s(\frac{\Delta\phi}{6}) ds + (\theta - c(d)) \int_0^t X_s(\phi) ds,$$

where

$$M_t(\phi) \text{ is a martingale satisfying } \langle M(\phi) \rangle_t = 2 \int_0^t X_s(\phi^2) ds.$$

Here ϕ has bounded continuous partial derivatives of order 2 or less.

If $\{U_n\}$ is a random walk with step size uniform on $[-1, 1]^d$, then

$$c(d) = 2^{-d} \sum_{n=1}^{\infty} P(U_n \in [-1, 1]^d) \quad \text{if } d > 2, \quad \text{and } c(2) = \frac{3}{2\pi}.$$

Item (b) can be used to refine the asymptotic bounds on $\lambda_c(M)$ given in Bramson et. al. [11]. In (a), Mueller and Tribe [93] showed a stronger result, namely the weak convergence of the approximating densities to X^N to the limiting density.

REMARK 6.1. See Mueller and Tribe [93] for a related limit theorem for the rescaled long range voter model in one dimension.

Theorem 6.2 gives us the existence and uniqueness of solutions to the stochastic partial differential equation in (a) and a formula for the Radon-Nikodym derivative of the law of the solution with respect to that of the density of super-Brownian motion ($d = 1$). Here one takes $\theta(s, \omega, x) = \theta - U(\omega_s, x) \leq \theta$, where $U(\mu, x)$ is the density (at x) of the absolutely continuous part of μ (c.f. Theorem 3.10 of Evans and Perkins [53]).

Application 2. Zero Range Interactive Killing.

(6.6)

$$\frac{\partial u^{(i)}(t, x)}{\partial t} = \frac{1}{2} u_{xx}^{(i)}(t, x) + \sqrt{u^{(i)}(t, x)} \dot{W}^{(i)}(t, x) - \lambda u^{(1)}(t, x) u^{(2)}(t, x), \quad i = 1, 2.$$

Consider the system of stochastic partial differential equations in \mathbb{R}^1 , where $\dot{W}^{(1)}$ and $\dot{W}^{(2)}$ are independent white noises on $\mathbb{R}_+ \times \mathbb{R}$. This could model the spatial distribution of two species competing for resources so that inter-species collisions are unhealthy for both types. It is easy to prove a two-type analogue of Girsanov's formula and use it to describe the law of the unique solution of (6.6) (c.f. Theorem 3.9 of Evans and Perkins [53]). The collision local time introduced in Section 4.8 allows us to formulate a martingale problem on Ω_X^2 which extends (6.6) to higher dimensions:

$$(6.7) \quad M_t^i(f) = X_t^i(f) - m_i(f) - \int_0^t X_s^i\left(\frac{\Delta f}{2}\right) ds - \lambda L_t(X^1, X^2)(f), \quad i = 1, 2$$

is a continuous martingale starting at zero such that

$$\langle M^i(f_i), M^j(f_j) \rangle_t = \delta_{ij} \int_0^t X_s^i(f_j^2) ds$$

for all $f, f_1, f_2 \in C^2(\mathbb{R}^d)$ which are constant outside a compact set. Refer to the above martingale problem as collision system (CS).

A solution (X^1, X^2) of (CS) is easily bounded above by a pair of independent super-Brownian motions (Y^1, Y^2) . Since Y^1 and Y^2 will only collide if $d < 6$ (c.f. Theorem 4.10). (CS) is only of interest when $d < 6$.

By analogy with Application 1, solutions should also arise as the weak limit of the following interacting particle systems: red and blue particles occupy sites in the lattice $L = \lambda^{1/d} N^{-2/d} \mathbb{Z}^d$. Each particle dies at rate N . With rate N a particle at x produces an offspring at a site chosen uniformly in $x + ([\frac{-\sqrt{3}}{\sqrt{N}}, \frac{\sqrt{3}}{\sqrt{N}}]^d \cap L)$. The offspring survives if the site is not occupied by a particle of the opposite color.

If $d \geq 4$, there can be some serious problems with this particle system. In fact, solutions to (CS) do not exist for $d \geq 4$ (c.f. Theorem 5.3 of Evans and Perkins [53]). There is a simple explanation for this. Choose a path y according to H_t^2 , the historical Brownian motion associated with $Y^2 \geq X^2$. Then y is a Brownian trajectory independent of the super-Brownian motion Y^1 . By Theorem 4.10) if $d \geq 4$, y will not encounter Y^1 and so will not be killed by the X^1 population.

This means that $Y_t^2 = X_t^2$ a.s. and so by symmetry and continuity, we deduce that $(X^1, X^2) = (Y^1, Y^2)$ a.s. Since (Y^1, Y^2) does not satisfy (CS) ($L(Y^1, Y^2) \neq 0$), (CS) has no solution.

If $d = 2$ or 3 , solutions to (CS) are constructed in Evans and Perkins [53] as weak limits of solutions to $(CS)^\varepsilon$ in which $L(X^1, X^2)$ is replaced by $L^\varepsilon(X^1, X^2)$ (c.f. Section 4.8). The approximating solutions exist by the two-species analogue of Theorem 6.2. Although uniqueness in (CS) remains unresolved for $d = 2, 3$, uniqueness in the martingale problem for the historical processes associated with solutions of (CS) was recently proved in Barlow, Evans and Perkins [6] using a strong equation approach. For $d = 3$ the solution is known not to be absolutely continuous with respect to $\mathbb{P}_{m_1} \times \mathbb{P}_{m_2}$, so the Girsanov approach used for $d = 1$ cannot work (c.f. Evans and Perkins [53]).

Application 3. Finite Carrying Capacity. The critical branching condition which corresponds to the assumption that the mean number of offspring produced when a particle branches is exactly one is not natural in the study of biological populations. A more realistic scenario is that the intrinsic population growth is supercritical, but due to competition for limited resources, the death rate increases with the population size. We will consider a variant of the usual logistic equation in which the “carrying capacity” is 1. This can be formulated in terms of the following martingale problem (MP_{I.C.}):

$$\forall f \in D(A) \quad M_t^{I.C.}(f) = X_t(f) - m(f) - \int_0^t \int (A + c(1 - X_s(1))) f(x) X_s(dx) ds$$

is a continuous martingale with increasing process

$$\langle M^{I.C.}(f) \rangle_t = \int_0^t X_s(\gamma f^2) ds.$$

Using Theorem 6.2, it is easy to show that this is a well-posed martingale problem (take $\theta(s, \omega, x) = c(1 - \omega_s(1))$). The parameter c measures how fast the corresponding deterministic system reaches equilibrium. If we consider the limit $c \rightarrow \infty$, then we obtain a measure-valued process which is equivalent to the super-Brownian motion conditioned to have total mass 1, that is, with state space $M_1(E)$. This process is called the Fleming-Viot process associated with parameters (A, γ) . The above limit result is proved below as Theorem 6.6, but first a few words are in order about Fleming-Viot processes.

This is a class of probability-measure-valued processes related to branching systems which is used in population genetics to model mutation, selection, and random sampling in the distribution of genotypes in a population. The locally compact state space, E , is the space of genotypes and the Feller processes, ξ (with generator A), represent mutation, that is, the displacement of the type of an offspring from that of its parent. The Fleming-Viot process, has a martingale problem characterization analogous to that of Theorem 1.1 (c.f. Chapter 10 of Ethier and Kurtz [49]). Let $\Omega_X^1 = C(\mathbb{R}^+, M_1(E))$ and we abuse notation slightly and let X_t and \mathcal{F}_t^X denote the coordinate maps and canonical filtration on Ω_X^1 , respectively.

THEOREM 6.4. *Let $\gamma : \mathbb{R}^+ \rightarrow (0, \infty]$ be a continuous function (called the sampling rate) such that $\gamma(t) = \infty$ for all $t \geq t_\gamma = \inf\{t : \gamma(t) = \infty\}$ and $0 < t_\gamma \leq \infty$. There*

exists a unique probability, $\bar{\mathbb{P}}_m^\gamma$ on Ω_X^1 such that for every $f \in D(A)$,

$$M_t(f) = X_t(f) - m(f) - \int_0^t X_s(Af)ds, \quad t < t_\gamma$$

is an (\mathcal{F}_t^X) -martingale starting at zero and with quadratic variation

$$\langle M(f) \rangle_t = \int_0^t \gamma_s [X_s(f^2) - (X_s(f))^2] ds, \quad t < t_\gamma,$$

and $X_t = X_{t_\gamma -}$ for $t \in [t_\gamma, \infty)$.

The above laws arise as weak limits of a constrained “branching particle” system in which the total number of particles is fixed. For example if $\gamma_t \equiv \gamma$ is constant, then at each generation the N “particles” produce an offspring vector which is multinomial with parameters N , $p_i = \frac{1}{N}$ ($i \leq N$). This models an essentially infinite number of potential offspring which are then culled down to the constant population size due to a limited carrying capacity of the environment (c.f. Hartl and Clark [66] where a special case is discussed in the context of population genetics). This particle model may be obtained by conditioning a branching particle system (use a Poisson (1) offspring law) on having constant population size. The same is true of the limiting martingale problems. Write \Rightarrow to denote weak convergence of probability laws.

THEOREM 6.5. (a) (Etheridge and March [48]). If $m(E) = 1$, \mathbb{P}_m is as in Theorem 1.1 (with branching rate γ) and $\bar{\mathbb{P}}_m$ is as in Theorem 6.4 with constant sampling rate γ , then

$$\mathbb{P}_m(X \in \cdot | |X_t(E) - 1| < \varepsilon \ \forall t \leq \varepsilon^{-1}) \xrightarrow{\varepsilon \rightarrow 0} \bar{\mathbb{P}}_m(\cdot).$$

(b) (Perkins [103]). More generally a \mathbb{P}_m -regular conditional distribution of $\bar{X}_\cdot = X_\cdot / X_\cdot(E)$ (set $\bar{X}_t = X_{t_X -}$ if $t > t_X = \inf\{t : X_t(E) = 0\}$ (c.f. Tribe [120])) given $X_\cdot(E) = f(\cdot)$ is the law of a Fleming-Viot process with sampling rate $\gamma f(\cdot)^{-1}$.

THEOREM 6.6. Let $m(E) = 1$, and \mathbb{P}_m^c be the law of the solution to (MP_c) . Then $\mathbb{P}_m^c \xrightarrow{c \rightarrow \infty} \bar{\mathbb{P}}_m^\gamma$ (γ constant).

Proof:

Use Theorem 6.2 and Theorem 6.5(b) to see that

$$\mathbb{P}_m^c(A) = \int \int 1(x, y_\cdot \in A) R_c(x, y_\cdot) \bar{\mathbb{P}}_m^{\gamma x^{-1}}(dy) P_1(dx), \quad A \in \mathcal{F}_t^X,$$

where P_1 is the law of Feller’s continuous state branching process with branching rate γ and

$$R_c(X) = \exp\left\{-cX_t(1)\left(\frac{X_t(1)}{2} - 1\right) + \frac{c\gamma}{2} \int_0^t X_s(1)(1 - c(X_s(1) - 1)^2)ds\right\}.$$

We have used some stochastic calculus to get the above expression for R_c . Note that $R_c(X) = \tilde{R}_c(X, (1))$ and so

$$(6.8) \quad \mathbb{P}_m^c(A) = \int \int 1(x, y_\cdot \in A) \bar{\mathbb{P}}_m^{\gamma x^{-1}}(dy) \tilde{R}_c(x) P_1(dx).$$

Under $\tilde{P}_c(dx) = \tilde{R}_c(x)P_1(dx)$, x satisfies the stochastic differential equation

$$x_t = 1 + \int_0^t \sqrt{\gamma x_s} dB_s + c \int_0^t (1 - x_s) ds \quad (B \text{ a Brownian motion}).$$

By Ito's lemma for $p \in \mathbb{N}$,

$$(6.9) \quad \begin{aligned} (x_t - 1)^p &= \int_0^t p(x_s - 1)^{p-1} \sqrt{\gamma} \sqrt{x_s} dB_s \\ &\quad - cp \int_0^t (x_s - 1)^p ds + \frac{p(p-1)}{2} \gamma \int_0^t (x_s - 1)^{p-2} x_s ds. \end{aligned}$$

An easy induction allows us to calculate $\tilde{P}_c((x_t - 1)^p)$, and the use of the Doob's strong L^p inequalities leads to

$$(6.10) \quad \sup_c \tilde{P}_c(\sup_{s \leq t} |x_s - 1|^p) < \infty \quad \forall p \in \mathbb{N}.$$

Takes means in (6.9) with $p = 2$ to see that

$$(6.11) \quad \tilde{P}_c((x_t - 1)^2) = \frac{\gamma}{2c}(1 - e^{-2ct}).$$

Take $p = 4$ in (6.9) to get

$$\sup_{t \leq T} (x_t - 1)^4 \leq \sup_{t \leq T} \int_0^t 4(x_s - 1)^3 \sqrt{\gamma} \sqrt{x_s} dB_s + 6\gamma \int_0^T (x_s - 1)^2 x_s ds.$$

This, together with (6.10) and (6.11) and a well-known square function inequality for martingales, leads to

$$\lim_{c \rightarrow \infty} \tilde{P}_c(\sup_{t \leq T} (x_t - 1)^4 > \varepsilon) = 0 \quad \text{for any } \varepsilon > 0.$$

Now use this in (6.8) together with the weak continuity in x of $\bar{\mathbb{P}}_m^{\gamma x^{-1}}$ at $x \equiv 1$ (c.f. Theorem 2(b) of Perkins [103]) to complete the proof. ■

6.3. Interactive Branching.

6.3.1. One-Dimensional Systems. In the one dimensional setting, consider the equation

$$(6.12) \quad \frac{\partial u}{\partial t} = \frac{1}{2} \Delta u + u^\gamma \dot{W}, \quad u \geq 0.$$

It is at times helpful to view (spde1) as a branching system with a density dependent branching rate given by $u(t, x)^{\gamma - \frac{1}{2}}$. Assume $u(0, \cdot)$ is a given continuous function on compact support which is not identically zero. If $\gamma \geq 1$, solutions to (spde1) are pathwise unique up to a possible explosion time σ at which time $u(t, \cdot)$ becomes unbounded on some compact interval ($u \rightarrow u^\gamma$ is locally Lipschitz) so one may easily adapt the arguments in Walsh [122]. If $0 < \gamma < 1$, there are globally defined solutions to (spde1) for some white noise \dot{W} (c.f. Shiga [115]). If $\gamma = 1/2$, u must be the density of super-Brownian motion. Uniqueness in law for $1/2 < \gamma < 1$ has recently been proved by Mytnik [96] by means of a duality argument which is a nontrivial generalization of that used for $\gamma = 1/2$. If $\gamma = \frac{1}{2}$, $u(t, \cdot)$ has compact support for all $t \geq 0$ a.s. (c.f. Remark 2.4). Here are some qualitative properties of solutions for other values of γ .

THEOREM 6.7. (a) (Mueller [89], and Mueller and Sowers [92]) If $1 \leq \gamma < \frac{3}{2}$, the solution to (spde1) exists for all $t \geq 0$. For sufficiently large γ , $\mathbb{P}(\sigma < \infty) > 0$.

(b) (Mueller [89], Mueller and Perkins [91], Shiga [115]). If $0 < \gamma < 1$, there is a solution to (spde1) with compact support for all $t > 0$ a.s. If $\gamma \geq 1$, then for each $t \geq 0$, with probability 1 either $t \geq \sigma$ or $u(t, x) > 0$ for all $x \in \mathbb{R}$.

Durrett (unpublished) and Krylov [75] have shown the conclusion in (b) for $0 < \gamma < 1$ is valid for any solution of (spde1). The largest value of γ for which global solutions exist in (a) is unknown. The compact support property in (b) suggests a delicate balance between critical branching which tends to kill off the population in underpopulated regions and diffusion which propagates instantaneously to the entire space. For $\gamma \geq 1$, the branching rate $u^{\gamma - \frac{1}{2}}$ becomes small quickly enough so that the diffusion term “wins” and the support propagates instantaneously.

6.3.2. Catalytic Branching. A different type of interactive branching is one in which the branching only occurs in the presence of a catalyst and in which the branching rate at x at time t is proportional to the amount of catalyst present, $\Gamma(t, dx)$, where $\Gamma(t)$ is a given continuous $M_p(\mathbb{R}^d)$ -valued process. The resulting catalytic super-Brownian motion is a measure-valued Markov process with Laplace functional $\mathbb{P}_{s,m}(\exp(-X_t(\varphi)) = \exp(-m(V_{s,t}\varphi)))$, where now for $\varphi \geq 0$, $u_{s,t} := V_{s,t}\varphi$ is the unique solution of

$$(6.13) \quad u_{s,t} = P_{t-s}\varphi - \int_s^t \int p_{t-r}(\cdot - y)(u_{r,t}(y))^2 \Gamma(r, dy) dr.$$

The function space equation would be expected to have the form

$$u_{s,t}(x) = P_{s,x}[\varphi(B_t) - \int_s^t u_{r,t}(B_r)^2 A_\Gamma(dr)],$$

where A_Γ would denote the collision local time between a Brownian particle and the medium Γ , namely,

$$A_\Gamma(t) = \left\langle \int_0^t \Gamma(r, dy) \delta_y(B_r) \right\rangle = \lim_{\varepsilon \downarrow 0} \int_0^t \int p_\varepsilon(B_r - y) \Gamma(r, dy) dr.$$

This limit exists and defines a continuous additive functional A_Γ under the regularity condition:

$$(6.14) \quad \sup_x \sup_{0 \leq s \leq N} \int_s^{s+\varepsilon} p_{r-s}(x - y) \phi_p(y) \Gamma(r, dy) \xrightarrow{\varepsilon \downarrow 0} 0 \quad \forall N > 0,$$

where $\phi_p(y) := (1 + |y|^2)^{-\frac{p}{2}}$. Under this condition the corresponding catalytic super-Brownian motion with values in $M_p(\mathbb{R}^d)$ also exists (c.f. Propositions 11 and 37 of Dawson and Fleischmann [24]). (Note that if $d = 1$ the condition (6.14) is always satisfied.) In this case \mathbb{P}_m^Γ denotes the law of this process, starting at m at $t = 0$.

The structure of the random measure X_t^Γ depends in an essential way on the local properties of the catalyst and is related to properties of the solutions of the log-Laplace equation. For example, in view of Theorem 6.7 it is natural to ask for which catalytic media in \mathbb{R} does the catalytic super-Brownian motion have the compact support property exhibited by ordinary super-Brownian motion. In fact, this compact support property is tied to a property of the solution of a related

nonlinear boundary value problem. For simplicity we restrict discussion to the case of a quenched medium $\Gamma(r, dx) = \Gamma(dx)$.

THEOREM 6.8. (*Dawson, Li, Mueller [32]*) Assume that Γ is a locally finite measure $\text{supp}(m) \subset [a_1, a_2]$.

(a) There exist sequences $\beta_{1,n}, \beta_{2,n} \uparrow \infty$ such that for each n the boundary value problem

$$(6.15) \quad \begin{aligned} \frac{1}{2}u_{xx} &= \Gamma u^2(x), \quad x \in (a_1, a_2) \\ u(a_1) &= \beta_{1,n}, \quad u(a_2) = \beta_{2,n}, \end{aligned}$$

(the first equality holds in the sense of equality of measures) has a unique solution $u_n(x)$.

(b) Let $S(X_t)$ be the closed support of X_t . Then

$$\mathbb{P}_m^\Gamma\{S(X_t) \cap [a_1, a_2]^c = \emptyset \quad \forall t \geq 0\}$$

$$= \lim_{n \rightarrow \infty} \exp\left[\int_{a_1}^{a_2} u_n(x)m(dx)\right].$$

The proof is too long to include here but we will sketch the main idea behind the proof of (b) which ties together the behavior of the range of the catalytic super-Brownian motion and the boundary value problem (6.15). It can be shown that

$$\mathbb{P}_{X(0)}^\Gamma \left(\exp\left(-\theta \int_0^t X_s([a_1, a_2]^c) ds\right) \right) = \exp\left(-\int v(0, t; \theta, x) X_0(dx)\right),$$

where $v(\cdot, t; \theta, x)$ satisfies the weak form of the backward equation

$$(6.16) \quad -\frac{\partial}{\partial s} v(s, t; \theta, x) = \frac{1}{2}\Delta v(s, t; \theta, x) - \Gamma v^2(s, t; \theta, x) + \theta 1_{[a_1, a_2]^c}(x)$$

for a.e. x and $s < t$,

with final condition

$$v(t, t; \theta, x) \equiv 0.$$

Then it can be verified that $v(0, t; \theta, x) \uparrow u(\theta, x)$ as $t \uparrow \infty$, where u satisfies the weak form of

$$\frac{1}{2}\Delta u = \Gamma u^2 + \theta 1_{[a_1, a_2]^c}.$$

The result is obtained by taking the limit as $\theta \rightarrow \infty$.

EXAMPLE 6.2. In one dimension we consider a “Lévy catalytic medium” in \mathbb{R}^1 given by

(6.17)

$$-\log E\{\exp(-\int f(x)\Gamma_1(dx))\} = \int \int_0^\infty (1 - e^{-\lambda f(x)})\nu(d\lambda)dx, \quad f \in p\mathcal{B}(\mathbb{R}).$$

The α -stable random measure corresponds to the special case $\nu(d\lambda) = \frac{d\lambda}{\lambda^{1+\alpha}}$, $0 < \alpha < 1$.

COROLLARY 6.1. (*Dawson, Li and Mueller [32]*) Let $(X_t^\Gamma : t \geq 0)$ be a catalytic super-Brownian motion in a Lévy random catalyst Γ given by (6.13) and let $\bar{R}(\infty)$ denote the closure of $\cup_{t>0} S(X_t)$.

(a) In the α -stable case, $\nu(d\lambda) = \frac{d\lambda}{\lambda^{1+\alpha}} 1(\lambda > 0)$, $\alpha \in (0, 1)$, then for a.e. realization of Γ the compact support property is satisfied, that is, $P_m^\Gamma(\bar{R}(\infty) \text{ is compact}) = 1$ for every initial measure, m , having compact support.

(b) If Γ has the rarefied Lévy measure $\nu(d\lambda) = \frac{1}{\lambda} 1(0 < \lambda \leq 1)$, then for a.e. realization of Γ , $P_m^\Gamma(\bar{R}(\infty) \text{ is compact}) = 0$ for every initial measure m .

A second natural question concerns the existence of a density process.

LEMMA 6.1. (*Theorems 2.5.2 and 2.6.2 of Dawson and Fleischmann [24]*) Absolute continuity of X_t^Γ is implied by the existence of a “fundamental solution” $V_t(\lambda\delta_x)$ of equation (6.16) (with $\theta \equiv 0$) such that

$$\begin{aligned} V_t &\rightarrow \lambda\delta_x \text{ as } t \downarrow 0 \\ \frac{\partial}{\partial \lambda} V_t(\lambda\delta_x)|_{\theta=0} &= p(t, x) \end{aligned}$$

for Lebesgue a.e. x .

Proof:

The idea of the proof is to consider the sequence of approximate delta functions $p_\frac{n}{n}(\cdot - x)$ and to verify that the corresponding solutions to the log-Laplace equation form a Cauchy sequence in an appropriate Banach space.

To illustrate this we will give a formal calculation of the second moment of the density (if it exists) at a point z when Γ is a finite singular measure that does not charge polar sets and for simplicity we also assume that it has bounded support and that the initial measure μ has a bounded density. Then

$$Var_{s,\mu}(x_T(z)) = 2 \int \mu(da) P_{s,a} \int_s^T A_\Gamma(dr) p^2(T-r, z - W_r).$$

The right side of the above expression behaves like $\int \frac{1}{|y-z|^{2d-\alpha}} \Gamma(dy)$. If this quantity is finite for Lebesgue almost every z , then we can verify that a density exists. ■

EXAMPLE 6.3. Consider the measure Γ_d^α which is supported by a dense family of hyperplanes, namely, $\Gamma_d^\alpha = \Gamma_1^\alpha \times \lambda_{d-1}$ in \mathbb{R}^d ($d > 1$), where Γ_1^α is the α -stable random measure and λ_{d-1} is $(d-1)$ -dimensional Lebesgue measure. There exists a density process if and only if $\alpha < \frac{1}{2d-1}$ (c.f. Dawson and Fleischmann [23]).

6.3.3. *Mutually Catalytic Branching.* Consider a two species model in which each species catalyzes the reproduction of the other species. Formally, this leads to the coupled pair of stochastic partial differential equations

$$\begin{aligned} \frac{\partial u(t, x)}{\partial t} &= \frac{1}{2} \Delta u(t, x) + \sqrt{u(t, x)v(t, x)} \dot{W}_1(t, x) \\ \frac{\partial v(t, x)}{\partial t} &= \frac{1}{2} \Delta v(t, x) + \sqrt{u(t, x)v(t, x)} \dot{W}_2(t, x), \end{aligned}$$

where \dot{W}_1, \dot{W}_2 are independent white noises. It is not hard to prove the existence of such solutions ($u \geq 0, v \geq 0$) for some \dot{W}_1, \dot{W}_2 if $d = 1$. Although one can use

collision local time to formulate a martingale problem for this model when $d > 1$, it is not at all clear if any solutions exist. It is therefore natural to study the analogous model on \mathbb{Z}^d , where existence holds for any d . If $\{p_t(j) : j \in \mathbb{Z}^d\}$ are the transition probabilities for the continuous time simple symmetric random walk on \mathbb{Z}^d , the lattice model leads to the system of stochastic integral equations: for $j \in \mathbb{Z}^d$

$$\begin{aligned} x_j(t) &= \sum_k p_t(j-k)x_k(0) + \sum_k \int_0^t p_{t-s}(j-k)M_k^1(ds) \\ y_j(t) &= \sum_k p_t(j-k)y_k(0) + \sum_k \int_0^t p_{t-s}(j-k)M_k^2(ds) \\ M_k^i(ds) &= \sqrt{x_k(s)y_k(s)}dW_k^i(s), \quad i = 1, 2 \end{aligned}$$

where $\{W_k^1, W_k^2 : k \in \mathbb{Z}^d\}$ are independent Brownian motions. Uniqueness in law of solutions to the above and their continuous analogue ($d = 1$) has recently been established by Mytnik [94] through a duality argument.

Although several questions remain unresolved, some results on the long-time behavior of the above lattice system have been recently derived.

THEOREM 6.9. (*Dawson and Perkins [35]*). *If $x_j(0) = x > 0$ and $y_j(0) = y$ for all j then*

- (a) *For $d \geq 3$, $(x, y)(t) \Rightarrow (x_\infty, y_\infty)$ as $t \rightarrow \infty$ (vague convergence), where (x_∞, y_∞) is a stationary equilibrium satisfying $P(x_\infty(j)) = x$, $P(y_\infty(j)) = y$ and $x_\infty(j) \neq 0, y_\infty(j) \neq 0 \forall j \in \mathbb{Z}^d$ a.s.*
- (b) *For $d \leq 2$, $(x, y)(t) \Rightarrow (x_\infty, y_\infty)$ as $t \rightarrow \infty$ (vague convergence), where $(x_\infty, y_\infty) \stackrel{\mathcal{D}}{\equiv} (B_T^1, B_T^2)\lambda$, (B^1, B^2) is a planar Brownian motion starting at (x, y) , $T = \inf\{t : B_t^1 B_t^2 = 0\}$ and λ is counting measure on \mathbb{Z}^d . (We consider $x(t)$ and x_∞ as $M(\mathbb{Z}^d)$ -valued processes.)*

Note the above theorem states that asymptotic coexistence of types is possible if and only if $d \geq 3$.

6.3.4. Multilevel Branching and Interaction. In this section we consider a hierarchically structured population consisting of objects at two levels. Objects at the first level are particles and a level 2 object is a collection of related particles. As an example consider a structured file system. Files are level one objects and directories are level 2 objects. We assign mass η_1 and branching rate $\frac{1}{\eta_1}$ to each level one object and have them evolve as independent time homogeneous A -motions in a Polish space E between branch times, and mass η_2 and branching rate $\frac{1}{\eta_2}$ to each level 2 object. We then consider the two level branching particle system

$$X_t^{(\eta_1, \eta_2)} := \eta_2 \sum_i \delta_{\eta_1 \sum_j \delta_{x_{i,j}(t)}}$$

and we view this as a $M_F^2(E) := M_F(M_F(E))$ -valued process.

As $\eta_1, \eta_2 \rightarrow 0$, $X^{\eta_1, \eta_2}(t)$ converges to a $M_F^2(E)$ -valued Markov process with Laplace transition functional

$$\begin{aligned} L_{t,\nu}(J) &= E \left\{ \exp(- \int J(\mu) X(t, d\mu)) | X(0) = \nu \right\} \\ &= \exp(- \int u(t, \mu) \nu(d\mu)), \end{aligned}$$

where $u(t, \mu)$ satisfies the integral equation

$$\begin{aligned} u(t, \mu) &= T_t u(0, \mu) - \frac{1}{2} \int_0^t [T_{t-s} u^2(s, \cdot)](\mu) ds \\ u(0, \mu) &= J(\mu) = f(\mu(\phi)), \quad \mu \in M_F(\mathbb{R}^d), \end{aligned}$$

where $\phi \in p\mathcal{B}(\mathbb{R}^d)$, $f \in p\mathcal{B}(\mathbb{R})$ and $\{T_t : t \geq 0\}$ is the semigroup of the A -superprocess.

In the case $E = \mathbb{R}^d$, $A = \frac{1}{2}\Delta$, the limiting process is in fact “super-super Brownian motion”. The aggregated process is defined to by

$$Z(t, dx) = \int \mu(dx) X(t, d\mu).$$

This is a $M_F(\mathbb{R}^d)$ -valued non-Markov process. Some small scale properties of this two level super-Brownian motion aggregate are described in Dawson, Hochberg, and Vinogradov [30]. In particular, the support dimension is 4 (compared to 2 for one level super-Brownian motion) in dimensions $d \geq 4$. The long time behavior has been studied when the initial state is $\lambda(dx)\delta_{x_0}$. With this initial condition it has been shown in Wu [125] and Gorostiza, Hochberg and Wakolbinger [61] that persistence holds if and only if $d > 4$ which corresponds to the dimensions in which there is clan transience.

We now proceed to give an example of inter-level interaction. Consider the limit of the two level system and in which files can be exported to all existing directories at any time. This leads to an additional term in the generator as follows:

$$G_2^c F(\nu) = G_2 F(\nu) + \int \int \int c(\mu, \mu', y) \frac{\delta}{\delta \mu(y) \delta \nu(\mu)} F(\nu) \mu'(dy) \nu(d\mu') \nu(d\mu),$$

where G_2 is the generator of the two-level system without file export. Here $c(\mu, \mu', y)$ is the rate at which a file of type y in a directory μ' is copied into a directory μ . Existence of solutions to this martingale problem can be obtained by a compactness argument. However, it is an interesting open problem to develop methods to establish that multilevel martingale problems such as this having an inter-level interaction are well-posed. It is shown in Dawson and Wu [36] that in the supercritical case this system can exhibit super-exponential growth and that there exists a time t_0 such that for all $t \geq t_0$,

$$E[\int \mu(E) X_t(\mu)] = \infty.$$

6.4. Fleming-Viot Processes with State Dependent Sampling. The analogue of interactive branching in the case of the Fleming-Viot process involves replacing Q_μ by

$$Q_\mu^G(du, dv) = \left(\int G(\mu, u, z) \mu(dz) \right) \mu(du) \delta_u(dv) - G(\mu, u, v) \mu(du) \mu(dv),$$

where

$$G(\mu, u, z) = G_{c_0+c_1}(\mu, u, z) = c_0 + c_1(\mu, u, z).$$

The idea is to consider a class of processes which can be viewed as perturbations of Fleming-Viot processes involving constant coefficients. Dawson and March [34] have proved that the corresponding martingale problem is well-posed for a class of coefficients G using a “parametrix” method. This involves representing the resolvent of the interactive process as a perturbation of the simple Fleming-Viot process as follows:

$$\begin{aligned} R_\lambda^{c_0+c_1} &= R_\lambda^{c_0}(I - K_\lambda)^{-1} \\ K_\lambda &= (G_{c_0+c_1} - G_{c_0})R_\lambda^{c_0}. \end{aligned}$$

In order to make this precise one introduces a normed space of functions of the form

$$\begin{aligned} F(\mu) &= \sum_{n=0}^{\infty} \int \cdots \int f_n(x_1, \dots, x_n) \mu(dx_1) \dots \mu(dx_n), \quad \text{with } f_n \in C(E^n) \\ \|F\|_b &:= \sum_{n=0}^{\infty} b^n \|f_n\|_{\infty} \quad \text{with } b > 1. \end{aligned}$$

Then the main step is to show that

$$(6.18) \quad \|K_\lambda\|_b < 1 \text{ for all large } \lambda$$

using the dual process representation of $R_\lambda^{c_0}$. This implies that the Neumann series for $(I - K_\lambda)^{-1}$ converges thus identifying $R_\lambda^{c_0+c_1}$ for all large λ which in turn determines the one dimensional marginal distributions of the solution of the perturbed martingale problem. In order to verify (6.18) it is necessary to assume that c_1 is sufficiently small in a suitable norm or that c_0 is sufficiently large. The stopping time argument in Stroock-Varadhan’s original arguments for d-dimensional diffusions no longer seems to give the result in general.

6.5. Multi-type Systems with Local Interaction. In studying the large space-time scale behavior of interactive systems we will turn to spatial discretizations to avoid the non-existence phenomena described above which arises in “zero-range” interactive systems. In particular we are concerned with multi-type systems in which there is local competition leading to constraints on the total population. This leads to systems of stochastic differential equations indexed by \mathbb{Z}^d .

To begin consider the two-type *stepping stone model* which is obtained when, at each lattice site, the death rate is proportional to the the total population. This comes from the system: $i \in S$, where S is a countable Abelian group

$$\begin{aligned} dx_i(t) &= \sum_{j \in S} q_{i-j}[x_j(t) - x_i(t)]dt + \sqrt{x_i(t)}dB_{1,i}(t) + c[1 - x_i(t) - y_i(t)]x_i(t)dt \\ dy_i(t) &= \sum_{j \in S} q_{i-j}[y_j(t) - y_i(t)]dt + \sqrt{y_i(t)}dB_{2,i}(t) + c[1 - x_i(t) - y_i(t)]y_i(t)dt. \end{aligned}$$

Let $r_i(t) = x_i(t) + y_i(t)$, $z_i(t) = \frac{x_i(t)}{x_i(t) + y_i(t)}$

$$\begin{aligned} dr_i(t) &= \sum_{j \in S} q_{i-j}[r_j(t) - r_i(t)]dt + c(r_i(t) - r_i^2(t))dt + \sqrt{r_i(t)}d\tilde{B}_{1,i}(t), \text{ and} \\ dz_i(t) &= \frac{1}{r_i(t)} \sum_{j \in S} q_{i-j}r_j(t)[z_j(t) - z_i(t)]dt + \sqrt{\frac{z_i(t)(1-z_i(t))}{r_i(t)}}d\tilde{B}_{2,i}(t). \end{aligned}$$

As in Theorem 6.6 in the limit $c \rightarrow \infty$ this gives rise to the Fisher-Wright stepping stone system

$$dz_i(t) = \sum_{j \in S} q_{i-j}[z_j(t) - z_i(t)]dt + \sqrt{z_i(t)(1-z_i(t))}dB_i(t).$$

This can be extended to the infinitely many types case, that is, a system $\{\mu_i(t)\}_{i \in S}$ of $M_1([0, 1])$ -valued processes in which state-dependent Fleming-Viot sampling takes place at each site. This corresponds to the martingale problem generator

$$\begin{aligned} \mathbb{G}F((\mu_i)_{i \in S}) &= \sum_{i,j \in S} q_{i,j} \int_{[0,1]} \frac{\delta F}{\delta \mu_i}(u)(\mu_j(du) - \mu_i(du)) \\ &\quad + \frac{\gamma}{2} \sum_{i \in S} \int_{[0,1]} \int_{[0,1]} \frac{\delta^2 F}{\delta \mu_i \delta \mu_i}(u, v) Q_{\mu_i}^G(du, dv), \end{aligned}$$

where $Q_{\mu_i}^G(du, dv) = \int G(\mu_i, u, z) \mu_i(dz) \mu_i(du) \delta_u(dv) - G(\mu_i, u, v) \mu_i(du) \mu_i(dv)$ and G are as in Section 6.4.

THEOREM 6.10. (*Dawson, Greven and Vaillancourt [28]*) Consider the infinitely many types stepping stone model (with $c \equiv 0$) and assume the initial random field is spatially stationary and ergodic and has single site mean measure $\mu \in M_1([0, 1])$. If $\{q_{i,j}\}$ is a symmetric transient random walk on S , then the stepping stone process X converges in distribution as $t \rightarrow \infty$ to a non-trivial ergodic invariant measure ν_μ and ν_μ has single site mean measure μ . If $\{q_{i,j}\}$ is recurrent, then the extreme invariant measures are δ_a , $a \in [0, 1]$, that is, there is local fixation.

The proof of this is based on a dual process representation which involves a system of elements performing coalescing random walks. For example if $\{q_{i,j}\}$ are the transition rates of a recurrent symmetric random walk on S , then it is clear that all elements eventually coalesce with probability one. This can then be used to show that any finite number of sites will in the limit as $t \rightarrow \infty$ be all occupied by a single type. On the other hand if the initial configuration is identically equal to μ in the transient case, then each of the partition elements which ultimately survive will sample independently from the distribution μ yielding the required moment measures.

6.5.1. Universality Classes. Many detailed results on the large space-time scale behavior have been obtained for super-Brownian motion and stepping stone Fleming-Viot models using the infinite divisibility, respectively, the dual process. However there is reason to believe that the structures which arise in these models are shared by a much wider class of systems. It is an open problem to establish this but some partial results in this direction have been obtained using the thermodynamic limit approach and hierarchical mean field introduced in Section 5.6.

In particular in the setting of the hierarchical mean-field limit Dawson and Greven [35] have shown that there is a close relationship between the large space-time scale behavior of the infinitely many types stepping stone model and the critical branching model.

Using the class of interactive stepping stone models with state dependent sampling described above Dawson, Greven and Vaillancourt [28] have obtained universality results in the hierarchical mean-field limit for large space-time behavior (including the persistence-extinction dichotomy and cluster formation) in the infinitely many types stepping stone model.

Bibliography

- [1] D. Aldous (1993). Tree-based models for random distribution of mass, *J. Stat. Phys.* 73, 625-641.
- [2] J.B. Baillon, Ph. Clément, A. Greven and F. den Hollander (1996). On the attracting orbit of a nonlinear transformation arising from normalization of hierarchically interacting diffusion. Part 2: The non-compact case, preprint.
- [3] D. Bakry and M. Emery (1985). Diffusions hypercontractives, in *Séminaire de Probabilités XIX*, Lecture Notes in Math. 1123, Springer-Verlag, 177-206.
- [4] P. Baras and M. Pierre. (1984). Problèmes paraboliques semi-linéaires avec données mesures. *Appl. Anal.* 18, 111-149.
- [5] M.T. Barlow, S.N. Evans, and E.A. Perkins. (1991). Collision local times and measure-valued processes. *Can. J. Math.* 43, 897-938.
- [6] M.T. Barlow, S.N. Evans, and E.A. Perkins (1997). Measure-valued stochastic processes and competing species. Preprint.
- [7] M.T. Barlow and E.A. Perkins (1994). On the filtration of historical Brownian motion. *Ann. Probab.* 22, 1273-1294.
- [8] D. Blount (1996). Diffusion limits for a nonlinear density dependent space-time population model, *Ann. Probab.* 24, 639-659.
- [9] M. Bramson, J.T. Cox and A. Greven (1993). Ergodicity of critical spatial branching processes in low dimensions, *Ann. Probab.* 21, 1946-1957.
- [10] M. Bramson, J.T. Cox and A. Greven (1997). Invariant measures of critical spatial branching processes in high dimensions, *Ann. Probab.* 25, 56-70.
- [11] M. Bramson, R. Durrett and G. Swindle (1989). Statistical mechanics of crab-grass, *Ann Prob.* 17, 444-481.
- [12] J.T. Cox and A. Greven (1994). Ergodic theorems for systems of locally interacting diffusions, *Ann. Probab.* 22, 833-853.
- [13] J.T. Cox, A. Greven, T. Shiga (1995). Finite and infinite systems of interacting diffusions, *Prob. Th. Rel. Fields* 102, 165-197.
- [14] J.T. Cox, K. Fleischmann and A. Greven (1996). Comparison of interacting diffusions and an application to their ergodic theory, *Prob. Th. Rel. Fields* 105, 513-528.
- [15] J.T. Cox and D. Griffeath (1986) Diffusive clustering in the two dimensional voter model, *Ann. Probab.* 14, 347-370.
- [16] D.A. Dawson (1975). Stochastic evolution equations and related measure processes. *J. Multivariate Anal.* 3, 1-52.
- [17] D.A. Dawson (1977). The critical measure diffusion, *Z. Wahr. verw Geb.* 40, 125-145.
- [18] D.A. Dawson (1978). Geostochastic calculus. *Can. J. Statistics* 6, 143-168.
- [19] D.A. Dawson (1992). Infinitely divisible random measures and superprocesses, in Proc. 1990 Workshop on Stochastic Analysis and Related Topics, Silivri, Turkey.
- [20] D.A. Dawson (1993). Measure-valued Markov processes, École d'Été de Probabilités de Saint Flour, 1991, Lect. Notes in Math. 1541, Springer, Berlin.
- [21] D.A. Dawson and K. Fleischmann (1988). Strong clumping of critical space-time branching models in subcritical dimensions, *Stoch. Proc. Appl.* 30, 193-208.
- [22] D.A. Dawson and K. Fleischmann (1994). A super-Brownian motion with a single point catalyst, *Stoch. Proc. Appl.* 49, 3-40.
- [23] D.A. Dawson and K. Fleischmann (1995). Super-Brownian motion in higher dimensions with absolutely continuous measure states, *J. Theor. Probab.* 8 (1995), 179-206.

- [24] D.A. Dawson and K. Fleischmann (1997) A continuous super-Brownian motion in a super-Brownian medium, *J. Theor. Probab.* 10.
- [25] D.A. Dawson, K. Fleischmann, Y. Li and C. Mueller (1995). Singularity of super-Brownian local time at a point catalyst, *Ann. Probab.* 23, 37-55.
- [26] D.A. Dawson and A. Greven (1993). Hierarchical models of interacting diffusions: multiple time scale phenomena. Phase transition and pattern of cluster-formation. *Prob. Th. Rel. Fields* 96, 435-473.
- [27] D.A. Dawson and A. Greven (1996). Multiple space-time scale analysis for interacting branching models, *Electronic Journal of Probability* 1.
- [28] D.A. Dawson, A. Greven and J. Vaillancourt (1995). Equilibria and quasi-equilibria for infinite systems of Fleming-Viot processes, *Trans. Amer. Math. Soc.* 347, 2277-2360.
- [29] D.A. Dawson and K.J. Hochberg (1979). The carrying dimension of a stochastic measure diffusion, *Ann. Probab.* 7, 693-703.
- [30] D.A. Dawson, K.J. Hochberg and V. Vinogradov (1996). High density limits of hierarchically structured branching-diffusing populations, *Stoch. Proc. Appl.* 62, 191-222.
- [31] D.A. Dawson, I. Iscoe and E.A. Perkins (1989). Super-Brownian motion: path properties and hitting probabilities. *Prob. Th. Rel. Fields* 83, 135-205.
- [32] D.A. Dawson, Y. Li, C. Mueller (1996). The support of measure-valued branching processes in a random environment, *Ann. Probab.* 23, 1692-1718.
- [33] D.A. Dawson and P. March (1995). Resolvent estimates for Fleming-Viot operators and uniqueness of solutions to related martingale problems, *J. Funct. Anal.* 132, 417-472.
- [34] D.A. Dawson and E.A. Perkins (1991). Historical processes. *Mem. Amer. Math. Soc.* 454.
- [35] D.A. Dawson and E.A. Perkins (1997). Long-time behavior and co-existence in a mutually catalytic branching model, preprint.
- [36] D.A. Dawson and Y. Wu (1996). Multilevel multi-type models of an information system, I.M.A. Volume 84, Classical and Modern Branching Processes, eds. K.B. Athreya and P. Jagers, 57-72, Springer-Verlag.
- [37] J.D. Deuschel and K. Wang (1994) Large deviations for occupation time functionals of branching particles and super-Brownian motion, preprint.
- [38] E. Derbez and G. Slade (1996). Lattice trees and super-Brownian motion, preprint.
- [39] P.E. Donnelly and T.G. Kurtz (1996a). A countable representation of the Fleming-Viot measure-valued diffusion, *Ann. Probab.* 24 698-742.
- [40] P.E. Donnelly and T.G. Kurtz (1996b). Particle representations for measure-valued population models, preprint.
- [41] E.B. Dynkin (1991). Branching particle systems and superprocesses, *Ann. Probab.* 19:1157-1194.
- [42] E.B. Dynkin (1993). Superprocesses and partial differential equations. *Ann. Probab.* 21, 1185-1262.
- [43] E.B. Dynkin (1994) An introduction to Branching Measure-Valued Processes, CRM Monograph Series, Vol. 6, Amer. Math. Soc., Providence.
- [44] E.B. Dynkin, S.E. Kuznetsov and A.V. Skorohod (1994). Branching measure-valued processes, *Probab. Th. Rel. Fields* 99, 55-96.
- [45] E.B. Dynkin and S.E. Kuznetsov (1996). Superdiffusions and removable singularities for quasilinear partial differential equations, *Comm. Pure and Appl. Math.* 49, 125-176.
- [46] M. Eigen (1993). Viral quasi-species, *Scientific American*, July 42-49.
- [47] N. El Karoui and S. Roelly (1991). Propriétés de martingales, explosion et représentation de Lévy-Kinchine d'une classes de processus de branchement à valeurs mesures, *Stoch. Proc. Appl.* 38, 239-266.
- [48] A. Etheridge and P. March (1991). A note on superprocesses, *Prob. Th. Rel. Fields* 89, 141-147.
- [49] S.N. Ethier and T.G. Kurtz (1986). *Markov Processes: Characterization and Convergence*, Wiley, New York.
- [50] S.N. Ethier and T.G. Kurtz (1993) Measure-valued diffusion processes in population genetics, *SIAM J. Control Optim.* 31, 345-386.
- [51] S.N. Evans (1993). Two representations of a conditioned superprocess, *Proc. Roy. Soc. Edinburgh*, 123A, 959-971.
- [52] S.N. Evans and E.A. Perkins (1991). Absolute continuity results for superprocesses, *Trans. Amer. Math. Soc.* 325, 661-681.

- [53] S.N. Evans and E.A. Perkins (1994) Measure-valued branching diffusions with singular interactions, *Can. J. Math.* 46, 120-168.
- [54] W. Feller (1939). Die grundlagen der Volterrassen theorie des kampfes ums dasein im wahrscheinlichkeitstheoretischer behandlung. *Acta Biotheor.* 5, 11–40.
- [55] P.J. Fitzsimmons (1988). Construction and regularity of measure-valued Markov processes, *Israel J. Math.* 64, 337-361.
- [56] P.J. Fitzsimmons (1992). On the martingale problem for measure-valued Markov branching processes. In *Seminar on Stochastic Processes 1991*, Birkhauser, Boston.
- [57] K. Fleischmann and I. Kaj (1994). Large deviation probabilities for some rescaled superprocesses, *Ann. Inst. Henri Poincaré* 30, 607-645.
- [58] K. Fleischmann, J. Gärtner and I. Kaj (1996). A Schilder type theorem for super-Brownian motion, *Can. J. Math.* 48, 542-568.
- [59] K. Fleischmann and J.-F. Le Gall (1995). A new approach to the single point catalytic super-Brownian motion, *Prob. Th. Rel. Fields* 102, 63-82.
- [60] K. Fleischmann and J. Gärtner (1986). Occupation time process at a critical point, *Math. Nachr.* 125, 275-290.
- [61] L.G. Gorostiza, K.J. Hochberg and A. Wakolbinger (1995). Persistence of a critical super-2 process, *J. Appl. Prob.* 32, 534-540.
- [62] L.G. Gorostiza and J.A. Lopez-Mimbela (1994). An occupation time approach for convergence of measure-valued processes, and the death process of a branching system, *Stat. and Prob. Letters* 21, 59-67.
- [63] G. Grimmett (1989). “Percolation”, Springer-Verlag, New York.
- [64] I. Gyöngy and E. Pardoux (1996). On quasi-linear stochastic partial differential equations, *Probab. Th. Rel. Fields* 94, 413-425.
- [65] T. Hara and G. Slade (1992). Self-avoiding walk in five or more dimensions. I. The critical behavior. *Commun. Math. Phys.* 147, 101-136.
- [66] D.L. Hartl and A.G. Clark (1989). Principles of Population Genetics, second ed., Sinauer Associates, Inc, Sunderland, Massachusetts.
- [67] I. Iscoe (1986a). A weighted occupation time for a class of measure-valued critical branching Brownian motion, *Probab. Theor. Rel. Fields* 71, 85-116.
- [68] I. Iscoe (1986b). Ergodic theory and a local occupation time for a class of measure-valued critical branching Brownian motion, *Stochastics* 18, 197-143.
- [69] I. Iscoe and T.Y. Lee (1993). Large deviations for occupation times of measure valued branching Brownian motions, *Stoch. and Stoch. Reports* 45, 177-209.
- [70] O. Kallenberg (1983.) Random Measures, 3rd ed., Akademie Verlag and Academic Press.
- [71] A. Klenke (1997). Multiple scale analysis of clusters in spatial branching models, *Ann. Probab.*, to appear.
- [72] N. Konno and T. Shiga(1988). Stochastic differential equations for some measure-valued diffusions, *Prob. Th. Rel. Fields* 79, 201–225.
- [73] P. Kotelenez (1995). A class of quasilinear stochastic partial differential equations of McKean-Vlasov type with mass conservation, *Probab. Th. Rel. Fields* 102, 159-1888.
- [74] S. Krone (1995). Conditioned superprocesses and their weighted occupation times, *Statistics and Probability Letters* 22, 59-69.
- [75] N. Krylov (1996). On a result of Mueller and Perkins, preprint.
- [76] N. Krylov and B. Rozovski (1981). Stochastic evolution equations, *J. Soviet Math. (Itogi Nauki i Techniki 14)*, 1233-1277.
- [77] G. Leduc (1995). Construction and characterization of superprocesses, Ph.D. thesis, Carleton University.
- [78] T.Y. Lee (1993). Some limit theorems for super-Brownian motion and semilinear differential equations, *Ann. Probab.* 21, 979-995.
- [79] J.-F. Le Gall (1991). Brownian excursions, trees and measure-valued branching processes, *Ann. Probab.* 19, 1399-1439.
- [80] J.-F. Le Gall (1993a). The uniform random tree in a Brownian excursion, *Prob. Th. Rel. Fields* 95, 25-46.
- [81] J.-F. Le Gall (1993b). A class of path-valued Markov processes and its applications superprocesses. *Prob. Th. Rel. Fields* 95, 25–46.
- [82] J.-F. Le Gall (1994). Mouvement brownien, processus de branchement et superprocessus, Prepublication 247 du Laboratoire de Probabilités.

- [83] J.-F. Le Gall (1995). The Brownian snake and solutions of $\Delta u = u^2$ in a domain, *Prob. Th. Rel. Fields* 102, 393-432.
- [84] J.-F. Le Gall (1996). Superprocesses, Brownian snakes and partial differential equations, Lecture Notes from 11th Winter School on Stochastic Processes, Siegmundsburg.
- [85] J.F. Le Gall and Y. Le Jan (1996) Branching processes in Lévy processes: The exploration process, preprint.
- [86] J.F. Le Gall and E.A. Perkins (1995). The Hausdorff measure of the support of two-dimensional super-Brownian motion, *Ann. Probab.* 23, 1719-1747.
- [87] Z. Li and T. Shiga (1995) Measure-valued branching diffusions: immigrations, excursions and limit theorems, *J. Math. Kyoto Univ.*
- [88] M. Lopez (1996). Path properties and convergence of interacting superprocesses, Ph.D. thesis, U.B.C.
- [89] C. Mueller (1991). Long time existence for the heat equation with a noise term, *Prob. Th. Rel. Fields* 90, 505-517.
- [90] C. Mueller (1991). On the support of solutions to the heat equation with noise, *Stochastics* 37, 225-246.
- [91] C. Mueller and E.A. Perkins (1992). The compact support property for solutions to the heat equation with noise, *Probab. Th. Rel. Fields* 93, 325-358.
- [92] C. Mueller and R. Sowers (1993). Blowup for the heat equation with a noise term, *Prob. Th. Rel. Fields* 97, 287-320.
- [93] C. Mueller and R. Tribe (1994). Stochastic p.d.e.'s arising from the long range contact and long range voter processes, *Prob. Th. Rel. Fields* 102, 519-546.
- [94] L. Mytnik (1996a). Superprocesses in random environments and related topics, Ph.D. thesis, Technion.
- [95] L. Mytnik (1996b). Superprocesses in random environments, *Ann. Probab.* 24, 1953-1978.
- [96] L. Mytnik (1997). Uniqueness result for an SPDE of C. Mueller and E. Perkins, preprint.
- [97] A. Pazy (1983). Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer-Verlag, Berlin.
- [98] Y. Peres (1996a). Intersection-equivalence of Brownian paths and certain branching processes, *Comm. Math. Phys.* 177, 417-434.
- [99] Y. Peres (1996b). The percolation approach to capacity, preprint.
- [100] E.A. Perkins (1988). A space-time property of a class of measure-valued branching diffusions, *Trans. Amer. Math. Soc.* 305, 743-795.
- [101] E.A. Perkins (1989). The Hausdorff measure of the closed support of super-Brownian motion, *Ann. Inst. Henri Poincaré* 25, 205-224.
- [102] E.A. Perkins (1990). Polar sets and multiple points for super-Brownian motion. *Ann. Probab.* 18, 453-491.
- [103] E.A. Perkins (1991). Conditional Dawson-Watanabe processes and Fleming-Viot processes. In Seminar in Stochastic Processes 1991, Birkhäuser, 142-155.
- [104] E.A. Perkins (1992) Measure-valued branching diffusions with spatial interactions, *Probab. Theory Relat. Fields* 94, 189-245.
- [105] E.A. Perkins (1995). On the martingale problem for interactive measure-valued branching diffusions. *Mem. Amer. Math. Soc.* 115, No. 549, 1-89.
- [106] M. Reimers (1989). One-dimensional stochastic partial differential equations and the branching measure diffusion. *Prob. Th. Rel. Fields* 81, 319-340.
- [107] S. Roelly-Coppoletta and A. Rouault (1989). Processus de Dawson-Watanabe conditioné par le futur lointain. *C.R. Acad. Sci. Paris* 309, Série I, 867-872.
- [108] L. Serlet (1995a). On the Hausdorff measure of multiple points and collision points of super-Brownian motion, *Stochastics*, to appear.
- [109] L. Serlet (1995b). Some dimension results for super-Brownian motion, *Prob. Th. Rel. Fields* 101, 371-391.
- [110] Y.-C. Sheu (1993). A Hausdorff measure classification of G-polar sets for the superdiffusions, *Prob. Th. Rel. Fields* 95, 521-534.
- [111] A. Schied (1996). Sample path large deviations for super-Brownian motion, *Prob. Th. Rel. Fields* 104, 319-347.
- [112] M. Sharpe (1988). General Theory of Markov Processes, Academic Press, Boston.
- [113] T. Shiga (1982). Continuous time multi-allelic stepping stone models in population genetics, *J. Math. Kyoto Univ.* 22, 1-40.

- [114] T. Shiga (1992). Ergodic theorems and exponential decay of sample paths for certain interacting diffusion systems, *Osaka J. Math.* 29, 789–807.
- [115] T. Shiga (1994). Two contrasting properties of solutions for one dimensional stochastic partial differential equations, *Can. J. Math.* 46, 415–437.
- [116] T. Shiga and A. Shimizu (1980). Infinite-dimensional stochastic differential equations and their applications, *J. Mat. Kyoto Univ.* 20, 395–416.
- [117] A. Stöckl and A. Wakolbinger (1994) On clan recurrence and transience in time stationary branching Brownian particle systems. In *Measure-valued Processes, Stochastic Partial Differential Equations and Interacting Systems*, ed. D. Dawson, CRM Lecture Notes and Monographs 5, 213–219.
- [118] S. Sugitani (1989). Some properties for the measure-valued branching diffusion processes. *J. Math. Soc. Japan* 41, 437–462.
- [119] S.J. Taylor and N.A. Watson (1985). A Hausdorff measure classification of polar sets for the heat equation. *Math. Proc. Cam. Phil. Soc.* 47, 325–344.
- [120] R. Tribe (1992) The behavior of superprocesses near extinction, *Ann. Prob.* 20, 286–311.
- [121] H. Wang (1995). Interacting branching particle systems and superprocesses, Ph.D. thesis, Carleton University.
- [122] J.B. Walsh (1986). An introduction to stochastic partial differential equations, *École d' Été de Probabilité de Saint Flour* 1984, Lect. Notes in Math. 1180, Springer, Berlin.
- [123] S. Watanabe (1968). A limit theorem of branching processes and continuous state branching, *J. Math. Kyoto U.* 8, 141–167.
- [124] Y. Wu (1993) A multilevel birth-death particle system and its continuous diffusion, *Adv. Appl. Prob.* 25, 549–569.
- [125] Y. Wu (1994) Asymptotic behavior of the two level measure branching process, *Ann. Probab.* 22, 854–874.
- [126] T. Yamada and S. Watanabe (1971). On the uniqueness of solutions of stochastic differential equations, *J. Math. Kyoto U.* 11, 155–167.

CHAPTER 3

Deterministic and Stochastic Hydrodynamic Equations Arising From Simple Microscopic Model Systems

GIAMBATTISTA GIACOMIN
Institut für Angewandte Mathematik
Universität Zürich–Irchel
Winterthurer Str. 190,
CH–8057 Zürich, Switzerland
gbg@amath.unizh.ch

JOEL L. LEBOWITZ
Department of Mathematics & Physics
Hill Center, Rutgers University
New Brunswick, N.J. 08903,
U.S.A.
lebowitz@math.rutgers.edu

ERRICO PRESUTTI
Dipartimento di Matematica
Università di Roma Tor Vergata
Via della Ricerca Scientifica
00133 Roma, Italy
presutti@mat.utovrm.it

1. Introduction

Nature has a hierarchical structure with strongly separated levels. This allows us, in many cases, to treat the higher levels (almost) independently of the lower ones. For example, the hydrodynamic laws which describe, with a high degree of accuracy, the large scale behavior of fluids were discovered before and are to a large extent independent of the detailed microscopic structure of matter.

These laws generally take the form of autonomous nonlinear partial differential equations,

$$(1.1) \quad \frac{\partial}{\partial t} \mathbf{M}(\mathbf{r}, t) = \mathbf{F}(\mathbf{M})(\mathbf{r}, t),$$

where $\mathbf{M}(\mathbf{r}, t)$ denotes an appropriate set of macroscopic variables depending on space and time and the structure of \mathbf{F} in (1.1) depends in general only on the phenomena considered and not on the nature of the microscopic constituents of the macroscopic objects studied. Thus, Fourier's law of heat conduction has the same form for solid gold and liquid water and the same Navier–Stokes equations describe

the flow of air and the flow of water: the details of microscopic structure generally enter in \mathbf{F} only through some parameters, e.g. heat conductivity, viscosity.

The origin of such generic hydrodynamic laws is a consequence of the existence of very large separation between the spatial and temporal scales of microscopic and macroscopic phenomena and some very general features of the microscopic dynamics. Chief among these are the approximate locality and additivity of the microscopic interactions and the resulting law of large numbers for the macroscopic fields. Consequently, our microscopic models can be rather crude, even blatantly wrong, and still give rise to correct macroscopic equations. All that is necessary is that the models contain the essential features responsible for the phenomena of interest.

The apparent robustness or universality of certain equations of type (1.1) is really quite remarkable, extending far beyond the description of *simple* macroscopic systems. Reaction–Diffusion equations of similar form can describe both chemically reacting molecular mixtures and the propagation of genetic traits in biological populations. The main aim of the present work is to discuss stochastic improvements to the deterministic description provided by (1.1).

The possibility and (occasional) need for improvement of this description comes from the fact that the different hierarchical levels are of course not completely isolated – there is no sharp demarcation line between an atomic beam and the jet stream. In fact, one of the basic dogmas of science is that the behavior at any level can be deduced, at least in principle, entirely from the dynamics of the level below it, i.e. there are no new physical laws, only new phenomena, as one goes from atoms to fluids to galaxies. The existence of the lower level will manifest itself in the apparently random fluctuations in the macroscopic fields about the solutions of (1.1): *the quantities fluctuate*, as one says (see [73], p.344). While these fluctuations are generally very small on the macroscopic scale in cases where the solutions of (1.1) are smooth and stable, they may become macroscopically relevant in cases where these solutions are unstable, singular, or non unique. There is thus a practical motivation for improving the deterministic equations in situations where the hydrodynamic laws fail, partially or entirely, to describe fully the phenomena of interest: the conceptual and mathematical motivations are of course not limited to such cases.

An *improved* version of (1.1) is often written in the form

$$(1.2) \quad \frac{\partial}{\partial t} \mathbf{M}(\mathbf{r}, t) = \mathbf{F}^\epsilon(\mathbf{M})(\mathbf{r}, t),$$

in which \mathbf{F}^ϵ is dependent on a small positive parameter ϵ , representing the contribution of the finer scale, and \mathbf{F}^ϵ approaches \mathbf{F} as ϵ goes to zero. A frequent choice for \mathbf{F}^ϵ , based mostly on heuristic considerations, is $\mathbf{F}^\epsilon = \mathbf{F} + \epsilon \omega(\mathbf{r}, t)$, in which $\omega(\mathbf{r}, t)$ is of white noise type (both in space and time) or the gradient, in the spatial variable, of a white noise. Examples (in which there is in fact no explicit small parameter ϵ) include the *fluctuating hydrodynamic theory* of Landau and Lifshitz (see [73], p.524) and the fluctuating Boltzmann equation, [95]. In these examples the covariances of the fluctuating terms are chosen on the basis of equilibrium statistical mechanical considerations. These equations are quite successful in predicting the general pattern of behavior occurring near the onset of instabilities or bifurcations. Most of these considerations follow however by considering linearized versions of

the original PDE's. This is just as well, since generally these type of noise terms make the nonlinear problem mathematically ill-posed.

To help give mathematical meaning to such SPDE's we shall here review their origin in simple models in which the transition from microscopic to macroscopic scales and ipso facto the corrections to (1.1) can be investigated with some mathematical rigor. To be more precise we will consider lattice systems with stochastic microscopic dynamics, the so called Interacting Particle Systems (IPS) [30],[81],[95], where the configurations are updated at random (Poisson) times according to some local rules. The importance of models is well established for equilibrium behavior, where the simple *Ising model*, has played a fundamental role in elucidating the nature of *phase transitions*, both from the physical and the mathematical viewpoint. There is unfortunately no comparable model for the more complex non equilibrium phenomena. This is not surprising since even the solutions of the macroscopic hydrodynamic equations are far from being understood in many interesting cases. Nevertheless, computer simulations and even some rigorous analysis have shown that very simple IPS models can capture the essence of large scale hydrodynamic behavior and so we shall assume that their behavior is of physical interest.

The IPS dynamics are divided into two main groups: the Glauber (or *spin-flip*) dynamics, which does not conserve the sum of the occupation variables (or any other analogous quantity), and the Kawasaki (or exchange) dynamics, which has at least one (particle) conservation law. For both classes the derivation of macroscopic behavior from microscopic model systems via hydrodynamical scaling limits is currently an active field of research. There are reviews of this work to which we refer the reader for background [30],[95].

1.1. The Micro–Macro Connection. The hydrodynamic limit (or hydrodynamic scaling limit: HSL) gives a reduced description of the collective behavior of some particle systems which involves only the slowly varying fields, e.g. those that satisfy local conservation laws: generally the same ones that describe the thermodynamic equilibria of the system. The hydrodynamic equations are obtained mathematically by suitable space-time scaling limits; we have a scaling parameter ϵ such that as $\epsilon \rightarrow 0^+$ the evolution converges, in a sense to be specified, to some autonomous deterministic limit laws of the form (1.1). Corrections may then be introduced to describe *physically important* effects present in the system when $\epsilon > 0$ which are lost after the limit and thus no longer present in the PDE. We shall always take ϵ to be the ratio of microscopic to macroscopic length scales ($\epsilon \sim 10^{-10} - 10^{-5}$ in many typical physical situations). The ratio of the macroscopic time scale to the microscopic one will be $\epsilon^{-\alpha}$, with α between 1 and 4 in the problems considered in this chapter; $\alpha = 1, 2$ are generally referred to as respectively the Euler and the diffusive scale. The HSL leaves out many interesting features of the IPS. We enumerate some of these here.

- 1). Small fluctuations in the hydrodynamical variables about their deterministic values provided by the solutions of (1.1). These will always be present since, by definition, the description given by the HSL is only valid when $\epsilon \rightarrow 0$, while in physical situations ϵ is always finite. These fluctuations have been studied most extensively for systems in equilibrium, when the right hand side of (1.1) vanishes. In the case of systems with unique equilibrium states (away from the critical point), the situation is analogous to the behavior of sums of (approximately) independent

random variables. The deterministic behavior corresponds to the law of large numbers while the fluctuations around stable evolutions obey the central limit theorem.

2). Deterministic corrections to (1.1) which change the qualitative behavior of the evolution on time scales larger than those on which (1.1) holds. A very striking example of such a situation are the expected corrections, on the time scale ϵ^{-2} , to the Euler equations of hydrodynamics, which give the correct behavior on the time scale ϵ^{-1} [38],[24]. These correction terms, containing the viscosity and heat conductivity, lead to the Navier-Stokes equations [73], whose solutions may have a behavior which is very different from that of the Euler equations. While the time evolution described by the Euler equations is, in the absence of shocks, entirely reversible, that described by the Navier-Stokes equations is time asymmetric, leading, in the absence of external forcing, to a final equilibrium independent of the initial conditions. We shall generally refer to such deterministic corrections to (1.1), leading to deterministic equations of form (1.2), as *Navier-Stokes corrections*. The rigorous derivation of such corrections for IPS has been achieved so far only in some very special cases, corresponding to the incompressible Navier-Stokes case ([38],[39],[74]). We shall discuss some such examples later.

3). Stochastic corrections to (1.1) which go beyond that of small fluctuations discussed in 1). These concern mostly cases in which the solutions of (1.1) are not linearly stable. They correspond in many ways to phase transitions and critical phenomena in equilibrium statistical mechanics. These corrections lead to the most difficult and also most interesting questions regarding the utility of SPDE's for improving the HSL. Unfortunately all we shall be able to do in most cases is raise questions rather than give answers.

1.2. Outline of the Chapter. The outline of the rest of the chapter is as follows. Part I (Sections 2 to 5) deals with asymmetric non-reversible models. In Section 2 we will introduce the HSL by starting from a *microscopic model* which is itself a deterministic PDE for a scalar field, i.e. the Burgers equation (2.1). The coarse graining leads to a hydrodynamic equation (1.1) which in this case is the same Burgers equation but without a viscosity term, (2.2). The corrected equation (1.2) will simply be the viscous Burgers equation (2.1) we started from.

In Section 3 we introduce two IPS, the asymmetric simple exclusion process (ASEP) and, for comparison, the independent particle process (IPP). The hydrodynamic equations for these systems are respectively the inviscid Burgers equation (2.1) and the free streaming equation (3.9). The first question we examine there concerns the Navier-Stokes correction, namely the analysis of effects in the particle system that require the addition of a viscosity term. Following this we discuss the behavior of the shocks, which, for the inviscid Burgers equation, are discontinuities of the density profile. Thus, at the shock, the basic assumption for hydrodynamic behavior, i.e. the slowly varying condition, fails. We will see that at a finer level of description (than the hydrodynamic one) stochastic effects become dominant. These may be explained still using the Burgers equation: the stochasticity will only involve the initial datum, namely the initial random fluctuations cannot be neglected, but those arising from the evolution are, at this order, negligible.

In Section 4 we study the ASEP in the case of weak asymmetry, the so called *weakly asymmetric simple exclusion process* (WASEP). For this system we can obtain, in suitable scalings, corrections that are captured by the addition of a stochastic forcing term to the Burgers equation. Some of the results for the WASEP

are translated in Section 5 into the language of growing interfaces [72],[67]. Such interfaces often display anomalous (nonlinear) fluctuations: the particular model considered here is the *restricted solid on solid* model. The microscopic analogy between the ASEP and an interface model is reflected by the macroscopic analogy between the stochastic Burgers equation and the Kardar–Parisi–Zhang equation [67].

In Part II (Sections 6 to 9) we consider IPS with *long range potentials*. These give rise to non trivial translation invariant states of the dynamics, which is now not necessarily conservative. We discuss first, in Section 6, the hydrodynamics, the small fluctuations and the relation between Ornstein–Uhlenbeck processes and *fluctuating hydrodynamic theory* for these systems. We then describe, in Section 7, the critical fluctuations (mostly for $d = 1$), i.e. fluctuations at the critical point, leading to one dimensional Reaction–Diffusion [93] and Cahn–Hilliard [13] equations with noise. In Section 8 we consider the fluctuations in unstable situations and the macroscopic effects that they can generate: in the context which we consider, the noise will trigger the origin of (random) interfaces. In Section 9 we look briefly at large deviation phenomena as a possible means for establishing the nature of the stochastic corrections to the HSL.

PART I

NON REVERSIBLE DYNAMICAL SYSTEMS: ASYMMETRIC MODELS WITH SHOCKS

2. The Burgers Equation

To make concrete some of the issues discussed in the Introduction we consider in this section the relatively trivial case of a *microscopic model* which is itself described by a deterministic PDE. This is the Burgers equation

$$(2.1) \quad \frac{\partial \rho}{\partial t} + v \cdot \nabla \rho (1 - \rho) = \nabla \cdot [D \nabla \rho],$$

where $\rho = \rho(r, t) \in [0, 1]$, $r \in \mathbb{R}^d$, $t \geq 0$, $v \in \mathbb{R}^d$, $v \neq 0$, and $D \geq 0$ a non negative, constant matrix. The equation with $D = 0$

$$(2.2) \quad \frac{\partial \rho}{\partial t} + v \cdot \nabla \rho (1 - \rho) = 0,$$

is called the inviscid Burgers equation. We recall, see for instance [93], that the Cauchy problem for (2.2) in general has no global, classical solution, as there are smooth initial data that develop singularities in a finite time, after which the solution is defined only in a weak sense. Uniqueness, which is lost at this stage, may be recovered by restricting to the class of *entropic solutions*. The equation has therefore a rich and intriguing mathematical structure and is of physical interest in fluid mechanics as a model for the formation and propagation of shocks. In the next section we will see that it is naturally related to an important class of stochastic IPS.

We begin with the HSL: while it is usually considered for IPS, the definition extends to several other dynamical systems. In particular we will next regard any of the Burgers equations (2.1) as defining the evolution of a microscopic system to which we want to apply the HSL procedure. Thus the space and time variables in

(2.1), that we denote by x and t , should be now regarded as microscopic coordinates. We then choose an initial datum that depends on a small, positive parameter ϵ and set $\rho(x, 0; \epsilon) = \rho_0(\epsilon x)$, ρ_0 a fixed, smooth function on \mathbb{R}^d . Thus when ϵ is small $\rho(x, 0; \epsilon)$ is a *slowly varying function* of x . Calling $\rho(x, t; \epsilon)$ the solution of (2.1) which starts from $\rho(x, 0; \epsilon)$, we write

$$(2.3) \quad \rho^{(\epsilon)}(r, \tau) := \rho(\epsilon^{-1}r, \epsilon^{-1}\tau; \epsilon),$$

With this change of variables in general $\rho^{(\epsilon)}(r, \tau)$ will no longer be a slowly varying function of its arguments, it is then natural to regard $r = \epsilon x$ and $\tau = \epsilon t$ as the macroscopic space and time coordinates and ϵ as the ratio between microscopic and macroscopic units. $\rho^{(\epsilon)}(r, \tau)$ and $\rho(x, t; \epsilon)$ are just the same density profile, but expressed in macroscopic and microscopic coordinates, respectively.

The HSL is, by definition, the limit behavior of $\rho^{(\epsilon)}(r, \tau)$ as $\epsilon \rightarrow 0$. The existence of a limit proves that the macroscopic description of the profile becomes less and less sensitive to the actual value of the ratio between micro and macro scales. Thus the limit $\epsilon \rightarrow 0$ and the HSL should be regarded as a macroscopic limit. If the evolution is given by (2.1) we have for any test function ϕ and any $\tau \geq 0$

$$(2.4) \quad \lim_{\epsilon \rightarrow 0^+} \int_{\mathbb{R}^d} dr \rho^{(\epsilon)}(r, \tau) \phi(r) = \int_{\mathbb{R}^d} dr \rho(r, \tau) \phi(r),$$

where $\rho(r, \tau)$ is the entropic solution of (2.2) starting from $\rho_0(r)$, [93].

Notice that the limit (2.4) holds independently of which D is used in (2.1) to define $\rho(x, t; \epsilon)$, namely all the equations (2.1) have the same hydrodynamic limit described by the same inviscid Burgers equation (2.2) (with the prescription of taking its entropic solution, a rule thus justified by this procedure). One of the main questions about hydrodynamic limits is therefore to determine domains of attraction to given hydrodynamic equations: as we will see, the domain of attraction of (2.2) is much larger than the class of evolutions defined by (2.1). Observe that, in particular, the hydrodynamic limit of (2.2) is (2.2) itself. In general the *pure hydrodynamic equations* are those identical to their hydrodynamic limit. They are thus fixed points of the scaling transformation (2.3) and they are at the origin of some scaling phenomena observed in macroscopic systems.

It is clear that for the *microscopic model* we have considered, namely (2.1), there are only Navier-Stokes corrections to the hydrodynamic equation (1.1), which is just (2.2). These corrections are in fact trivial: the analog of (1.2) is here just the same as (2.1) written in macroscopic coordinates. In fact after the substitution (2.3), (2.1) becomes

$$(2.5) \quad \frac{\partial \rho^{(\epsilon)}}{\partial \tau} + v \cdot \nabla \rho^{(\epsilon)} (1 - \rho^{(\epsilon)}) = \epsilon \nabla \cdot [D \nabla \rho^{(\epsilon)}].$$

The HSL procedure has been applied to a large variety of deterministic equations, including cases where they are *far* from their hydrodynamic limits: the analysis is then much more intricate than in the simple case we have considered above. We just mention, very briefly, two of the most interesting examples. In the first one the Boltzmann equation plays the role of the microscopic model and the *true* Euler and Navier Stokes equations, [14],[23], are its hydrodynamic limits, respectively in the Euler and in the *diffusive-incompressible* scaling limits. The second example is rather different since the microscopic model, in contrast to the previous cases, has no conservation laws. The evolution is defined by the

Allen-Cahn equation

$$(2.6) \quad \frac{\partial u}{\partial t} = \Delta u - U'(u)$$

in \mathbb{R}^d , $d \geq 2$ with $U(\cdot)$ a double well, symmetric potential: to be definite let us take $U(u) = u^4/4 - u^2/2$. The HSL procedure, object of many studies in the recent years, see for instance [41] and references therein, is applied to initial data that are equal (or approximately equal) to either of the two phases ($u = \pm 1$, the minimizers of $U(u)$) in macroscopically large regions. The diffusive scaling limit gives rise to a motion by mean curvature which describes the macroscopic forces (due to the surface tension) between the two phases at the interface. In the second part of this chapter we will come back to some problems connected to (2.6).

3. The Asymmetric Simple Exclusion and the Independent Particle System

A single random walk is the process on \mathbb{Z}^d where the particle (say at x) waits for an exponential time of mean 1 and then jumps to a site y chosen with probability $P(x, y)$, x, y in \mathbb{Z}^d . After that it starts anew from y , independently of the past. We suppose that $P(x, y)$ is translation invariant, $P(x, y) = P(0, y - x)$, nearest neighbor, $P(x, y) = 0$ unless $|x - y| = 1$, and asymmetric:

$$(3.1) \quad \sum_y P(0, y)y = v \neq 0.$$

The ASEP is a Markov process which describes the evolution of random walks on \mathbb{Z}^d which interact only by exclusion: the exclusion rule (hard core interaction) prohibits the occupation of any lattice site by more than one particle. Starting from any configuration satisfying the exclusion rule we stipulate that if the site chosen by a particle for jumping on is already occupied by another particle then that jump is suppressed and the particle stays where it was, waiting for the next attempt. The ASEP is therefore a process on $\{0, 1\}^{\mathbb{Z}^d}$; its elements, denoted by $\eta = \{\eta(x), x \in \mathbb{Z}^d\}$, are particle configurations, $\eta(x) = 1$ meaning that x is occupied by a particle (particles are here indistinguishable). Let L be the (pre)generator of the process, which is defined by its action over the cylindrical functions $f(\eta)$, i.e. functions which depend only on the values of $\eta(x)$, for x in a finite set:

$$(3.2) \quad Lf(\eta) = \sum_{x,y} P(x, y)\eta(x)(1 - \eta(y))[f(\eta^{x,y}) - f(\eta)],$$

where $\eta^{x,y}$ is the configuration obtained from η by exchanging the occupation numbers at x and y , that is

$$(3.3) \quad \eta^x(y) = \begin{cases} \eta(y) & \text{if } y \neq x \\ -\eta(y) & \text{if } y = x. \end{cases}$$

The closeness of the ASEP to the Burgers equation can be readily understood with the help of the following heuristic argument. Let $f(\eta) = \eta(x)$ in (3.2), then

$$(3.4) \quad L\eta(x) = - \sum_{i=1}^d [j_i(x, x + e_i) - j_i(x - e_i, x)],$$

where e_i is the unit vector along the i -th direction and $j_i(x, x + e_i)$ is the (expected) current through the bond $(x, x + e_i)$:

$$(3.5) \quad j_i(x, x + e_i) = P(x, x + e_i)\eta(x)(1 - \eta(x + e_i)) - P(x + e_i, x)\eta(x + e_i)(1 - \eta(x)).$$

Under a propagation of chaos ansatz, supported by the fact that the Bernoulli measures (a Bernoulli measure on a product space is a product measure) ν_ρ on \mathbb{Z}^d (with density $\rho \in [0, 1]$) are invariant for the ASEP, we get

$$(3.6) \quad \begin{aligned} E(j_i(x, x + e_i, t)) &\approx P(x, x + e_i)E(\eta(x, t))[1 - E(\eta(x, t))] \\ &\quad - P(x + e_i, x)E(\eta(x + e_i, t))[1 - E(\eta(x, t))]. \end{aligned}$$

Assuming that $E(\eta(x, t))$ is a slowly varying function of x and keeping only the leading orders we get from (3.4) and (3.6) the inviscid Burgers equation (2.2) with v as in (3.1).

The statement is made precise using the HSL procedure described in the previous section. Let then ρ_0 be a smooth function on \mathbb{R}^d and let the ASEP start from the product measure $\mu^{(\epsilon)}$ on $\{0, 1\}^{\mathbb{Z}^d}$ with averages

$$(3.7) \quad E_{\mu^{(\epsilon)}}(\eta(x)) = \rho_0(\epsilon x) \quad x \in \mathbb{Z}^d$$

This choice makes the local equilibrium condition (reasonably well) satisfied at time 0, since the Bernoulli measures ν_ρ with constant average ρ are the true equilibria of the ASEP. Let $P_{\mu^{(\epsilon)}}$ be the law of the ASEP starting from $\mu^{(\epsilon)}$, then, recalling (2.3) and (2.4),

$$(3.8) \quad \lim_{\epsilon \rightarrow 0} P_{\mu^{(\epsilon)}} \left(\left| \epsilon^d \sum_{x \in \mathbb{Z}^d} \phi(\epsilon x) \eta(x, \epsilon^{-1}\tau) - \int_{\mathbb{R}^d} dr \phi(r) \rho(r, \tau) \right| < \delta \right) = 1,$$

where $\rho(r, \tau)$ is the entropic solution of (2.2) with v as in (3.1) and initial datum ρ_0 .

Theorem 3.1 is proved in [91] where more general initial data are also allowed. The proof follows previous works where particular cases had been studied in great detail, see references in [91].

The most remarkable fact about Theorem 3.1 is its validity even after the formation of shocks when the condition of slow variations in the previous heuristic argument fails. The use of the test functions ϕ in the limit (3.8) allows for a coarse graining analysis that avoids a finer examination of the behavior of the system at the shock and leaves open the question of why the system still follows (2.1) even when the slow variation condition is not satisfied. A possible and instructive guess is that the shock is such only at the macroscopic level and that on a finer scale, intermediate between micro and macro, it is resolved into a smooth function. This turns out to be wrong for the ASEP but right for the independent particle process (IPP) where an analogue of Theorem 3.1 holds, as we shall see next.

3.1. The Independent Particle Process: Hydrodynamics and Navier-Stokes Corrections. The IPP is the Markov process on $\mathbb{N}^{\mathbb{Z}^d}$ defined by random walks on \mathbb{Z}^d which evolve independently of each other. The extremal, invariant measures (still denoted by) ν_ρ , $\rho \in \mathbb{R}_+$, are the product measures where each $\eta(x)$ has the Poisson distribution of mean ρ . The initial condition $\mu^{(\epsilon)}$ is then chosen as the product of Poisson distributions with averages $\rho_0(\epsilon x)$. Theorem 3.1 holds with $\rho(r, \tau)$ in (3.8) satisfying

$$(3.9) \quad \frac{\partial \rho}{\partial t} + v \cdot \nabla \rho = 0.$$

No shocks are present here since (3.9) is a linear equation, but we can always simulate a shock by choosing ρ_0 equal to

$$(3.10) \quad \chi_{\rho_-, \rho_+}(r) = \rho_- \mathbf{1}_{\{r_1 \leq 0\}}(r) + \rho_+ \mathbf{1}_{\{r_1 > 0\}}(r), \quad r_1 = r \cdot e_1, \quad 0 \leq \rho_- < \rho_+.$$

We suppose $v \cdot e_1 \neq 0$. The solution of (3.9) with initial datum χ_{ρ_-, ρ_+} is

$$(3.11) \quad \chi_{\rho_-, \rho_+}(r, t) = \chi_{\rho_-, \rho_+}(r - vt)$$

which is the same step function moving rigidly with velocity $v \cdot e_1$.

To determine a (possible) smoothening of the shock, which would be incompatible with (3.9), we need a criterion finer than that used in (3.8). To this end it is convenient to extend the test functions ϕ to include characteristic functions of suitable sets. Let $C^{(\ell)}(x)$ be the cube in \mathbb{R}^d of side ℓ and center x , $|C^{(\ell)}(x)|$ its volume, $\rho(y)$ a function on \mathbb{R}^d . We define the distance between a configuration η and a density $\rho(y)$, $y \in \mathbb{R}^d$, as

$$(3.12) \quad d_{\ell, x}(\eta, \rho) = \frac{1}{|C^{(\ell)}(x)|} \left| \sum_{y \in C^{(\ell)}(x) \cap \mathbb{Z}^d} \eta(y) - \int_{C^{(\ell)}(x)} dy \rho(y) \right|.$$

We use the distance (3.12) to compare the random particle configurations $\eta(\cdot, t)$ and the density $\rho(y, t; \epsilon) = \rho(\epsilon y, \epsilon t)$, where $\rho(r, \tau)$ is the solution of (3.9) starting from ρ_0 , i.e. $\rho(y, t; \epsilon)$ is the same $\rho(r, \tau)$, but in microscopic coordinates (recall the discussion in Section 2). It can be proved that for any $\tau > 0$, $\ell > 0$ and $R > 0$

$$(3.13) \quad \lim_{\epsilon \rightarrow 0} \sup_{|x| \leq \epsilon^{-1} R} P_{\mu^{(\epsilon)}} \left(d_{\epsilon^{-1} \ell, x} (\eta(\cdot, \epsilon^{-1} \tau), \rho(\cdot, \epsilon^{-1} \tau; \epsilon)) < \delta \right) = 1.$$

Without the supremum, taking $x = \epsilon^{-1} \bar{r}$, \bar{r} a fixed point in \mathbb{R}^d , (3.13) becomes (3.8) with $\phi(r) = \mathbf{1}_{\{|r - \bar{r}| \leq \ell\}}$, see for instance Ch.2 of [30], to which we also refer for the statements below that concern the IPP. We also mention that (3.13) holds as well with the supremum inside the probability.

Let us next go back to the shocks, taking as the initial datum the function χ_{ρ_-, ρ_+} in (3.10) and improving our accuracy by replacing in (3.13) $\epsilon^{-1} \ell$ with $\epsilon^{-\alpha}$, $\alpha \in (0, 1)$. The result changes drastically: if $R > |v|\tau$, $\alpha \leq 1/2$ and $\delta > 0$ is small enough, then

$$(3.14) \quad \lim_{\epsilon \rightarrow 0} \sup_{|x| \leq \epsilon^{-1} R} P_{\mu^{(\epsilon)}} \left(d_{\epsilon^{-\alpha}, x} (\eta(\cdot, \epsilon^{-1} \tau), \chi_{\rho_-, \rho_+}(\cdot, \epsilon^{-1} \tau; \epsilon)) < \delta \right) = 0$$

At this level of accuracy, therefore, (3.9) does not describe anymore the behavior of the IPP.

Consider the new equation

$$(3.15) \quad \frac{\partial \rho^*}{\partial t} + v \nabla(\rho^*) = \nabla \cdot [D \nabla \rho^*],$$

with D the constant matrix

$$(3.16) \quad D_{i,j} = \frac{1}{2} \delta_{i,j} [P(0, e_i) + P(0, -e_i)].$$

Call $\rho^*(x, t)$ the solution of (3.15) starting from χ_{ρ_-, ρ_+} , then for any $R > 0$, $\tau \geq 0$ and $\alpha > 0$,

$$(3.17) \quad \lim_{\epsilon \rightarrow 0} \sup_{|x| \leq \epsilon^{-1}R} P_{\mu^{(\epsilon)}} \left(d_{\epsilon^{-\alpha}, x} (\eta(\cdot, \epsilon^{-1}\tau), \rho^*(\cdot, \epsilon^{-1}\tau)) < \delta \right) = 1.$$

By comparing (3.14) and (3.17) we conclude that (3.15) gives a better description of the IPP than (3.9), while in the pure hydrodynamic limit (3.13) they are indistinguishable.

An alternative, maybe more convenient way to select (3.15) is to look at longer times while keeping the same coarse grained accuracy as in (3.13). We have for any $\tau \geq 0$, $\ell > 0$ and $R > 0$:

$$(3.18) \quad \lim_{\epsilon \rightarrow 0} \sup_{|x| \leq \epsilon^{-2}R} P_{\mu^{(\epsilon)}} \left(d_{\epsilon^{-1}\ell, x} (\eta(\cdot, \epsilon^{-2}\tau), \rho^*(\cdot, \epsilon^{-2}\tau)) < \delta \right) = 1,$$

while the limit is 0 if ρ^* is replaced by $\chi_{\rho_-, \rho_+}(\cdot, \epsilon^{-2}\tau)$, see (3.11) for notation.

This formulation allows for a scaling limit procedure similar to that in Theorem 3.1. To be more precise let $\rho^*(x, 0; \epsilon) = \rho_0(\epsilon x)$ and set

$$(3.19) \quad \rho^{(\epsilon)}(r, \tau) = \rho^*(\epsilon^{-1}r + \epsilon^{-2}v\tau, \epsilon^{-2}\tau; \epsilon),$$

then $\rho^{(\epsilon)}$ is actually independent of ϵ , $\rho^{(\epsilon)} = \tilde{\rho}$ where

$$(3.20) \quad \frac{\partial \tilde{\rho}}{\partial \tau} = \nabla \cdot D \nabla \tilde{\rho}$$

with $\tilde{\rho}(r, 0) = \rho_0(r)$. We then have that, for a general class of initial data ρ_0 ,

$$(3.21) \quad \lim_{\epsilon \rightarrow 0} P_{\mu^{(\epsilon)}} \left(\left| \epsilon^d \sum_{x \in \mathbb{Z}^d} \phi(\epsilon[x - v\epsilon^{-2}\tau]) \eta(x, \epsilon^{-2}\tau) - \int_{\mathbb{R}^d} dr \phi(r) \tilde{\rho}(r, \tau) \right| < \delta \right) = 1.$$

Thus a diffusive scaling (with the introduction of *Lagrangian coordinates*) gives rise to the heat equation (3.20), while the Euler scaling with space and time scaled by the same factor gives the hyperbolic equation (3.9): longer times in the IPP yield a non zero Navier Stokes correction.

3.2. Random Fluctuations of the Shock in the ASEP. It is natural at this point to ask whether similar Navier-Stokes corrections arise also in the ASEP, the answer, as we will see, is negative. Let $d = 1$ and, going back to (3.2), set

$$(3.22) \quad p = P(0, 1) > 1/2, \quad q = P(0, -1), \quad p + q = 1.$$

The shocks in the inviscid Burgers equation have also the form (3.10)-(3.11) with $0 \leq \rho_- < \rho_+ \leq 1$ and

$$(3.23) \quad \chi_{\rho_-, \rho_+}(x, t) = \chi_{\rho_-, \rho_+}(x - ct), \quad c = (p - q)(1 - \rho_- - \rho_+).$$

Since $\chi_{\rho_-, \rho_+}(\epsilon x)$ is independent of ϵ , the family of initial measures $\mu^{(\epsilon)}$ is independent of ϵ , $\mu^{(\epsilon)} \equiv \mu$ being the product measure with averages $\chi_{\rho_-, \rho_+}(x)$. (3.14) holds also for the ASEP, but, contrary to what happens in the IPP, it cannot be

fixed, as in (3.18), by replacing $\chi_{\rho_-, \rho_+}(x, t)$ by the solution of the equation with an additional viscosity. In fact let $\rho(x, t)$ be the solution of (2.1) with some $D > 0$ (in $d = 1$, D is a scalar) starting from χ_{ρ_-, ρ_+} : it is then possible to show that

$$(3.24) \quad \limsup_{t \rightarrow \infty} \sup_{x \in \mathbb{R}} |\chi_{\rho_-, \rho_+; K}(x, t) - \rho(x, t)| = 0,$$

where the functions $\chi_{\rho_-, \rho_+; K}(x, t)$ are the traveling waves

$$(3.25) \quad \chi_{\rho_-, \rho_+; K}(x, t) = \frac{\rho_- + \rho_+}{2} + \frac{\rho_+ - \rho_-}{2} \tanh \left\{ \frac{x - ct}{K} \right\}, \quad K = \frac{(\rho_+ - \rho_-)D}{4(p - q)}$$

(ρ_- , ρ_+ and c as before).

The discontinuity at $x = 0$ of $\chi_{\rho_-, \rho_+}(x)$ disappears instantaneously in (2.1) with $D > 0$ just as in (3.15), but, contrary to what happens in the latter, the shock width does not diverge as $t \rightarrow \infty$, instead it approaches that in (3.25). $\chi_{\rho_-, \rho_+; K}(\cdot, t)$ behaves just like $\chi_{\rho_-, \rho_+}(\cdot, t)$ as far as the criterion (3.14) is concerned and since the inviscid Burgers equation does not approximate correctly the ASEP, in the sense of (3.14), then none of the Burgers equations, no matter what is $D > 0$, gives a correct approximation either. The shocks (3.25) have an essentially finite width (of the order of K) and they could only be distinguished from the shock χ_{ρ_-, ρ_+} by a finite coarse graining, i.e. if $\epsilon^{-1/2}\ell$ is substituted by ℓ . But then the statistical fluctuations intrinsic to the microscopic description would be no longer negligible.

The failure of the Burgers equation to describe the ASEP is not at all due to a problem of Navier-Stokes corrections, but rather to a completely new phenomenon of stochastic nature:

THEOREM 3.2. *Given any ρ_- and ρ_+ as in (3.23) there is a probability ν on $\{0, 1\}^{\mathbb{Z}}$ and a random process ξ_t , $t \geq 0$, $\xi_0 = 0$, with values in \mathbb{Z} , adapted to $\eta(\cdot, t)$, $t \geq 0$, such that for any $\delta > 0$, $\alpha > 0$, $\tau \geq 0$ and $R > 0$*

$$(3.26) \quad \lim_{\epsilon \rightarrow 0} \sup_{|x| \leq \epsilon^{-1}R} P_{\nu} \left(d_{\epsilon^{-\alpha}, x} (\eta(\cdot, \epsilon^{-1}\tau), \chi_{\rho_-, \rho_+}(\cdot + \xi_{\epsilon^{-1}\tau}, 0)) < \delta \right) = 1.$$

Moreover

$$(3.27) \quad \lim_{\epsilon \rightarrow 0} E_{\nu} (\epsilon \xi_{\epsilon^{-1}\tau}) = c,$$

where c is the speed of the shock, given in (3.23).

This beautiful result, proved in [43], shows that in all the typical configurations of the ASEP which starts from ν , there is always, somewhere, a shock. Initially it is at 0 then it moves randomly as ξ_t .

By (3.26) with $\tau = 0$ we can say that ν is an acceptable choice for describing the initial macroscopic profile χ_{ρ_-, ρ_+} , just like the traditional product measure μ . (3.26) can then be seen as the analogue of (3.18) in the case of the ASEP. The result proved in [43] is even stronger since it states also that the measure ν is invariant when the ASEP is seen from ξ_t . In particular this means that we can take any time, say $\epsilon^{-100}\tau$, and still find the shock somewhere, in all the typical configurations of the process.

There are sharper results about the location of the shock. We restrict ourselves for simplicity to the totally asymmetric ASEP, i.e. $p = 1$ in (3.22), and refer to [44] for a more detailed survey on the argument and for references. It is proved that the deviation

$$(3.28) \quad \epsilon^{1/2} \left\{ \xi_{\epsilon^{-1}\tau} - c\epsilon^{-1}\tau \right\}$$

converges weakly to a Brownian motion. The main contribution to these fluctuations can be ascribed to the randomness of the initial datum and the behavior of the shocks in the ASEP can be described, at the level of accuracy (3.26), by the Burgers equation with any given value D of the viscosity in terms of the solution $\rho = \rho(r, t; \eta)$, $r \in \mathbb{R}$, $t \geq 0$, of the following Cauchy problem

$$(3.29) \quad \frac{\partial \rho}{\partial t} + v \cdot \frac{\partial}{\partial x} \rho(1 - \rho) = D \frac{\partial^2}{\partial x^2} \rho,$$

with initial datum:

$$(3.30) \quad \rho(x, 0; \eta) = \eta([x]),$$

where $[x]$ denotes the integer part of x . Here the $\eta(x)$'s are independent random variables with average $\chi_{\rho_-, \rho_+, K}(x, 0)$ and K is related to D as in (3.25). To be more specific, define

$$(3.31) \quad x_{\pm}(t) = \pm(p - q)[1 - 2\rho_- - (1 - \rho_- - \rho_+)]t$$

($p - q = 1$ in the actual case we are considering) and

$$(3.32) \quad \xi_t^* = \frac{1}{\rho_+ - \rho_-} \sum_{x \in \mathbb{Z} \cap [x_{-(t)}, x_{+(t)}]} \eta(x).$$

The main result is that if ν (the same as in Theorem 3.2) is the distribution of the η 's, then

$$(3.33) \quad \lim_{\epsilon \rightarrow 0} E_\nu \left(\epsilon^{1/2} |\xi_{\epsilon^{-1}\tau} - \xi_{\epsilon^{-1}\tau}^*| \right) = 0$$

(the same result holds starting from the product measure with averages χ_{ρ_-, ρ_+}). Moreover for any $\delta > 0$, $\alpha > 0$, $\tau \geq 0$ and $R > 0$, we expect that

$$(3.34) \quad \lim_{\epsilon \rightarrow 0} \sup_{|x| \leq \epsilon^{-1}R} P_\nu \left(d_{\epsilon^{-\alpha}, x} (\rho(\cdot, \epsilon^{-1}\tau; \eta), \chi_{\rho_-, \rho_+, K}(\cdot + \xi_{\epsilon^{-1}\tau}^*, 0)) < \delta \right) = 1.$$

We do not know of a proof of the above statement which should thus be regarded as a conjecture (partial results in this direction can however be found in [34]).

According to Theorem 3.2, the random microscopic position of the shock, ξ_t , has a nice interpretation in terms of second class particles, since it is the position of a second class particle added to the system. The dynamics of a second class particle is like that of the other particles in exchanges with vacancies, but it acts like a vacancy in exchanges with regular (i.e. first class) particles. In this way the second class particle does not affect the dynamics of the first class particles. It can be shown (and this is the reason for introducing it here) that the second class particle moves with velocity $1 - 2\rho_-$ (respectively $1 - 2\rho_+$) to the left and right of the shock and hence it is attracted by the shock. The particles around it develop a pattern which is described by the measure ν , whose densities are asymptotically ρ_+ and ρ_- respectively to the right and to the left of the second class particle. A complete characterization of ν , first introduced in [43], has been obtained in [33] where it is shown that, looking from the second class particle, the asymptotic

convergence to the Bernoulli measures with densities ρ_- and ρ_+ is exponentially fast.

We have seen so far how to localize the shock, we next discuss about its shape, in particular whether the smoothed shock $\chi_{\rho_-, \rho_+, K}$, (3.25) is better than the sharp step χ_{ρ_-, ρ_+} , (3.10). *Better* here means a better approximation to the profiles described by the measure ν . Since we need large averaging blocks $C^{(\ell)}$ to dampen the statistical fluctuations of the empirical density, the shape of the shock is essentially undetectable unless its width is much larger than the side of the averaging block. The Burgers profile $\chi_{\rho_-, \rho_+, K}$ has width proportional to K . K can be made larger by decreasing the asymmetry $p - q$, see (3.25). In [33] it is shown that as $p - q \rightarrow 0$ the profiles of the typical configurations of ν agree with $\chi_{\rho_-, \rho_+, K}$. The statement is made precise under a proper rescaling where $p - q = \epsilon$ is the micro-macro space ratio. The ASEP in such a case is called the *weakly asymmetric simple exclusion process* (WASEP), a process that will be discussed in the next section.

We conclude this section with a few words on the problem of the Navier-Stokes correction for the ASEP. We have seen that contrary to what happens in the IPP the analysis of the shock does not select any value for the viscosity coefficient. There is however evidence that the correct equation for the ASEP is a Burgers equation with viscosity. This comes from studying the ASEP in $d = 3$, starting from an initial datum $\rho^{(\epsilon)}(x) = 1/2 + \epsilon\theta(\epsilon x)$, where θ is independent of ϵ . It has been proved in [38] that at times $\epsilon^{-2}\tau$ the correct profile is, to leading orders, $1/2 + \epsilon\theta(\epsilon x, \tau)$ with $\theta(r, \tau)$ satisfying a Burgers equation with a density dependent diffusion coefficient obtained from a Green-Kubo formula [95].

4. The Weakly Asymmetric Simple Exclusion Process: Hydrodynamics and Stochastic Corrections

We have seen in the previous section that there are stochastic effects which survive in the hydrodynamic limit. They appear as small, random displacements of the shock profile caused primarily by the fluctuations of the initial data (and not by the randomness in the dynamics). They are therefore more appropriately treated in the hydrodynamic description by the Burgers equation with random initial condition rather than by adding a random forcing term to the Burgers equation.

In this section we will instead examine a variation of the ASEP, whose behavior appears to be captured by the following one dimensional SPDE

$$(4.1) \quad \frac{\partial}{\partial t} \rho^{(\epsilon)} = \frac{1}{2} \frac{\partial^2}{\partial x^2} \rho^{(\epsilon)} - \frac{\partial}{\partial x} \left[\rho^{(\epsilon)} (1 - \rho^{(\epsilon)}) + \sqrt{\epsilon \rho^{(\epsilon)} (1 - \rho^{(\epsilon)})} \dot{W}_t^\epsilon \right],$$

with

$$(4.2) \quad \dot{W}_t^\epsilon = \varphi_\epsilon * \dot{W}_t,$$

φ_ϵ an approximate δ -function, i.e. $\varphi_\epsilon(\cdot) = \epsilon^{-1}\varphi(\cdot\epsilon^{-1})$, φ a non negative, even, compactly supported, smooth function with $\int_{\mathbb{R}} \varphi(r) dr = 1$, and $*$ denotes the convolution. \dot{W}_t is the space-time white noise, that is the Gaussian process with mean zero and covariance

$$(4.3) \quad E \left(\dot{W}_t(r) \dot{W}_{t'}(r') \right) = \delta(t - t')\delta(r - r').$$

The particle system we consider here is the one dimensional weakly asymmetric simple exclusion process (WASEP) already mentioned in the previous section.

Recall that this is the exclusion process with p and q in (3.22) chosen as

$$(4.4) \quad p = \frac{1+\epsilon}{2}, \quad q = \frac{1-\epsilon}{2}.$$

This means that the bias v in (3.1) is equal to ϵ and goes to zero in the hydrodynamic limit. However, if we use the diffusive scaling with times proportional to ϵ^{-2} , then the effect of the drift survives in the limit and in fact the hydrodynamic equation for the WASEP is the viscous Burgers equation (2.1). There is also a Central Limit Theorem which describes the small deviations from the hydrodynamic limit as an Ornstein-Uhlenbeck process and thus making rigorous, in this context, the *fluctuating hydrodynamic theory*. We will then explain how a nonlinear SPDE, i.e. the stochastic Burgers equation with viscosity and with additive noise, (4.1), arises as a particular scaling limit of the WASEP. This scaling is obtained by looking at much longer times (proportional to ϵ^{-4}) and very special initial data.

4.1. The Hydrodynamic Limit for WASEP. To understand the scalings let us go back to Section 2 and to (2.1). Let $\rho = \rho(x, t; \epsilon)$, $x \in \mathbb{R}$, $D = 1$ and $v = \epsilon$, then (2.1) becomes

$$(4.5) \quad \frac{\partial \rho}{\partial t} + \epsilon \cdot \frac{\partial}{\partial x} \rho(1 - \rho) = \frac{1}{2} \frac{\partial^2}{\partial x^2} \rho.$$

As in Section 2 we first look at (4.5) as the microscopic system and perform the HSL. We thus set $\rho(x, 0; \epsilon) = \rho_0(\epsilon x)$ and, in contrast to (2.3), we scale times diffusively. We define

$$(4.6) \quad \rho(r, \tau) := \rho(\epsilon^{-1}r, \epsilon^{-2}\tau; \epsilon).$$

As the notation suggests, the rescaled density $\rho(r, \tau)$ is independent of ϵ and, moreover, it solves the viscous Burgers equation

$$(4.7) \quad \frac{\partial \rho}{\partial \tau} + \frac{\partial}{\partial r} \rho(1 - \rho) = \frac{1}{2} \frac{\partial^2}{\partial r^2} \rho,$$

with initial condition $\rho_0(r)$, which is therefore the hydrodynamic equation associated to (4.5).

The same conclusions carry through to the WASEP. We take as usual $\mu^{(\epsilon)}$ to be the product measure on $\{0, 1\}^{\mathbb{Z}}$ with $\mu^{(\epsilon)}(\eta(x)) = \rho_0(\epsilon x)$, $\rho_0 \in C^2(\mathbb{R})$, $0 \leq \rho_0(r) \leq 1$ for all $r \in \mathbb{R}$. We call P_ν the law of the WASEP starting from a probability ν .

THEOREM 4.1. *For any $\tau \geq 0$, any $\phi \in C_0^\infty(\mathbb{R})$ and any $\delta > 0$ we have*

$$(4.8) \quad \lim_{\epsilon \rightarrow 0} P_{\mu^{(\epsilon)}} \left(\left| \epsilon \sum_{x \in \mathbb{Z}} \eta(x, \epsilon^{-2}\tau) \phi(\epsilon x) - \int_{\mathbb{R}} \rho(r, \tau) \phi(r) dr \right| \leq \delta \right) = 1,$$

where $\rho \in C^{1,2}(\mathbb{R}^+, \mathbb{R})$ is the unique (classical) solution of (4.7) with initial condition ρ_0 .

This result says that the hydrodynamic limit of the WASEP, on the proper diffusive time scale, is the Burgers equation with viscosity. For a proof see [31] and

[53]. See also [70], where the large deviations from the hydrodynamic limit are established (in macroscopically finite volumes). Theorem 4.1 holds in any dimension.

4.2. The Normal Fluctuations. Let us now look at the fluctuations around the limit behavior given by (4.7). To do this we consider the distribution valued process $X^\epsilon \equiv \{X_\tau^\epsilon, \tau \geq 0\}$

$$(4.9) \quad \langle X_\tau^\epsilon, \phi \rangle \equiv \epsilon^{1/2} \sum_{x \in \mathbb{Z}} [\eta(x, \epsilon^{-2}\tau) - \rho(\epsilon x, \tau)] \phi(\epsilon x),$$

where ϕ is a test function in \mathcal{S} . For any $\tau \in \mathbb{R}^+$, X_τ^ϵ is an element in \mathcal{S}' (Schwarz distributions) and X^ϵ is an element in $D([0, \infty); \mathcal{S}')$, the Skorohod space of distribution valued functions: we denote by $\langle \cdot, \cdot \rangle$ the duality between \mathcal{S} and \mathcal{S}' .

If we assume that $\mu^{(\epsilon)}$ is the product measure considered in Theorem 4.1, then $X_0^\epsilon \in \mathcal{S}'$ converges weakly to a \mathcal{S}' -valued process X_0 , where $\langle X_0, \phi \rangle$ is a centered Gaussian variable with variance $\int \rho_0(r)(1 - \rho_0(r))\phi(r)^2 dr$ (i.e. X_0 is a *spatially modulated white noise*).

We have the following result (see [31],[35])

THEOREM 4.2. $X^\epsilon \in D([0, \infty); \mathcal{S}')$ converges weakly to $X \in C([0, \infty); \mathcal{S}')$, where X is the unique solution in $C([0, \infty); \mathcal{S}')$ of

$$(4.10) \quad \langle X_t, \phi \rangle = \langle X_0, \phi \rangle + \int_0^t \langle Y_s, \left(\frac{1}{2}\Delta + (2\rho_s - 1)\nabla \right) \phi \rangle ds + \langle M_t, \phi \rangle,$$

for all $\phi \in \mathcal{S}$ and $t \geq 0$. $\langle M_t, \phi \rangle$ is the continuous martingale with independent increments and quadratic variation given by

$$(4.11) \quad \int_0^t ds \langle \rho(\cdot, s)(1 - \rho(\cdot, s)), \frac{\partial}{\partial r} \phi \rangle^2.$$

In other words, the fluctuations of the particle distribution around its deterministic limit is given by an Ornstein–Uhlenbeck process in which the drift term is the linearization of the Burgers equation: X_t satisfies the linear SPDE

$$(4.12) \quad \frac{\partial}{\partial t} X_t = \frac{1}{2}\Delta X_t - \nabla[(2\rho(\cdot, t) - 1)X_t] + \nabla \left(\sqrt{\rho(\cdot, t)(1 - \rho(\cdot, t))} \dot{W}_t \right).$$

Theorem 4.2 can be extended to $d > 1$, but some care has to be taken in defining (4.9): the fluctuation field in general should be defined by subtracting the mean of $\eta(x, \epsilon^{-2}\tau)$ and not its hydrodynamic value, which will differ by terms of $O(\epsilon)$ from the mean, since, for example as in the proof of Theorem 4.1, we are replacing a discrete version of (4.7) on a grid of spacing ϵ with its continuum limit. This problem is not seen in $d = 1$, where this difference is masked by the magnitude of the fluctuations which are of $O(\sqrt{\epsilon})$, but, for general d , the fluctuations will be of order $\epsilon^{d/2}$, so the difference between the mean of the process its hydrodynamic limit is important for $d \geq 2$.

The result analogous to Theorem 4.2 for the ϵ -dependent SPDE problem (4.1) is showing that

$$\epsilon^{-1/2} \left[\rho^{(\epsilon)}(\cdot, t) - \rho(\cdot, t) \right]$$

converges weakly to to X_t , a fact that can be easily verified at a formal level.

We want to stress that the appearance of a *linear* SPDE as the law of the fluctuations is no surprise and the structure of Theorem 4.2 is expected to hold in great generality: the field of fluctuations is, in the limit, a diffusion with drift given by the linearization of the hydrodynamic equation. There is however, for the WASEP, a scaling limit in which a suitable fluctuation field converges to a *nonlinear* SPDE and we describe this next.

4.3. Nonlinear Fluctuations for WASEP: the Space–Time Scaling. The next results are restricted to $d = 1$. Let us start by explaining the limit that we will consider. Theorem 4.1 and Theorem 4.2 show that (4.1) is correct up to first order in ϵ in the hydrodynamic scaling. We will then use (4.1) (beyond the regime where its validity has been proved) for some formal manipulations aimed at understanding what could be a scaling in which both noise and nonlinearity appear together. In particular, for simplicity, we will replace \dot{W}_t^ϵ by \dot{W}_t . \dot{W}_t denotes different processes which coincide in law.

Let $m_t^{(\epsilon)}(r) = \rho^{(\epsilon)}(\epsilon^{-1}r, \epsilon^{-2}t)$, this corresponds to looking at the empirical density on the longer space-time scale $t \sim \epsilon^{-4}$, $x \sim \epsilon^{-2}$: recall that the forcing term is $\sim \epsilon$, so that this new scaling, in spite of being still diffusive, is different from the previous one. Using the scaling properties of white noise, we get from (4.1)

$$(4.13) \quad \frac{\partial m_t^{(\epsilon)}}{\partial t} = \frac{1}{2} \Delta m_t^{(\epsilon)} - \partial_r \left[\epsilon^{-1} m_t^{(\epsilon)} (1 - m_t^{(\epsilon)}) + \epsilon \sqrt{m_t^{(\epsilon)} (1 - m_t^{(\epsilon)})} \dot{W}_t \right],$$

which gives a diverging current when $\epsilon \rightarrow 0$. In order to make this term finite we shall consider what has been called, in the context of the derivation of the Navier-Stokes equation, the *incompressible limit*, i.e. we analyze a small perturbation of the global equilibrium of constant density. We assume that $m^{(\epsilon)}$ has the form

$$(4.14) \quad m_t^{(\epsilon)} = \frac{1}{2} \left[1 + \epsilon u_t^{(\epsilon)} \right],$$

and that $u_t^{(\epsilon)}$ has a limit as ϵ tends to zero, i.e. we are considering a perturbation of the same order as the noise in (4.13), for this reason we *may be able* to see both the nonlinearity and the random force in the scaling limit. Recall that at the end of Section 3 we have considered initial profiles of the form (4.14): they were used in the ASEP in $d \geq 3$ to study the Navier Stokes correction, [38], in a diffusive scaling limit. In that case in fact we do not see stochastic effects, but an extra viscosity shows up.

With the choice (4.14) the diverging term in (4.13) disappears and, in the limit $\epsilon \rightarrow 0$, the equation that we obtain by substituting (4.14) into (4.13) formally converges to

$$(4.15) \quad \frac{\partial u_t}{\partial t} = \frac{1}{2} \Delta u_t + \frac{1}{2} \nabla(u_t^2) + \nabla \dot{W}_t.$$

The stochastic Burgers equation (4.15) should thus describe the first order correction to a global equilibrium profile.

In this heuristic discussion we are always *assuming* that the form (4.14), which can be assumed at $t = 0$, will be stable under the time evolution, but this is to a certain extent the real issue. In other words we are asking what happens to a perturbation of order ϵ (on the space scale ϵ^{-2}) after a time ϵ^{-4} .

At this point, in conformity with the existing literature, we change slightly the definition of the WASEP process. From now on in this section the WASEP process will be generated by

$$(4.16) \quad L_\epsilon f(\eta) = \sum_{x \in \mathbb{Z}} \left\{ \left(\frac{1}{2} + \epsilon \right) [f(\eta^{x-1,x}) - f(\eta)] + \frac{1}{2} [f(\eta^{x,x+1}) - f(\eta)] \right\},$$

where f is, as before, a local function of the configuration η . The difference with the process introduced before, generated by (3.2) with the choices (3.22) and (4.4), is that here the bias is toward the left, i.e. $q > p$, and the parameter ϵ which appears in (4.4) should be replaced by $\epsilon/(1+\epsilon) = \epsilon + O(\epsilon^2)$ and the time should be speeded up of a factor $1+\epsilon$, i.e. the generator (3.22) should be multiplied by $1+\epsilon$. The change of time as well as the small change in the size of the asymmetry do not change the statements of Theorem 4.1 and Theorem 4.2, apart for the different sign in the drift term of the Burgers equation, i.e. the sign of the second term in the right-hand side of (4.7) should be exchanged (and analogous change in Theorem 4.2). Some of the result we are going to present now are more refined and the results would have to be restated in a slightly different way for the process with jump rates given by (4.4), but no substantial change would be needed. The bias toward the left has been introduced because it gives a more direct picture of the growing interface process that will be introduced in the next section.

The heuristic considerations we made above suggest that in order to derive the stochastic Burgers equation from the WASEP, we should consider the fluctuation field

$$(4.17) \quad \langle X_t^\epsilon, \phi \rangle = \epsilon \sum_{x \in \mathbb{Z}} \phi(\epsilon^2 x) [2\eta(x, \epsilon^{-4}t) - 1],$$

where $\{\eta\}$ is the WASEP. The initial condition $\mu^{(\epsilon)}$ is the product measure on $\{0,1\}^{\mathbb{Z}}$ with marginals

$$(4.18) \quad \mu^{(\epsilon)}(\eta(x)) = H \left(\frac{1}{2} + \epsilon^{-1} [h(\epsilon^2 x) - h(\epsilon^2(x-1))] \right),$$

$$H(r) = \begin{cases} 1 & \text{if } r > 1 \\ r & \text{if } r \in [0, 1] \\ 0 & \text{if } r < 0, \end{cases}$$

with $h = h(r)$ a α -Hölder continuous function ($\alpha > 1/2$) on \mathbb{R} which satisfies the condition that there is $a > 0$ such that, for every $r \in \mathbb{R}$, $|h(r)| \leq a(1 + |r|)$. In [5] a much larger class of initial data is considered, but here we will restrict ourselves to this case. It is by no means obvious how to give a meaning to (4.15) and, as a matter of fact, we will be selecting one of the possible *regularizations* of (4.15). We therefore first discuss this point.

4.4. The Burgers Equation with Conservative Noise. We consider the space $C^0(\mathbb{R})$ of continuous functions on the real line equipped with the topology of uniform convergence over compact sets. As test functions we use the space \mathcal{D} , i.e. the function in $C_0^\infty(\mathbb{R})$ with the inductive limit topology; its strong topological dual \mathcal{D}' is the space of distributions over \mathbb{R} . We denote by $\langle \cdot, \cdot \rangle$ the duality between the test functions and the distributions. In this notation the space-time W_t in (4.3) is the distribution valued centered Gaussian process such that

$$(4.19) \quad E(\langle W_t, \phi_1 \rangle \langle W_s, \phi_2 \rangle) = t \wedge s \langle \phi_1, \phi_2 \rangle,$$

where $\phi_1, \phi_2 \in \mathcal{D}$, are test function, (\cdot, \cdot) is the scalar product in $L^2(\mathbb{R})$ and $a \wedge b := \min\{a, b\}$. We denote by \mathcal{A}_t the natural filtration of W_t and denote by P the law of W_t .

For the initial datum of the stochastic Burgers equation we take a random field with trajectories u_0 in \mathcal{D}' of the form that we now explain. We specify a random field \mathcal{P}_0 with trajectories $h_0 \in C(\mathbb{R})$ which is independent of P and satisfies the condition: for all $p > 0$ there exists a such that

$$(4.20) \quad \sup_{r \in \mathbb{R}} e^{-a|r|} \mathcal{E}_0 \left(e^{-ph_0(r)} \right) < \infty.$$

For $\phi \in \mathcal{D}$ we then define

$$(4.21) \quad \langle u_0, \phi \rangle = - \int h_0(r) \phi'(r) dr$$

To summarize, u_0 is the derivative, in distributional sense, of a continuous trajectory $h_0 = h_0(r)$ which has for every $r \in \mathbb{R}$ an exponential moment that grows at most exponentially in r . One can obviously take h_0 to be a deterministic function with at most linear growth.

We will characterize the solution of the stochastic Burgers equation through a limiting procedure. Accordingly we introduce a regularized version of the cylindrical Wiener process which will define a family of approximating problems.

Let $\varphi \in C_0^\infty(\mathbb{R})$ be as in (4.2). Introduce, for $\kappa > 0$, the approximate identity $\delta_r^\kappa(r') = \varphi_{1/\kappa}(r-r') = \kappa\varphi(\kappa(r-r'))$ and here we use the notation $W_t^\kappa(r) = \langle W_t, \delta_r^\kappa \rangle$; its covariance is then

$$(4.22) \quad E \left(W_t^{1/\kappa}(r) W_s^{1/\kappa}(r') \right) = (t \wedge s) C_\kappa(r-r'), \quad C_\kappa(r) = \int \delta^\kappa(r-r') \delta^\kappa(r') dr'.$$

We then write the stochastic Burgers equation with this regularized Wiener process against test functions as

$$(4.23) \quad \langle u_t^\kappa, \phi \rangle = \langle u_0, \phi \rangle + \int_0^t \frac{1}{2} \{ \langle u_s^\kappa, \phi'' \rangle + \langle (u_s^\kappa)^2, \phi' \rangle \} ds - \langle W_t^{1/\kappa}, \phi' \rangle,$$

which is formally obtained from (4.15) by integrating by parts. The limit $\kappa \rightarrow \infty$ is now taken according to the following strategy. When the cutoff κ is finite $W_t^{1/\kappa}(r)$ is smooth so that (4.23) makes sense in the space of continuous functions; we thus obtain a processes u^κ in $C((0, T]; C^0(\mathbb{R}))$. Since a limiting process will not be continuous in space we have no hope to get a convergent sequence in the topology of $C((0, T]; C^0(\mathbb{R}))$; however $\{u^\kappa\}_{\kappa>0}$ does form a weak convergent family as $\kappa \rightarrow \infty$ in the topology of $C([0, T]; \mathcal{D}')$.

In fact we have the following

PROPOSITION 4.1. *Under the above assumptions on the initial condition u_0 , for all $\kappa > 0$ there exists a unique process $u_t^\kappa = u_t^\kappa(r)$ in $C((0, \infty); C^0(\mathbb{R}))$, adapted to the filtration \mathcal{A}_t , which solves almost surely (4.23) for all $\phi \in \mathcal{D}$ and every $t > 0$. Moreover the family $\{u^\kappa\}_{\kappa > 0}$, $u_t^\kappa \in C([0, T]; \mathcal{D}')$, converges weakly as $\kappa \rightarrow \infty$ to a limit process denoted by u .*

Although this result characterizes uniquely the solution of the stochastic Burgers equation (4.15) through the approximating problems (4.23) it is not completely satisfactory since it avoids the issue of showing that u satisfies a limit equation. Below we will give an alternative definition of the process u . But let us give first the result linking the WASEP with the stochastic Burgers equation.

4.5. WASEP and the Stochastic Burgers Equation. Let us go back to the particle system and consider (4.17). We regard $X^\epsilon = (X_t)_{t \in [0, T]}$ as a random element in $D([0, T]; \mathcal{D}')$. In the scaling limit $\epsilon \rightarrow 0$ the distribution of the fluctuation field (4.17) is determined by the Burgers equation with random noise on the current, namely

THEOREM 4.3. *The family $\{X^\epsilon\}_{\epsilon > 0}$ is weakly convergent in $D([0, T]; \mathcal{D}')$ as $\epsilon \rightarrow 0$. Furthermore the weak limit is concentrated on $C([0, T]; \mathcal{D}')$ and coincides with the process u constructed in Proposition 4.1, with $\langle u_0, \phi \rangle$ a Gaussian variable with mean $-\langle m, \phi' \rangle$ and variance (ϕ, ϕ) , for any $\phi \in \mathcal{D}$.*

The initial condition is easily derived from (4.18) and it simply says that $u_0 = \nabla m + W$, where W is a white noise. We have thus found a scaling in which the fluctuations are nonlinear. We are now going to say a few words about the derivation of this result.

4.6. The Cole–Hopf Transform and the Stochastic Heat Equation. The proof of Theorem 4.3 follows a strategy parallel to the one in [53]. The heuristic observation is simply that if u_t is a solution of the stochastic Burgers equation (4.15) and h_t is a function such that $u_t = \nabla h_t$, then the process $\theta_t := \exp\{-h_t\}$ solves the linear SPDE

$$(4.24) \quad d\theta_t = \frac{1}{2} \Delta \theta_t dt - \theta_t dW_t,$$

with the stochastic differential interpreted in the Ito sense. Equation (4.24) is usually referred to as the *stochastic heat equation*, see [98],[84],[4] for existence, uniqueness and other properties of the equation, with the stochastic differential interpreted in the Ito sense. Here we just recall that if we set

$$(4.25) \quad C_+^0(\mathbb{R}) = \{f \in C^0(\mathbb{R}) : f(r) > 0 \text{ for all } r \in \mathbb{R}\}$$

and $\theta_0 \in C_+^0$ a.s., then $\theta \in C([0, \infty); C_+^0(\mathbb{R}))$ a.s. (this has been proved by C. Müller [84]). We can then define the logarithm of θ and in [5] it is proven that the process u given by Proposition 4.1 is in law the same as $-\nabla \log \theta$, thus providing a better characterization of the limit behavior of the particle system. The choice of the initial condition for (4.24) to match (4.21) is straightforward.

The idea is now to perform a transformation similar to the Cole–Hopf on the particle system, as it is done in [53]. The resulting particle model is simpler than the original one and one can prove convergence of the new process to the solution of (4.24). Theorem 4.3 is then obtained by performing the inverse Cole–Hopf transform.

5. Driven Surfaces and Fluctuations

We describe here one particular model out of an extremely wide field, explored by physicists and recently also by mathematicians. There is however, up to now, very little mathematical understanding of the basic physical results, in spite of the very many connections with active fields of probability, like random walks in random environments and polymer problems. An interesting review of the physical ideas is in [72].

5.1. Interface Models of Solid on Solid Type. The WASEP can be easily mapped into a one-dimensional interface growth model and we will explain how the results obtained in the context of the stochastic Burgers equation map into results for this interface model.

The microscopic interface model has state space

$$(5.1) \quad \hat{\Omega} = \{ \zeta \in \mathbb{Z}^{\mathbb{Z}} : |\zeta(x+1) - \zeta(x)| = 1 \text{ for all } x \in \mathbb{Z} \}.$$

Its time evolution is governed by the jump Markov process generated by

$$(5.2) \quad \hat{L}_{\epsilon} f(\zeta) = \sum_{x \in \mathbb{Z}} \{ c_{\epsilon}^{+}(x, \zeta) [f(\zeta + 2\delta_x) - f(\zeta)] + c_{\epsilon}^{-}(x, \zeta) [f(\zeta - 2\delta_x) - f(\zeta)] \},$$

where f is a cylindrical function on $\hat{\Omega}$, $\delta_x \in \mathbb{Z}^{\mathbb{Z}}$ is defined by $\delta_x(y) = \delta_{x,y}$ (the Kronecker symbol) and

$$c_{\epsilon}^{+}(x, \zeta) = \begin{cases} 1/2 + \epsilon & \text{if } \zeta(x+1) - 2\zeta(x) + \zeta(x-1) = 2 \\ 0 & \text{otherwise,} \end{cases}$$

$$c_{\epsilon}^{-}(x, \zeta) = \begin{cases} 1/2 - \epsilon & \text{if } \zeta(x+1) - 2\zeta(x) + \zeta(x-1) = -2 \\ 0 & \text{otherwise.} \end{cases}$$

The dynamics is easily explained: view $\zeta(x)$ as the height of the interface at the site x . There is the constraint that, moving along x , the interface is never flat and changes its height by ± 1 at each step. The interface evolves according to a Poisson process and the only changes allowed are the ones which make a local minimum into a local maximum or vice versa. This is performed with a slight bias (ϵ) toward increasing the height, that is more minima become maxima than vice versa. Notice that the dynamics preserves the single step constraint. This model goes under the name of restricted solid on solid process (SOS). We shall consider ζ as a continuous function by linear interpolation on its value on the lattice \mathbb{Z} .

To give the relation between the SOS process and the WASEP we put a tag on the particle that at time zero is the closest to the origin on the positive half-axis. Let x_t^0 be the position of the tagged particle under the time evolution. It

is straightforward to verify that we can construct a version of the SOS process, starting from the WASEP, essentially by summing from the tagged particle

$$(5.3) \quad \zeta_t(x) = \begin{cases} \sum_{x_i^0 < y \leq x} (2\eta_t(y) - 1) - x_t^0 & \text{if } x > x_t^0 \\ -\sum_{x < y \leq x_t^0} (2\eta_t(y) - 1) - x_t^0 & \text{if } x < x_t^0 \\ -x_t^0 & \text{if } x = x_t^0 \end{cases}$$

where η_t is the WASEP.

In the reverse direction, given ζ_t a realization of the SOS process, we obtain a version of the WASEP by looking at the increments, i.e.

$$\eta_t(x) = \frac{(\zeta_t(x) - \zeta_t(x-1)) + 1}{2}.$$

5.2. Continuum Interfaces. We first translate (briefly and informally) Theorem 4.1 and Theorem 4.2 into the interface language (the precise statements can be easily recovered).

The law of large numbers this time involves $h_t^\epsilon(r) \equiv \epsilon \zeta_{t\epsilon^{-2}}([\epsilon^{-1}r])$, an element in $D(\mathbb{R}^+; C^0(\mathbb{R}))$. It is then an easy corollary of Theorem 4.1 to show that $h_t^\epsilon \in C^0(\mathbb{R})$ converges in probability to $h \in C^{1,2}(\mathbb{R}^+, \mathbb{R})$, the unique classical solution of

$$(5.4) \quad \frac{\partial h_t}{\partial t} = \frac{1}{2} \Delta h_t + \frac{1}{2} [1 - (\nabla h_t)^2]$$

with initial condition $h_0 = h$.

In a similar way the normal fluctuations consist in looking at

$$(5.5) \quad X_t^\epsilon(r) = \epsilon^{-1/2} [\epsilon \zeta_{t\epsilon^{-2}}([\epsilon^{-1}r]) - h_t(r)]$$

and the weak limit X of $X^\epsilon \in D(\mathbb{R}^+; C^0(\mathbb{R}))$ solves the linear SPDE

$$(5.6) \quad \frac{\partial X_t}{\partial t} = \frac{1}{2} \Delta X_t - \nabla h_t \nabla X_t + \sqrt{1 - (\nabla h_t)^2} \dot{W}_t.$$

We now pass to the nonlinear fluctuations. These are of particular interest because they give rise to the celebrated Kardar–Parisi–Zhang (KPZ) equation [67]. In this case the object of interest turns out to be

$$(5.7) \quad Z_t^\epsilon(r) = \sqrt{\epsilon} (\zeta_{t\epsilon^{-2}}(\epsilon^{-2}r) - v_\epsilon t),$$

where $v_\epsilon = (1/2)\epsilon^{-3} - (1/24)\epsilon^{-1}$ and we regard Z^ϵ as a random element in $D([0, T]; C(\mathbb{R}))$. The statement is then that the family $\{Z^\epsilon\}_{\epsilon>0}$ is weakly convergent as $\epsilon \rightarrow 0$ to the limit $Z = -\log \theta$, where θ is the solution of the stochastic heat equation.

Also in this case we can write a mollified KPZ equation against test functions as

$$(5.8) \quad \langle h_t^\kappa, \phi \rangle = \langle h_0, \phi \rangle + \int_0^t \frac{1}{2} \{ \langle h_s^\kappa, \phi'' \rangle - \langle [(\nabla h_s^\kappa)^2 - C_\kappa(0)], \phi \rangle \} ds + W_t^{1/\kappa}(\phi)$$

where $\phi \in \mathcal{D}(\mathbb{R})$. The weak limit of h^κ coincides with $-\log \theta$ (this is the analog of Proposition 4.1).

We notice that, to a certain extent, the result on the KPZ equation is more complex than the one for Burgers. First of all there is the term $C_\kappa(0) \sim \kappa^{-1}$, which corresponds to a *renormalization* of the nonlinearity: this term disappeared in the Burgers case, due to the extra space gradient. Moreover to define the fluctuation field (5.7) we had to subtract not only the average speed of the interface, that is

the term of order ϵ^{-3} in v_ϵ , but also an extra term of order ϵ^{-1} . While the origin of the first term is easily understood, since the interface has a drift proportional to ϵ acting for a time $O(\epsilon^{-4})$, the fact that there is no $O(\epsilon^{-2})$ term and, in particular, the value of the constant in front of the ϵ^{-1} term (1/24) are highly non obvious. Using the result we just (informally) stated, it has been possible to prove [6] that the value of *sample Lyapunov exponent* ([83],[15]) of the stochastic heat equation (4.24) is 1/24. This result can be found in the physical literature ([72],[66]), where the quenched free energies of the directed polymer and other interface models are computed via a non rigorous *replica computation*: the stochastic heat equation can be interpreted as the partition function of a one dimensional directed polymer in random environment (t in this case becomes the size of the system) and the sample Lyapunov exponent is then the *quenched free energy* of the system.

The KPZ equation is expected to describe the large scale fluctuations of a very general class of driven interfaces. These fluctuations are *anomalous*: in particular they are not of Gaussian type and this phenomenon goes under the name of *kinetic roughening*. This behavior is intimately related to the appearance of superdiffusive behavior of a random walk in random environment (the directed polymer mentioned above: see e.g. [72] and [66]). The physical phenomena are very dependent on the dimensionality of the space and the one-dimensional case is definitely different from the higher dimensional cases (from the physical viewpoint, the one-dimensional and two-dimensional cases are the most complex ones). The results that we have just reviewed, besides establishing the validity of the physicists' picture in one particular case, provide a rigorous mathematical framework (first of all a precise definition of *the solution to the KPZ equation*) within which one could try to formulate and eventually prove the physical statements.

PART II

REVERSIBLE DYNAMICAL SYSTEMS: SYMMETRIC MODELS WITH LONG RANGE INTERACTIONS

6. Ising Models with Kac Potentials: Glauber and Kawasaki Dynamics

The systems considered so far have trivial translation invariant stationary states which are just the Bernoulli measures (indexed by the density parameter $\rho \in [0, 1]$) and their linear combinations. We now switch to systems whose stationary states have a much richer structure, in particular we will be concerned with systems which can undergo phase transitions. In the cases we are going to consider it is convenient to replace the particle density $\rho \in [0, 1]$ by the magnetization $m = 2\rho - 1 \in [-1, +1]$; the analogous change is also done at the microscopic level, where from now on the occupation value $\eta(x) = 0, 1$ at $x \in \mathbb{Z}^d$ is replaced by the spin variable $\sigma(x) = \pm 1$. We will look both at systems with exchange dynamics, like all the particle models considered in Part I, and at systems with spin-flip dynamics, i.e. where the elementary event is the change of the value (flip) of a single spin. We will briefly consider also evolutions where both mechanisms are present.

The easiest examples of these dynamics are given by the following two *non interacting systems*: the first one is the independent spin-flip process, i.e. a process $\{\sigma_t\}_{t \in \mathbb{R}^+}$, $\sigma_t \in \{-1, +1\}^{\mathbb{Z}^d}$, which is a collection of independent processes

$\{\sigma_t(x)\}_{t \geq 0}$ indexed by $x \in \mathbb{Z}^d$, and for each x the only possible transition, which happens at exponentially distributed random times (with mean 1), is the spin-flip $\sigma(x) \rightarrow -\sigma(x)$. It is obvious that this dynamics has only one invariant measure: the Bernoulli measure with average 0. The second one is the Symmetric Simple Exclusion Process, that is the process generated by (3.22) with $p = q = 1/2$ and $\sigma = 2\eta - 1$. For this model, as we have already pointed out, the combinations of Bernoulli measures with average $m \in [-1, 1]$ are invariant (they are actually the only invariant measures).

More complex systems are naturally built by choosing the rate of transition at time t dependent on the configuration at time t^- (indeed the exclusion process does have a dependence, given by the exclusion rule, but one can view it as an independent process on the bonds). We will choose configuration dependent rates, but we will do it in such a way that the *Gibbs measures*, the equilibrium measures according to statistical mechanics, are stationary for our systems. We thus know from the beginning (at least some of) the invariant measures, and this is a real advantage since often the existence and characterization of the stationary states is a very hard problem. A Gibbs measure is usually specified by an *interaction potential*, related to the transition rates of the dynamics.

The structure of the Gibbs states, which is the main object of study in equilibrium statistical mechanics, may indeed be very complex. In the simpler case of the Ising system that we consider here, if the temperature is sufficiently large or the *interaction* weak, there is a unique Gibbs measure (and a unique invariant measure for the spin flip dynamics as well) which is not too different from the Bernoulli measure with zero average. But at low temperatures and 0 magnetic field, i.e. in a *phase transition situation*, this will no longer be true and there will be (at least) two stationary Gibbs measures with non zero average magnetization $\pm m$. Correspondingly the spin flip evolution has two (or more) extremal, invariant measures. In the case of the exchange dynamics, the total magnetization is invariant and consequently, contrary to the case of the spin flip dynamics, any Gibbs measure, irrespectively of the value of the magnetic field, is invariant. Thus, analogously to the case of the symmetric SEP, we have a family of extremal invariant measures (the Gibbs measures corresponding to all the possible values of the magnetic field) which will be indexed by the magnetization m . If there is no phase transition, m takes all the values in the interval $[-1, 1]$, but when there is a phase transition there is an interval $I \subset [-1, +1]$ of values of m that are forbidden, namely for no value of the magnetic field there is an extremal, translation invariant state with magnetization in I . As one can easily guess, this has dramatic consequences on the macroscopic behavior of the system leading to phenomena like *spinodal decomposition*, free boundary and Stefan problems, interface dynamics, formation of mushy regions, metastable behavior and critical fluctuations. Here we will concentrate on the spinodal decomposition and the critical fluctuations.

The spinodal decomposition is a phenomenon that appears when the temperature of a system is suddenly lowered (*quenched*) from above to below its *critical* value (namely the temperature where there is a bifurcation from absence to presence of phase transition). The magnetization, which during the (very rapid) quenching process does not change significantly, is chosen so that it lies, at the final temperature, in the forbidden interval I . In the interesting cases, the state is then stationary, but unstable for the deterministic macroscopic equation that describes the system. The stochastic corrections to this equation have then the fundamental

role of inducing the escape from such an unstable stationary state with the formation of the pure phases in a non trivial, random spatial pattern. The critical fluctuations refer to the anomalously large magnetization fluctuations that occur at the critical temperature and they are described by nonlinear SPDE's, like the Cahn–Allen equation with additive white noise.

6.1. Gibbs Measures. We use the following notation: for any $\Lambda \subset \mathbb{Z}^d$, we denote by σ_Λ a configuration in Λ , i.e. an element in $\{-1, +1\}^\Lambda$. When $\Lambda = \mathbb{Z}^d$, we simply write σ . We call $\Omega = \{-1, +1\}^{\mathbb{Z}^d}$ and equip it with the product topology.

Since the analysis of models with phase transitions presents great difficulties in general, we will restrict drastically the class of systems by considering Ising systems with Kac potentials. For any value of the Kac parameter $\gamma > 0$, we define the *Gibbs measure* as a probability measure which can be formally written as

$$(6.1) \quad \frac{e^{-\beta H_\gamma(\sigma)}}{\sum_{\sigma \in \Omega} e^{-\beta H_\gamma(\sigma)}},$$

$\beta > 0$ is the inverse temperature and

$$(6.2) \quad H_\gamma(\sigma) = \frac{1}{2} \sum_{x,y \in \mathbb{Z}^d} J_\gamma(x,y) \sigma(x) \sigma(y) + h \sum_{x \in \mathbb{Z}^d} \sigma(x),$$

where $h \in \mathbb{R}$ is the external magnetic field and $J_\gamma(x,y)$ the coupling strength:

$$(6.3) \quad J_\gamma(x,y) = \gamma^d J(\gamma(x-y)).$$

We suppose that $J \in C^\infty(\mathbb{R}^d)$ is a non negative function supported in the unit ball, with $\int_{\mathbb{R}^d} J(r) dr = 1$. We are therefore looking at a system with interaction range γ^{-1} and interaction strength γ^d , we will take γ small. Clearly neither (6.1) nor (6.2) make sense, since they contain unbounded terms.

There are various ways to define properly *Gibbs measures* (for a large amount of information on this one can look at [55],[37]) and we will choose to define a Gibbs measure by giving its conditional expectation at each site. After introducing the molecular magnetic field $h_\gamma(x; \sigma_\Delta)$ at $x \notin \Delta$ due to the spin configuration σ_Δ in Δ as

$$(6.4) \quad h_\gamma(x; \sigma_\Delta) = \sum_{y \in \Delta} J_\gamma(x,y) \sigma(y),$$

we define a Gibbs measure $\mu_{\beta,h,\gamma}$ as a probability on the Borel sets of Ω such that for any x the conditional probability that $\sigma(x) = \bar{\sigma} \in \{-1, +1\}$, given the configuration in x^c (short-hand notation for $\{x\}^c$) is

$$(6.5) \quad \mu_{\beta,h,\gamma}(\sigma(x) = \bar{\sigma} | \sigma_{x^c}) = \frac{e^{-\beta \bar{\sigma}[h_\gamma(x; \sigma_{x^c}) + h]}}{e^{-\beta[h_\gamma(x; \sigma_{x^c}) + h]} + e^{\beta[h_\gamma(x; \sigma_{x^c}) + h]}}$$

almost surely w.r.t. $\mu_{\beta,h,\gamma}(d\sigma_{x^c})$ (which is the marginal of the measure $\mu_{\beta,h,\gamma}$ on $\{-1, 1\}^{x^c}$). The conditioning is on the event $\{\sigma' : \sigma'(y) = \sigma_{x^c}(y) \text{ for all } y \in x^c\}$.

The set of Gibbs measures is non empty, because Ω is compact, and it may contain more than one element, in which case we say that the system undergoes a phase transition. By general theorems on Ising ferromagnetic systems with finite range interactions, for any $\gamma > 0$ there is a unique Gibbs measure, $\mu_{\beta,h,\gamma}$, for any h in $d = 1$ and for $h \neq 0$ in $d \geq 2$. Moreover([10],[18],[11]) in $d \geq 2$ and $h = 0$ there is a unique Gibbs measure if $\beta < 1$ and γ sufficiently small, while if $\beta > 1$, and γ

sufficiently small there are two different, extremal, invariant Gibbs measures, $\mu_{\beta,\gamma}^{\pm}$. They are translationally invariant and any other translationally invariant Gibbs measure is a convex combination of $\mu_{\beta,\gamma}^{\pm}$. For simplicity, even if uniqueness does not hold, we will keep using the notation $\mu_{\beta,h,\gamma}$ to indicate any Gibbs measure with parameters β , h and γ .

6.2. Glauber and Kawasaki Dynamics. We will consider two classes of reversible processes: Glauber (spin–flip) dynamics and Kawasaki (spin–exchange) dynamics. The elementary events in the former are spin–flips (i.e. changes of sign of a single spin, (6.7) below) while, in the Kawasaki case, two spins exchange their position, like in the dynamics presented in Part I. In both cases, we build the dynamics from jump rates which will only depend on the difference of energy before and after an elementary transition. To simplify the exposition, from now on a Gibbs measure is always a *translationally invariant* Gibbs measure and an *extremal* Gibbs measure is a measure ergodic with respect to spatial translation.

The dynamics we will be looking at are Markov processes in $D(\mathbb{R}^+; \Omega)$ generated by self–adjoint operators in $L^2(\Omega; \mu)$, where μ is a Gibbs measure at given β , h and γ . Self–adjointness (of the generator of the dynamics in $L^2(\Omega; \mu)$) is sometimes called *reversibility* and the process *reversible*, see [95]). Self–adjointness trivially implies that μ is invariant for the dynamics. All the details on the construction of these processes are given in [81].

The generator of the particular Glauber dynamics we shall consider here, L_γ^G , acts on the local (cylinder) functions f as

$$(6.6) \quad L_\gamma^G f(\sigma) = \sum_{x \in \mathbb{Z}^d} c_\gamma(x; \sigma) [f(\sigma^x) - f(\sigma)],$$

where

$$(6.7) \quad \sigma^x(y) = \begin{cases} \sigma(y) & \text{if } y \neq x \\ -\sigma(y) & \text{if } y = x, \end{cases}$$

$$(6.8) \quad c_\gamma(x; \sigma) = \frac{e^{-\beta\sigma(x)[h_\gamma(x, \sigma|_{x^c}) + h]}}{e^{-\beta[h_\gamma(x, \sigma|_{x^c}) + h]} + e^{\beta[h_\gamma(x, \sigma|_{x^c}) + h]}}.$$

Observe that the *flip rate* $c_\gamma(x; \sigma)$ in (6.8) is a nonnegative function $\Phi(E)$ of the difference of energy E before and after a spin–flip

$$(6.9) \quad H_{\gamma, \{x\}}(-\sigma|_{\sigma|_{x^c}}) - H_{\gamma, \{x\}}(\sigma|_{\sigma|_{x^c}}) = 2\sigma(x) [h_\gamma(x, \sigma|_{x^c}) + h].$$

The choice (6.8) is not the only possible one: any flip rate defined in terms of a function $\Phi(E)$ (with some regularity conditions) that satisfies the detailed balance condition (see [95], page 161, and [81], Section 4)

$$(6.10) \quad \Phi(E) = \exp(-\beta E)\Phi(-E) \quad \text{for all } E \in \mathbb{R},$$

leads to a reversible process with invariant measure μ . (6.8) is obtained by choosing $\Phi(E) = e^{-\beta E/2}/(e^{\beta E/2} + e^{-\beta E/2})$.

The generator of the Kawasaki dynamics acts on the local functions as

$$(6.11) \quad L_\gamma^K f(\sigma) = \sum_{x,y \in \mathbb{Z}^d} c_\gamma(x, y; \sigma) [f(\sigma^{x,y}) - f(\sigma)].$$

By $\sigma^{x,y} \in \Omega$, as in (3.3), we mean the configuration obtained from σ by exchanging the spins at the sites x and y ,

$$(6.12) \quad \sigma^{x,y}(z) = \begin{cases} \sigma(x) & \text{if } z = y \\ \sigma(y) & \text{if } z = x \\ \sigma(z) & \text{otherwise.} \end{cases}$$

We set

$$(6.13) \quad c_\gamma(x, y; \sigma) = \begin{cases} \Phi(\Delta_{x,y} H_\gamma) & \text{if } |x - y| = 1 \\ 0 & \text{otherwise,} \end{cases}$$

with $\Phi \in C^2$ which satisfies (6.10) and

$$(6.14) \quad \Delta_{x,y} H_\gamma(\sigma) = H_{\gamma, \{x,y\}} (\sigma^{x,y} | \sigma|_{\{x,y\}^c}) - H_{\gamma, \{x,y\}} (\sigma | \sigma|_{\{x,y\}^c}).$$

For convenience we set $\Phi(0) = 1$, which by (6.10) implies $\Phi'(0) = -\beta/2$ and, as it will be explained below, at the hydrodynamic level all that can be observed of Φ is its value and the value of its first derivative at 0. The general case is obviously recovered by a time change.

It is easy to realize that the rates in (6.13) are h independent. Due to the condition (6.10), any measure in the family $\{\mu_{\beta,h,\gamma}\}_{h \in \mathbb{R}}$ is invariant. This is the first important difference between the Glauber and the Kawasaki dynamics. In the Glauber dynamics for each different choice of the parameters β , h and γ (we have fixed once for all J) we have different jump rates. Correspondingly only the Gibbs measures $\mu_{\beta,h,\gamma}$ with the same magnetic field as the one used in the definition of the dynamics are invariant (apart for degenerate cases). For the Kawasaki dynamics, instead, the magnetic field h does not appear in the definition of the rates: this is an immediate consequence of the fact that the magnetization is locally conserved in a spin exchange. As a consequence all the Gibbs measures, no matter what is the value of h , are invariant and we thus have the one parameter family of invariant measures $\{\mu_{\beta,h,\gamma}\}_{h \in \mathbb{R}}$ with the magnetic field h that plays for the spin systems the role that the *chemical potential* has for the particle systems.

The presence of a *first order* phase transition at $h = 0$ for $\beta > 1$ and h sufficiently small means that, for the same value of β and γ , the function $m(h) = E_{\beta,h,\gamma}(\sigma(x))$, well defined for $h \neq 0$ (and in fact analytical and, due to the condition $J \geq 0$, monotonic increasing in $\mathbb{R} \setminus \{0\}$), see [55], [37]), has left limit $-m_{\beta,\gamma}$ and right limit $+m_{\beta,\gamma}$, with $m_{\beta,\gamma} > 0$. Therefore if we call I the interval $(-m_{\beta,\gamma}, +m_{\beta,\gamma})$ (the *forbidden* interval), the equation $m(h) = \bar{m} \in I$ cannot be solved. Any Gibbs measure with average $m \in I$ is a linear combination of the extremal states with magnetization $m \notin I$.

6.3. Macroscopic Limits and Non Local Evolution Equations. We take γ^{-1} as our macroscopic space unit and we are interested in the limit as $\gamma \rightarrow 0$. Let $m_0 \in C^\infty(\mathbb{R}^d; [-1, 1])$ be a macroscopic initial profile which at the spin level is represented by the product measures $\mu^{(\gamma)}$ on Ω with averages

$$(6.15) \quad E_{\mu^{(\gamma)}}(\sigma(x)) = m_0(\gamma x)$$

(just as in Section 3, (3.7) with γ instead of ϵ). Denoting by $P_\nu^{(\gamma, G)}$ and by $P_\nu^{(\gamma, K)}$ the processes starting from ν with generators L_γ^G and, respectively, L_γ^K we have ([26] for Glauber, [78],[58],[59] for Kawasaki¹)

THEOREM 6.1. *For any $\delta > 0$, $\tau \geq 0$ and $\phi \in C_0^\infty(\mathbb{R}^d)$ the following holds.*

(i). Glauber dynamics.

$$(6.16) \quad \lim_{\gamma \rightarrow 0} P_{\mu^{(\gamma)}}^{(\gamma, G)} \left(\left| \gamma^d \sum_{x \in \mathbb{Z}^d} \phi(\gamma x) \sigma(x, \tau) - \int_{\mathbb{R}^d} dr \phi(r) m^{(G)}(r, \tau) \right| < \delta \right) = 1,$$

where $m^{(G)}(r, \tau)$ is the unique solution of

$$(6.17) \quad \frac{\partial m^{(G)}}{\partial \tau}(r, \tau) = -m^{(G)}(r, \tau) + \tanh \left\{ \beta((J * m^{(G)})(r, \tau) + h) \right\},$$

with initial datum m_0 . Here $(J * m^{(G)})(r, \tau)$ denotes the convolution of J and $m^{(G)}(\cdot, \tau)$ evaluated at r .

(ii). Kawasaki dynamics.

$$(6.18) \quad \lim_{\gamma \rightarrow 0} P_{\mu^{(\gamma)}}^{(\gamma, K)} \left(\left| \gamma^d \sum_{x \in \mathbb{Z}^d} \phi(\gamma x) \sigma(x, \gamma^{-2}\tau) - \int_{\mathbb{R}^d} dr \phi(r) m^{(K)}(r, \tau) \right| < \delta \right) = 1,$$

where $m^{(K)}(r, 0) = m_0(r)$ and $m^{(K)}(r, \tau)$ is the unique solution of

$$(6.19) \quad \frac{\partial m^{(K)}}{\partial t}(r, \tau) = \Delta m^{(K)}(r, \tau) - \beta \nabla \left[(1 - m^{(K)}(r, \tau)^2) (\nabla J * m^{(K)})(r, \tau) \right],$$

and $*$ is again the convolution in the spatial variable r .

A more accurate analysis of the process shows that for both dynamics the spins are to leading order (as $\gamma \rightarrow 0$) mutually independent. This is a mean field effect due to the fact that for small γ the direct interaction between two spins is negligible. Such a property is also true for the extremal Gibbs measures, but, while in equilibrium the magnetization has constant values, in Theorem 6.1 the value is determined by the initial datum and the evolution equation and may thus be an arbitrary function. This is a short time effect, as it is clear in the Glauber case where times are not scaled at all and each spin undergoes only a finite number of flips in a finite time. Thus Theorem 6.1 describes the early stage of the evolution while the main interest is the long time behavior, in particular if and how the true equilibrium is reached. The natural question is then if such effects are still well described by (6.17) and (6.19). Behind this there is an inversion of limits: we want first to run the process for a long time and then take $\gamma \rightarrow 0$, while the opposite is done if we look at the long time behavior of (6.17) and (6.19). It is not obvious at all that the two procedures agree. In fact we will see that there are cases where this is not true and we will concentrate on those where the corrections to (6.17) and (6.19) are of stochastic nature. Before considering the fluctuations from the

¹The result stated for the case of the Kawasaki dynamics does not appear in the literature. The correlation function techniques in [78],[58] can however be used to establish it. The result in [59] is the one quoted here, but in finite macroscopic volumes. In order to cover the infinite volume statement presented here one needs some extra arguments, like in [47].

deterministic evolutions (6.17) and (6.19), we give a heuristic explanation of why the relevant time scale in the Kawasaki case turns out to be γ^{-2} . This can be understood in the following way: the jump rate can be expanded to obtain

$$(6.20) \quad \begin{aligned} c_\gamma(x, y; \sigma) &= 1 - \frac{\beta}{2}[\sigma(y) - \sigma(x)] \sum_{z \neq x, y} (J_\gamma(z, x) - J_\gamma(z, y))\sigma(z) + O(\gamma^2) \\ &= 1 - \frac{\beta\gamma}{2}[\sigma(y) - \sigma(x)]\gamma^d \sum_{z \neq x, y} (y - x) \cdot \nabla J(\gamma(x - z))\sigma(z) + O(\gamma^2), \end{aligned}$$

which in particular says that the process is a $O(\gamma)$ correction of the symmetric SEP. It is in fact similar to the WASEP considered in Section 4, but the asymmetry, i.e. the $O(\gamma)$ term in (6.20), is configuration dependent. This dependence upon the configurations is nice, since it is averaged over a length of order γ^{-1} , exactly the scale on which we are looking for the result. Recalling that the hydrodynamic limit of the WASEP process is (2.1), it does not come as a surprise that in this case the hydrodynamic equation is (6.19).

6.4. Small Fluctuations. In agreement with the Fluctuating Hydrodynamic Theory, the small deviations from the macroscopic limit are described by the solution of a linear SPDE obtained by linearizing the macroscopic equation and adding a stochastic noise. Using the notation and assumptions of Theorem 6.1 we introduce the fluctuation fields

$$(6.21) \quad X_\tau^{(\gamma, G)}(\phi) = \gamma^{d/2} \sum_{x \in \mathbb{Z}^d} \phi(\gamma x)[\sigma(x, \tau) - m^{(G)}(\gamma x, \tau)],$$

$$(6.22) \quad X_\tau^{(\gamma, K)}(\phi) = \gamma^{d/2} \sum_{x \in \mathbb{Z}^d} \phi(\gamma x)[\sigma(x, \gamma^{-2}\tau) - m^{(K)}(\gamma x, \tau)].$$

THEOREM 6.2. *For any $T > 0$ the family $\{X^{(\gamma, G)}\}_{\gamma > 0}$ is weakly convergent in $D([0, T]; \mathcal{S}')$ as $\gamma \rightarrow 0$. The limit process $\{\xi^{(G)}(r, t)\}$ is in $C([0, T]; \mathcal{S}')$ and satisfies the following SPDE*

$$(6.23) \quad \begin{aligned} \frac{\partial \xi^{(G)}(r, t)}{\partial t} &= \int_{\mathbb{R}} \mathcal{L}(r, r', t) \xi^{(G)}(r', t) dr' \\ &+ \sqrt{2[1 - m^{(G)}(r, t) \tanh\{\beta(J * m^{(G)})(r, t) + h\}]} \dot{W}(r, t) \end{aligned}$$

and

$$(6.24) \quad \mathcal{L}(r, r', t) = -\delta(r - r') + \frac{\beta J(r - r')}{\cosh^2[\beta(J * m^{(G)})(r, t) + h]}.$$

We stress once again that the drift term of the diffusion process $\xi^{(G)}$ is simply given by the linearization of the operator on the right-hand side of (6.17) around the solution $m^{(G)}(r, t)$.

Also in the Kawasaki case the fluctuations are expected to scale to an infinite dimensional diffusion process whose drift is given by the linearization of (6.19) around a solution $m^{(K)}(r, \tau)$ and the martingale part is the gradient of white noise,

in space and time, with a spatial modulation. In SPDE terms they should scale to the solution of

$$(6.25) \quad \begin{aligned} \frac{\partial \xi^{(K)}}{\partial t} = & \Delta \xi^{(K)} - \beta \nabla \left[(1 - [m^{(K)}]^2) \nabla J * \xi^{(K)} \right] \\ & + 2\beta \nabla \left[\xi^{(K)} \nabla J * m^{(K)} \right] + \nabla \left[\sqrt{1 - [m^{(K)}]^2} \dot{W} \right]. \end{aligned}$$

This result is not proven, but it should be inside the realm of the known techniques [24],[31],[26],[19].

7. Nonlinear Fluctuations: Stochastic Allen–Cahn and Cahn–Hilliard Equations

In this section we will see that in a neighborhood of the critical temperature both the Glauber and the Kawasaki dynamics considered in the previous section have anomalous fluctuating behaviors. We will show that, for certain initial conditions, a suitable random function on the $d = 1$ Glauber spin flip process converges to the solution of a stochastic Reaction–Diffusion equation with additive white noise. We will also discuss partial results for the $d = 2$ Glauber process and present some heuristics about the Kawasaki dynamics that indicate convergence to a fourth order PDE with *additive conservative white noise*.

7.1. The General Strategy. We will only study deviations from the constant profile $m(r) = 0$, represented at the spin level by the product measure μ with 0 averages, i.e. $E_\mu(\sigma(x)) = 0$ for all x . At the macroscopic level the state $m \equiv 0$ is stationary because $m(r, t) = 0$ is a stationary solution both for (6.17) (with $h = 0$) and (6.19). The first order corrections to the state $m \equiv 0$ are described by the limit behavior of the fluctuation fields, characterized by (6.23) and (6.25). These are linear equations that can be solved quite explicitly. The interesting case for us is when their solutions diverge as $t \rightarrow \infty$, so that, according to the linear theory, the deviations from the state $m \equiv 0$ grow indefinitely. This cannot be correct at the spin level because the validity of the linear approximation, which leads to (6.23) and (6.25), fails when the fluctuations become too large. Exactly at this point the non linear effects, that could be neglected so far, become relevant and we may have both non linear and stochastic terms all together, hence non linear SPDE's.

We start from the Glauber dynamics with $h = 0$. Since $m^{(G)}(r, t) = 0$ we get from (6.23) that the limit fluctuation field $\xi^{(G)}(r, t)$, which we simply write here as $\xi(r, t)$, satisfies the equation

$$(7.1) \quad \frac{\partial \xi(r, t)}{\partial t} = -\xi(r, t) + \beta(J * \xi)(r, t) + \dot{W}(r, t).$$

The covariance is

$$(7.2) \quad C_t(r, r') = E(\xi(r, t)\xi(r', t)), \quad C_0(r, r') = \delta(r - r').$$

Setting $\alpha = 1 - \beta$ and recalling that the integral of J is equal to 1 (we look here at the case in which J is isotropic, i.e. it depends on r only via $|r|$), we have

$$(7.3) \quad C_t(r, r') = 2\beta \int_0^t ds e^{-2\alpha s} \int dr_1 \int dr_2 e^{Gs}(r, r_1) e^{Gs}(r', r_2) J(r_1 - r_2)$$

where G is the generator of the following jump process on \mathbb{R}^d :

$$(7.4) \quad (Gf)(r) = \beta \int dr' J(r - r')[f(r') - f(r)].$$

7.2. The Stable Case, $\beta < 1$. We have from (7.3)

$$(7.5) \quad \lim_{t \rightarrow \infty} C_t(r, r') = C_\infty(r, r')$$

By (6.21) and Theorem 6.2 we then have

$$(7.6) \quad \lim_{t \rightarrow \infty} \lim_{\gamma \rightarrow 0} E_\mu^{(\gamma, G)} \left(X_t^{(\gamma, G)}(\phi) X_t^{(\gamma, G)}(\psi) \right) = \int \int dr dr' \phi(r) \psi(r') \{ \delta(r - r') + C_\infty(r, r') \}.$$

In [16] it is proved that

$$(7.7) \quad \lim_{\gamma \rightarrow 0} E_{\mu_{\beta, \gamma}} \left(X_0^{(\gamma, G)}(\phi) X_0^{(\gamma, G)}(\psi) \right) = \int \int dr dr' \phi(r) \psi(r') \{ \delta(r - r') + C_\infty(r, r') \},$$

where $\mu_{\beta, \gamma}$ is the unique Gibbs measure in \mathbb{Z}^d , recall that $\beta < 1$.

In conclusion when $\beta < 1$ the fluctuations remain *small* and converge to their equilibrium value (7.7): we do not have in this case the desired growth of the fluctuations nor the nonlinear SPDE's.

7.3. The Unstable Case, $\beta > 1$. We have $\alpha = 1 - \beta < 0$, so that, contrary to the stable case, the covariance in (7.3) diverges exponentially as $t \rightarrow \infty$. This seems the best setup for finding nonlinear effects and indeed they will appear. But the instability is so strong that when they enter into play, the random effects are already gone. We will deal with this case in Section 8.

7.4. The Critical Case, $\beta = 1$. In this case $\alpha = 0$ so that the drift term in (7.1) is given by the generator G defined in (7.4) which, for what follows, should be thought of as a sort of Laplacian. The magnitude of the fluctuations $\xi(r, t)$ depends then in a highly non trivial way on the size of the system and when this is defined on the whole space, the number of dimensions of the space become critically relevant. We will consider first the system in a unit torus of \mathbb{R}^d , then in the whole \mathbb{R}^d , first with $d = 1$ and finally with $d = 2$.

In the unit torus $\mathbb{R}^d/\mathbb{Z}^d$ the total fluctuation is

$$(7.8) \quad \Xi_t = \int_{\mathbb{R}^d/\mathbb{Z}^d} dr \xi(r, t),$$

and its typical magnitude is

$$(7.9) \quad K_t = E \left(\Xi_t^2 \right)^{1/2} = \left(\int dr C_t(0, r) \right)^{1/2},$$

which by (7.3) grows like $t^{1/2}$. In the whole \mathbb{R}^d the covariance $C_t(0, r)$ decays on the scale $r \approx \sqrt{t}$. We then replace (7.8) by averaging $\xi(r, t)$ over the ball $B(\sqrt{t})$ of volume $t^{d/2}$ which is centered at 0. We thus set

$$(7.10) \quad \Xi_t = t^{-d/2} \int_{B(\sqrt{t})} dr \xi(r, t).$$

Analogously to (7.9), its typical magnitude is

$$(7.11) \quad K_t = E\left(\Xi_t^2\right)^{1/2} \approx \left(t^{-d/2} \int_{\mathbb{R}^d} dr C_t(0, r)\right)^{1/2}.$$

Using (7.3) we then conclude that K_t grows proportionally to $t^{1/4}$ in $d = 1$ while it remains bounded in $d = 2$.

However while in $d = 1$ (and in the unit torus as well) $\xi(r, t)$ is close to Ξ_t , this is no longer true in $d = 2$ where in fact $\xi(r, t)$ is defined only as a distribution. This reflects on the fact that a *local fluctuation*

$$(7.12) \quad \Xi_t^* = \int_B dr \xi(r, t)$$

B a unit ball, has typical magnitude of the order of $(\log t)^{1/2}$ (recall that $d = 2$). When averaging over $B(\sqrt{t})$, all these logarithmic divergences compensate so that Ξ_t is typically bounded. But when the non linear terms become important there is no longer an exact compensation and, as we will see, such a logarithmic divergence will play a crucial role, with effects similar to those emerging in the Euclidean field theory.

7.5. Rescaled and Renormalized Fluctuations. Recalling the normalization in (6.21), the deviations from $m^{(G)} = 0$ in the regime of validity of (7.1) are given by $M_t^{(\gamma)} = \gamma^{d/2} \Xi_t$. Since the non linearity in (6.17) around $m^{(G)} = 0$ is cubic, we may expect this term to be relevant on a time interval t if $[M_t^{(\gamma)}]^3 t$ has the same order as $M_t^{(\gamma)}$ itself, i.e. what produced by (7.1) in the same time interval. Recalling the previous computations of the typical value, K_t , of Ξ_t , we find that the above criterion yields for the critical time scale in the unit torus of \mathbb{R}^d

$$(7.13) \quad \gamma^{d/2} t^{1/2} = [\gamma^{d/2} t^{1/2}]^3 t$$

namely

$$(7.14) \quad t = \gamma^{-d/2}, \quad M_t^{(\gamma)} = \gamma^{d/4}$$

The same argument in \mathbb{R} gives

$$(7.15) \quad \gamma^{1/2} t^{1/4} = [\gamma^{1/2} t^{1/4}]^3 t, \quad t = \gamma^{-2/3}, \quad M_t^{(\gamma)} = \gamma^{1/3}$$

while, in \mathbb{R}^2 , since K_t remains bounded,

$$(7.16) \quad \gamma = [\gamma]^3 t, \quad t = \gamma^{-2}, \quad M_t^{(\gamma)} = \gamma.$$

In all the above cases $M_t^{(\gamma)}$ is still infinitesimal as $\gamma \rightarrow 0$, it is then natural to conjecture that the only relevant non linearity is cubic and that this critical regime is described by

$$(7.17) \quad \frac{\partial X}{\partial t} = -X + J \star X - \frac{1}{3}(J \star X)^3 + \gamma^{d/2} \dot{W}$$

Besides its relation with the spin system, equation (7.17) looks like a natural regularization of the Cahn–Allen equation with noise which is therefore interesting in

its own right. The heuristics behind (7.14), (7.15) and (7.16) suggests to study (7.17) with the following change of variables:

$$(7.18) \quad Z_t^{(\gamma)}(r) = \begin{cases} \gamma^{-d/4} X(r, \gamma^{-d/2}t) & \text{in } \mathbb{R}^d / \mathbb{Z}^d \\ \gamma^{-1/3} X(\gamma^{-1/3}r, \gamma^{-2/3}t) & \text{in } \mathbb{R} \\ \gamma^{-1} X(\gamma^{-1}r, \gamma^{-2}t) & \text{in } \mathbb{R}^2 \end{cases}$$

According to the previous heuristics, $Z_t^{(\gamma)}$ should (or better may) converge as $\gamma \rightarrow 0$ to a process described by a nonlinear SPDE.

We conclude this subsection with the particle analog of $Z_t^{(\gamma)}$:

$$(7.19) \quad Y_t^{(\gamma)}(\phi) = \begin{cases} \gamma^{-d/4+d} \sum \sigma(x, \gamma^{-d/2}t) \phi(\gamma x) & \text{in } \mathbb{Z}^d / [\gamma^{-1}] \mathbb{Z}^d \\ \gamma^{-1/3+(1+1/3)} \sum \sigma(x, \gamma^{-2/3}t) \phi(\gamma^{1+1/3}x) & \text{in } \mathbb{Z} \\ \gamma^{-1+4} \sum \sigma(x, \gamma^{-2}t) \phi(\gamma^2 x) & \text{in } \mathbb{Z}^2. \end{cases}$$

7.6. The Limit Equations. The above conjecture is proved in the torus and in $d = 1$ dimensions for the spin system, we are not aware of the analogous result for $Z_t^{(\gamma)}$. The case in \mathbb{R} is solved in [7], [49], the analysis however easily extends to the case of the torus. Starting from the latter, we have that $\{Y_t^{(\gamma)}\}_{t \geq 0}$ converges weakly to the solution of the ordinary stochastic differential equation

$$(7.20) \quad \frac{dX}{dt} = -\frac{1}{3} X^3 + \dot{W}$$

(i.e. long times in bounded regions suppress all the spatial structure). In \mathbb{R} instead $\{Y_t^{(\gamma)}\}_{t \geq 0}$ converges weakly to the solution of the Cahn–Allen equation with noise

$$(7.21) \quad \frac{\partial X}{\partial t} = \Delta X - \frac{1}{3} X^3 + \dot{W}.$$

7.7. Critical Fluctuations in $d = 2$. It may look at first sight surprising that in $d = 2$, for any $a > 0$ $\{Y_t^{(\gamma)}\}_{t \geq a}$ converges weakly to the field identically equal to 0. The same phenomenon happens in equilibrium, [16]: calling $\mu_{\beta,\gamma}$, here $\beta = 1$, the (unique) Gibbs measure on $\{-1, 1\}^{\mathbb{Z}^2}$ (uniqueness at $\beta = 1$ is also proved in [16]), we have that for any test function ϕ

$$(7.22) \quad \lim_{\gamma \rightarrow 0} E_{\mu_{\beta,\gamma}}(Y_0^{(\gamma)}(\phi)^2) = 0,$$

where $Y_0^{(\gamma)}(\phi)$ is given by the last expression in (7.19). The same conclusions obviously apply to $Y_t^{(\gamma)}(\phi)$, for any $t > 0$, if the process starts from the invariant measure $\mu_{\beta,\gamma}$. This shows that the normalization and rescaling in (7.19) are not correct in $d = 2$ at $\beta = 1$. The question has been studied in [48] where it is shown that if the Glauber process starts from the equilibrium $\mu_{\beta,\gamma}$, then the correct equation that describes to leading orders the magnetic density field is

$$(7.23) \quad \frac{\partial X_t^{(\gamma)}}{\partial t} = [J * X_t^{(\gamma)} - X_t^{(\gamma)}] - C \gamma^2 \log \gamma^{-1} J * X_t^{(\gamma)} + \gamma \dot{W}, \quad C = \frac{1}{\pi D}$$

$$(7.24) \quad D = \int J(r) r^2 dr$$

Observe that as $\gamma \rightarrow 0$, the solution of (7.23) vanishes so that the 0-th order approximation to $X_t^{(\gamma)}$ is the stationary state $m_\beta \equiv 0$. By multiplying $X_t^{(\gamma)}$ by

$\gamma^{-d/2} = \gamma^{-1}$, we recover the fluctuation fields and in fact in the limit as $\gamma \rightarrow 0$ with time running in a compact interval, we obtain (7.1) (with $\beta = 1$). The new result in [48] is that the spin systems follows (6.22) also in the regime defined by the following renormalization procedure

$$(7.25) \quad u(r, t) = \gamma^{-1} \lambda^{-2} t(\lambda^{-1} r), \quad \lambda = \gamma[\sqrt{\log \gamma^{-1}}]^{-1}.$$

If $X_t^{(\gamma)}$ solves (7.23), the function u in (7.25) is actually independent of γ and it solves

$$(7.26) \quad \frac{\partial u}{\partial t} = \Delta u - Cu + \dot{W}.$$

Thus the limit of the Glauber process under the renormalization as in (7.25) is described by the solution of (7.26).

The analogous limit procedure applied to (7.1) leads to (7.26) without the term $-Cu$ and since (7.1) describes the linear approximation, we may say that the extra term $-Cu$ is due to non linear effects. The time and space rescaling in the definition of $Y_t^{(\gamma)}$, see (7.19), is obtained by an extra rescaling with respect to that in (7.25). Again in [48] it is shown that there is still agreement between (7.23) and the Glauber process under this extra rescaling, thus the Glauber process is described by the long time behavior of (7.26). The extra mass term $-Cu$ determines a decay of correlations in (7.26) which therefore explains the triviality of the limit (7.22).

To summarize, we have seen that the non linearity in (7.17) reflects in an extra mass term $-Cu$ in (7.26) and that this is responsible for a decay of correlations not present in (7.1) which ultimately causes the triviality of the limit of $Y_t^{(\gamma)}$ and the failure of the heuristic argument for the derivation of a non linear SPDE. We can say a little more about this issue, but unfortunately only at a heuristic level. Indeed (7.23) can be formally derived from (7.17) by the substitution

$$(7.27) \quad -\frac{1}{3}(J \star X)^3 \longrightarrow C\gamma^2 \log \gamma^{-1} (J \star X).$$

On the other hand the covariance $C_t(0, 0)$ when $\beta = 1$ and $d = 2$ behaves for $t = \gamma^{-2}$ just as $C \log \gamma^{-1}$ (the extra factor γ^2 comes from the normalization of the fluctuations). The result may thus be interpreted by saying that the leading contribution to $(J \star X)^3/3$ is given by $J \star X$ times the covariance of $(J \star X)$ computed using the linear approximation (at the times under consideration). As already observed, in $d = 2$ such a value is much larger (by a factor $\log \gamma^{-1}$) than the average magnitude K_t , $t = \gamma^{-2}$, given by (7.11) (which remains bounded). The non linear effects are here important and they occur when the noise is still effective, but they manifest via (7.27) only as a linear extra mass term in the final equation (7.26). To kill the extra mass term we go back to (6.17) and repeat the expansion that leads to (7.17) when the inverse temperature is $\beta = 1 + a\gamma^2 \log \gamma^{-1}$, $a > 0$. We formally obtain to leading order in γ the new equation

$$(7.28) \quad \frac{\partial X}{\partial t} = \beta[J \star X - X] + [a\gamma^2 \log \gamma^{-1}]X - \frac{1}{3}(J \star X)^3 + \gamma^{d/2}\dot{W}.$$

The conjecture is that by letting $a \rightarrow C$, C as in (7.23), the second term in (7.28) will cancel the extra mass term in (7.23). Unfortunately there is not yet a proof of such a statement, neither at the spin level nor starting from (7.28) itself.

The second term on the right hand side of (7.28) is the Wick regularization of the cubic term which provides a nice (and we think enlightening) interpretation of

the Wick regularization in terms of critical phenomena in statistical mechanics. In fact the value $\beta = 1 + C\gamma^2 \log \gamma^{-1}$ has a particular meaning in the theory of Gibbs states. In [16] it is proven that the inverse critical temperature equal to 1 in the mean field limit, is for $\gamma > 0$ not smaller than $1 + C\gamma^2 \log \gamma^{-1}$. It is conjectured that this is the true value of the critical temperature, possibly modulo corrections of higher order in γ .

7.8. Kawasaki Dynamics. An analogous scheme of analysis (stable, critical, unstable) can be repeated for the Kawasaki case. Most of the rigorous results are however still missing. Here we look at the heuristics in the critical case ($\beta = 1$) and in the next section we will briefly consider the unstable case ($\beta > 1$).

As in the last part of the previous subsection we will consider temperatures which are γ -dependent, more precisely we will look at cases in which the temperature is close to the critical one: this is done because we want to obtain a more general form of the limiting equation. We will always work in $d = 1$. We set $\beta = \beta(\gamma) = 1 + a\gamma^{2/3}$ ($a \in \mathbb{R}$), $a \in \mathbb{R}$, and start the process with the product measures $\mu^{(\gamma)}$ such that for any $x \in \mathbb{Z}$

$$(7.29) \quad \mu^{(\gamma)}(\sigma(x)) = \gamma^{1/3} m^*(\gamma^{4/3} x),$$

where m^* is a continuous bounded function. The formal analysis below suggests that the law of the random function

$$(7.30) \quad \sigma_t^\gamma(r) \equiv \gamma^{-1/3} \sigma_{t\gamma^{-3-1/3}}([\gamma^{-4/3} r]),$$

converges (weakly and for the sake of definiteness in the same sense as in the Glauber case) to the law \mathcal{P} on $C(\mathbb{R}^+; C^0(\mathbb{R}))$ of the solution of the following SPDE

$$(7.31) \quad \frac{\partial m}{\partial t} = \Delta \left[U'(m) - \frac{D}{2} \Delta m \right] + \sqrt{2} \nabla \dot{W},$$

$$(7.32) \quad m(0, \cdot) = m^*(\cdot),$$

where D is defined in (7.24) and

$$(7.33) \quad U(m) = \frac{m^4}{4} - \frac{am^2}{2}.$$

It is already a nontrivial issue to give a meaning to (7.31) and hence to define \mathcal{P} . At first sight it may seem that the noise is too irregular to allow the existence of a solution which is continuous in space and time. This is however not true: the right-hand side of (7.31) contains the operator $-\Delta^2$, which has very strong regularizing properties. In particular it is easy to see that if U' is linear, m will be in $C(\mathbb{R}^+; C^0(\mathbb{R}))$. The nonlinear case is of course much less trivial, but in [50] it is shown that if $U(m) = m^2 + V(m)$, where V is a bounded C^3 function, (7.31) has a unique solution in $C(\mathbb{R}^+; C^0)$ for a suitable class of initial data. The extra problem that we encounter here is the fact that U' is not globally Lipschitz. Choosing properly the original lattice domain, we could have convergence to equation (7.31) on a torus and, in this case, one can deal with the lack of a global Lipschitz bound quite directly (see [42] for the case of stochastic Reaction-Diffusion equations and [22] for the case of a stochastic Cahn-Hilliard equation with additive noise, which, however, is not of the conservation law type). In the case of (7.31), as stated above, the method in [65] should apply.

The heuristics of the derivation of (7.31) goes as follows. In analogy with what we did above, to we start off with the following SPDE:

$$(7.34) \quad \partial_t m = \Delta m - \beta \nabla [(1 - m^2) \nabla J * m] + \gamma^{1/2} \sqrt{2} \nabla \left(\sqrt{1 - m^2} \dot{W}_\gamma \right).$$

Which is the analog of (4.1). where there is ϵ instead of γ : the deterministic part is given by (6.19) and the stochastic part is the same as in (4.1), apart for the factor $\sqrt{2}$. The system we are dealing with now, as explained after formula (6.20), is still a weakly asymmetric process and it is rather easy to see that, to highest order, the noise term is determined by the $O(1)$ part of the generator, which is a simple exclusion process, but speeded up of a factor 2 (whence the $\sqrt{2}$ in front of the noise part). The very same heuristics can be worked out directly for the particle model. As before, in order to use directly the exact scaling relations for the white noise without having to worry about the corrections given by the γ -regularization, we will replace \dot{W}_γ with \dot{W} .

Now we scale space with $\gamma^{-1/3}$ and time with $\gamma^{-4/3}$: this scaling is non diffusive and reflects the fact that the limit equation is of 4^{th} order. The strange exponents are a consequence of the fact that the nonlinearity is given by a cubic power. More precisely we set

$$(7.35) \quad u(x, \tau) = \gamma^{-1/3} m(\gamma^{-1/3} x, \gamma^{-4/3} \tau),$$

with $x \in \mathbb{R}$ and $\tau \in \mathbb{R}^+$, and we use the scaling properties of the white noise ($\dot{W}(\gamma^{-1/3} x, \gamma^{-4/3} \tau)$ is equal in law to $\gamma^{1/6} \gamma^{2/3} \dot{W}(x, \tau)$) we obtain (recall that $\beta = 1 + a\gamma^{2/3}$)

$$(7.36) \quad \begin{aligned} \partial_\tau u &= \gamma^{-2/3} \left[\Delta u - (1 + a\gamma^{2/3}) \nabla \left((1 - \gamma^{2/3} u^2) J_{\gamma^{1/3}} * \nabla u \right) \right] \\ &\quad + \nabla \left(\sqrt{1 - \gamma^{2/3} u^2} \dot{W} \right), \end{aligned}$$

where $J_{\gamma^{1/3}}(\cdot) = \gamma^{-1/3} J(\gamma^{-1/3} \cdot)$. Expand $J_{\gamma^{1/3}} * \nabla u$ to obtain

$$(7.37) \quad \nabla u + \gamma^{2/3} \frac{D}{2} \nabla \Delta u + \dots,$$

where D is defined in (7.24). Once we insert (7.37) in (7.36), we realize that the terms of order $\gamma^{-2/3}$ cancel and, if we take only the terms of order 1, we find (7.31). Recalling that equation (6.19) was derived by scaling space with γ^{-1} and time with γ^{-2} , we have thus justified, at a formal level, our claim that the limit behavior of (7.30) is given by (7.31).

8. Macroscopic Effects of Small Fluctuations: the Origin of Spatial Patterns

We will now outline some of the results obtained in [28] as a prototype of the following situation: when a solution of a PDE, obtained as macroscopic limit of an IPS or a more general particle system, is *unstable* under small perturbations, the behavior of the IPS, on long macroscopic times, can deviate substantially from the solution of the PDE. Crucial, in the cases discussed here, are the small stochastic corrections which disappear on the space-time scaling of the hydrodynamic limit. Once again we are facing a problem involving the exchange of a limit: one can study the behavior of the macroscopic law, obtained by sending (in this case) γ to zero, for t large, but this may not coincide with the limit that we obtain as γ tends to zero by letting t depend on γ , with $t \rightarrow \infty$ as $\gamma \rightarrow 0$.

In the notation of Section 7, we are dealing with the case $\alpha < 0$. Therefore, for notational convenience we replace α by $-\alpha$ so that $\alpha = \beta - 1$ is strictly positive.

In [28] the authors consider the Glauber dynamics with a Kac potential as in Section 5. We recall that the hydrodynamic limit for this system is given in Theorem 6.1, equation (6.17), and we will look at the case $h = 0$. One can easily see that the generator of the linearized evolution around the stationary solution $m \equiv 0$, considered as an operator in $L^2(\mathbb{R}^d; dr)$, contains α in its spectrum which is contained in $(-\infty, \alpha]$ (recall that $\int J(r)dr = 1$) and therefore we are in an unstable case. If $\beta = \alpha + 1 > 1$, the nonlocal equation (6.16) has three stationary constant solutions $(-m_\beta, 0, +m_\beta)$, $m_\beta > 0$, of

$$(8.1) \quad m = \tanh(\beta m), \quad m \in [-1, 1],$$

and a straightforward linear analysis shows that $\pm m_\beta$ are stable (m_β coincides with the limit as $\gamma \rightarrow 0$ of the equilibrium magnetization $m_{\beta, \gamma}$, [11]). We will call the states with magnetization $\pm m_\beta$ the *phases*. If we start with $m \equiv 0$, under small perturbations the system will leave this state and it will try to rearrange itself into regions of the two different phases $\pm m_\beta$. We will now study how the system makes this rearrangement. Let us stress however that Theorem 6.1 holds for this initial condition and therefore, for any fixed $\tau \geq 0$, $m = 0$ gives the large scale behavior of the system. On the other hand, Theorem 6.2 is telling us that fluctuations are present and that they are of order $\gamma^{1/2}$, if $d = 1$, and of order $\gamma^{d/2}$ in dimension d (see [28]). Due to the instability, the fluctuations grow at exponential rate, as one can easily understand from the linear analysis of (6.17). At time τ they are of order

$$(8.2) \quad \gamma^{d/2} \exp(\alpha\tau).$$

Theorem 6.2, like Theorem 6.1, is proven for τ finite. However, if we assume its validity at longer times, we obtain that the fluctuations will be $o(1)$ if $\tau = t|\log \gamma|$ and $t < d/2\alpha$. If the above heuristics are correct, the linearization of (6.16) should be a faithful description of the system up to $t_c|\log \gamma|$. We then set

$$(8.3) \quad t_c = \frac{d}{2\alpha},$$

when the nonlinearity should take over.

Going back to the fluctuations results of Section 7 we can make a more accurate analysis, by starting again from (7.1), (7.2) and (7.3). In view of the growth of $C_t(r, r')$ it is convenient to change variables in (7.1), writing, for $\lambda > 0$,

$$(8.4) \quad X_t^{(\lambda)}(r) = \lambda^{-d/2} e^{-\lambda^2 \alpha t} \xi(\lambda r, \lambda^2 t).$$

It is not difficult to prove that for any $s > 0$ the process $\{X_t^{(\lambda)}\}_{t \in [s, \infty)}$ converges weakly in $D([s, \infty); \mathcal{S}')$ as $\lambda \rightarrow \infty$ to the restriction to $[s, \infty)$ of the process $\hat{\xi}(r, t)$ which is the solution of

$$(8.5) \quad \frac{d\hat{\xi}(r, t)}{dt} = \frac{D}{2} \Delta \hat{\xi}(r, t),$$

with $\hat{\xi}(r, 0)$ a Gaussian distribution with mean 0 and covariance

$$(8.6) \quad E \left(\hat{\xi}(r, 0) \hat{\xi}(r', 0) \right) = \delta(r - r') \left[1 + \frac{2}{\alpha} \right]$$

and D is defined in (7.24). Thus $\hat{\xi}(r, t)$ evolves deterministically and the only stochasticity comes from the initial datum. The result we just stated does not hold if $s = 0$: observe in fact that the covariance of $\hat{\xi}(\cdot, 0)$, given by (8.6), is larger than the limit as $\lambda \rightarrow \infty$ of the covariance of $X_0^{(\lambda)}$. This initial layer effect is the only remnant of the white noise present in (7.1) and absent in (8.5).

When we go beyond the linear analysis, see [27],[28], we eventually see the non linear effects, but they appear only when the deterministic regime described above has already been established. We thus have first small fluctuations and a linear SPDE, then a linear deterministic PDE and finally a non linear PDE still without noise. At this latter stage, described by the nonlinear equation, the stochasticity appears only as a random initial datum. The relation with the IPS is obtained by setting

$$(8.7) \quad \lambda = \sqrt{|\log \gamma|}, \quad \epsilon = \gamma/\lambda, \quad \tau_\lambda = t\lambda^2,$$

and introducing the renormalized fluctuation field

$$(8.8) \quad Z_t^\gamma(\phi) = e^{-\alpha\tau_\lambda} \epsilon^{d/2} \sum_{x \in \mathbb{Z}^d} \phi(\epsilon x) \sigma_{\tau_\lambda}(x)$$

(to be compared with (6.21)). Then for any $0 < s \leq t_c = d/2\alpha$ the process $\{Z_t^\gamma\}_{t \in [s, t_c]}$ converges weakly in $D([s, t_c]; S')$ as $\gamma \rightarrow 0$ to the restriction to $[s, t_c]$ of the process $\hat{\xi}(r, t)$ which solves (8.5). After t_c the non linear effects appear and the evolution is described by (6.17).

The above analysis gives a conjecture on the exit time from the unstable state $m = 0$. It turns out that this conjecture can actually be proven and, more surprisingly, that the nonlinear stage of the escape can be treated in great detail too.

We recall that our initial condition for the IPS is the product measure μ with zero average magnetization. We write $P^{(\gamma)} = P_\mu^{(\gamma, G)}$ and $E^{(\gamma)}$ for its expectation.

THEOREM 8.1. Set $R_\gamma = R\lambda/\gamma$. For any $n \in \mathbb{Z}^+$ and $R > 0$ we have that

$$(8.9) \quad \lim_{\gamma \rightarrow 0} \sup_{x_1 \neq \dots \neq x_n, |x_i| \leq R_\gamma} E^{(\gamma)} \left(\prod_{i=1}^n \sigma_{t\lambda^2}(x_i) \right) = 0,$$

for all $t \in [0, t_c]$. Moreover

$$(8.10) \quad \lim_{\gamma \rightarrow 0} \sup_{x_1 \neq \dots \neq x_n, |x_i| \leq R_\gamma} \left| E^{(\gamma)} \left(\prod_{i=1}^n \sigma_{t_c\lambda^2 + (\log \lambda)^2}(x_i) \right) - \mathcal{E} \left(\prod_{i=1}^n m_\beta \text{sign}(\rho(x_i \gamma/\lambda)) \right) \right| = 0,$$

where $\{\rho(r)\}_{r \in \mathbb{R}^d}$ is a centered Gaussian field with covariance

$$(8.11) \quad \mathcal{E}(\rho(r)\rho(r')) = \exp \left\{ -\frac{\alpha(r-r')^2}{2d\beta D} \right\},$$

and D is defined in (7.24).

Theorem 8.1 is a simplified version of some of the results in [28]. The essence of its content can be summed up in the following four points:

1. The relevant time scale for the growth of the fluctuations is $|\log \gamma|$.
2. On this new time scale nothing really happens, i.e. $m = 0$, up to t_c .
3. Slightly after t_c the phase separation has taken place and the magnetization is locally concentrated around the values $\pm m_\beta$, i.e. on the phases.
4. The emergent spatial patterns have a typical length scale of $\gamma^{-1} \sqrt{|\log \gamma|} = \epsilon^{-1}$ lattice sites. On this spatial scale the zero-level set of the Gaussian field ρ gives, in distribution, the random (hyper-)surfaces which separate homogeneous regions. This level set is a.s. a smooth manifold (see [57]).

Remark: The evolution after the appearance of the interfaces is understood too. It is on the time scale λ^2 and it is given by motion by mean curvature of the interfaces between the phases (see [68],[26]).

We now go back to some simple heuristics in order to connect the results in Theorem 8.1 with PDEs with random initial condition. It would be more correct to consider SPDEs with random initial condition, since in the IPS there is both a dynamical noise and a random initial condition. We have already seen that the effect of the dynamical noise, on the time scale we are interested in, is only to change the initial condition and we will see that the observed geometrical patterns are independent of the change in the initial condition. We will thus for simplicity consider only the noise in the initial condition. This analogy with a PDE with random initial datum will turn out to be particularly helpful in understanding better Theorem 8.1 and in trying to outline the main difficulties which come up in the case with a conservation law.

We will exploit below the similarities between the integral equation (6.17) and the Reaction–Diffusion equation (RD). Equation (6.17) can be cast into the form

$$(8.12) \quad \frac{\partial m}{\partial \tau} = \mathcal{D}(m) + \mathcal{R}(m),$$

where

$$(8.13) \quad \mathcal{R}(m) = -m + \tanh(\beta m)$$

and

$$(8.14) \quad \mathcal{D}(m) = \tanh(\beta(J * m)) - \tanh(\beta m).$$

We will interpret \mathcal{D} as a diffusion term and \mathcal{R} as a reaction term. The form (8.12) hence suggests an analogy with the Reaction–Diffusion (or Cahn–Allen) equation [2],[93]]

$$(8.15) \quad \frac{\partial m}{\partial \tau} = \frac{\beta D}{2} \Delta m - U'(m),$$

with $-U'(m) = \mathcal{R}(m)$. Observe that $U(m)$ is a symmetric double-well potential with minima at $\pm m_\beta$ and $U(m) = -\alpha m^2/2 + O(m^4)$, near $m = 0$. We have drawn here an analogy between (8.12) and (8.15) at an intuitive level, but in fact the goes well beyond (see e.g, [26],[27],[28]). From now on we will stick to the RD equation: there would be no essential change in the heuristics below keeping the integral equation (6.17), but we expect that the reader may be more accustomed to (8.15).

We then consider equation (8.15) with random initial condition,

$$(8.16) \quad \begin{cases} \partial_\tau m = (\beta D/2)\Delta m - U'(m) \\ m(r, 0) = \gamma^{d/2}(W * \varphi_\gamma)(r), \end{cases}$$

in which the random field $\{W(r)\}_{r \in \mathbb{R}^d}$ is a white noise. The choice $\varphi(r) = (1/\sqrt{2\pi})^d \exp(-r^2/2)$ simplifies the notation. Repeating in this context the analysis described above, we look at the solution $\ell(r, \tau)$ of the linearization of (8.16). We have

$$(8.17) \quad \ell(r, \tau) = \gamma^{d/2} e^{\alpha\tau} (\mathcal{G}_{\tau D\beta} * (W * \phi_\gamma))(r),$$

where $\mathcal{G}_\tau(r)$ is the heat kernel ($\mathcal{G}_0(r) = \delta(r)$ and $\partial_\tau \mathcal{G}_\tau(r) = (1/2)\Delta \mathcal{G}_\tau(r)$). Hence $\ell(r, t)$ is, for each $t \geq 0$, a centered spatial Gaussian field with covariance

$$(8.18) \quad E(\ell(r, \tau)\ell(r', \tau)) = \gamma^d \exp\{2\alpha\tau\} \mathcal{G}_{2(\tau D\beta + \gamma)}(r - r').$$

As before, if we look at times of order $\lambda^2 (= |\log \gamma|)$, the natural space scale turns out to be λ . In fact, if we set $\ell_\gamma(x, t) = \ell(\lambda x, \lambda^2 t)$ the scaling properties of the heat kernel immediately yield

$$(8.19) \quad E(\ell_\gamma(x, t)\ell_\gamma(x', t)) = \gamma^{d-2\lambda t} \lambda^{-d} \mathcal{G}_{2[D\beta t + (\gamma/\lambda^2)]}(x - x').$$

It is not too difficult to bound the difference between ℓ and the solution $m(r, t)$ of (8.16) for all times $t \leq t_c$ (see e.g. [29],[56]) and what one obtains at t_c is that the solutions of (8.16) can then be roughly described as distributed according to

$$(8.20) \quad \lambda^{-d/2} \rho(x),$$

where ρ is the Gaussian random field introduced in Theorem 8.1.

What happens right after t_c ? Loosely speaking the magnetization is of order $\lambda^{-d/2}$ and due to the instability, which gives an exponential growth, it will take a very short time ($O(|\log \lambda|)$) to make the magnetization profile of order 1, so that the nonlinearity will no longer be negligible. However this argument does not take into account that the simultaneous presence of regions with positive and negative magnetizations may cause, for example, a balance of forces preventing the final growth. To analyze this question, suppose that at time $\lambda^{-2}t_c$ we have a positive magnetization in r . Then with large probability, see (8.20), there will be a large ball (of radius $O(\lambda)$) in which the magnetization is also positive. If the magnetization was positive everywhere, a straightforward application of the monotonicity properties of (8.15) (see e.g. [93]) would show that it would approach the value $+m_\beta$ exponentially fast. The magnetization in our case is positive (at least) in a ball of radius $O(\lambda)$ centered in r , which diverges fast enough as γ goes to zero, to make the situation about the same as if the ball were infinite, because the evolution of (8.15) is almost local and one can control the effect of the magnetization outside the ball (see the Barrier Lemma in [32] and its applications in [29],[56],[57],[28]). The same conclusions hold of course starting from a negative magnetization. It is therefore clear that the one phase domains will be distributed at a time slightly longer than $\lambda^2 t_c$ (see Theorem 8.1) on the space scale λ^{-1} as $\text{sign}(\rho(r))$.

The analysis we proposed is very simplified: in [57] and [28] one can find much finer results on the structure of the interface between the one-phase domains, but these results go beyond the aims of the present work. All these results, properly translated, should apply also to (8.16).

We conclude this section with three remarks.

Remark 1: We gave a formal analysis using the RD equation (8.15), which is not the right hydrodynamic equation for the long range IPS we are dealing with. RD equations can however be derived ([24],[30],[70],[36]) from dynamics of *mixed-type*, i.e. dynamics which are a superposition of Glauber and Kawasaki processes. The generator is chosen of the form

$$(8.21) \quad L_\gamma = \gamma^{-2} L_0 + L_G,$$

where L_0 is the generator of the simple exclusion process and L_G is a local Glauber dynamics:

$$(8.22) \quad L_G f(\sigma) = \sum_{x \in \mathbb{Z}^d} c(x; \sigma) [f(\sigma^x) - f(\sigma)],$$

where $c(x; \sigma) = c(0; \sigma(\cdot + x))$, with $c(0; \cdot)$ a local function from $\{-1, +1\}^{\mathbb{Z}^d}$ to \mathbb{R}^+ . The freedom of choosing the Glauber rates c allows us to reproduce various potentials U appearing in (8.15), including the two well potentials considered here (see [24]) for an easy formula that links the rates c and the potential U).

The dynamics generated by (8.21) is essentially that of a local mean field. The Glauber process is short range, but the spins are stirred around by the exchange part of the generator, which is speeded up by a factor γ^{-2} . Hence every spin will effectively perform a random walk with jump rates γ^{-2} , so that in a time of order 1 (the time-scale for a Glauber event) each spin will have traveled a distance $O(\gamma^{-1})$. Therefore, loosely speaking, a spin is not really interacting with its immediate neighbors, but with some other spins chosen at random in a ball of radius γ^{-1} . The dynamics (8.21) has a nice interpretation in terms of birth-death spatial processes (see e.g. [36],[71] for connections with biological problems and [36] also for other results in the same spirit as the ones presented here).

Remark 2: The case with a conservation law, that is the pure Kawasaki case is very relevant for the applications, for example to certain coarsening processes in metallurgy (see e.g. [12],[77],[86]). Kawasaki dynamics are typically harder to analyze, but in this case most of the difficulties are present already at the PDE level (for a discussion with numerical simulations on the IPS see [58]). In analogy with the Glauber case, we look at the initial value problem (see Theorem 6.1)

$$(8.23) \quad \begin{cases} \partial_\tau m = \Delta m(r, \tau) - \beta \nabla [(1 - m^2)(\nabla J * m)] \\ m(r, 0) = \gamma^{d/2}(W * \varphi_\gamma)(r), \end{cases}$$

with $\beta > \beta_c = 1$, or, replacing the nonlocal equation in (8.23) with the Cahn–Hilliard equation [13]]

$$(8.24) \quad \begin{cases} \partial_\tau m = \Delta [U'(u) - \Delta u] \\ m(r, 0) = \gamma^{d/2}(W * \varphi_\gamma)(r), \end{cases}$$

in which U is a symmetric double well potential, as before. We observe moreover that the random initial condition problems (8.16), (8.23), (8.24), once properly stated, are well posed even if the initial condition is not regularized, i.e. if $m(r, 0) = \gamma^{d/2}W(r)$, and the solution at any $\tau > 0$ is smooth. We also note that the analogy between the integrodifferential equation in (8.23) and the CH equation (8.24) can be pushed quite far [60] and we do not expect any significant difference in the behavior of (8.23) and (8.24) on the relevant space–time scales.

Again it is easy to realize that $m = 0$ is an stationary unstable state for both the equation in (8.23) and the equation (8.24) ($m \equiv c$, with $c \in [-1, 1]$, is a stationary solution and it is unstable for $|c| \leq \bar{c}$, where \bar{c} is easily computed). However, unlike the case without a conservation law, large positive values of the spectrum of the linearized evolution operator (the strongest instabilities) are not associated with long wavelength perturbations, but they are concentrated around a definite (finite) wavelength. In fact if we look at the linearization of (8.24) around $m = 0$ for the Fourier transform $\hat{m}(k, \tau)$ ($k \in \mathbb{R}^d$), we obtain ($U'(m) = \alpha m + O(m^3)$)

$$(8.25) \quad \frac{\partial \hat{m}(k, \tau)}{\partial \tau} = k^2[\alpha - k^2]\hat{m}(k, \tau),$$

and it is evident that the *most unstable* mode is associated with $k \in \mathbb{R}^d$ such that $|k| = \sqrt{\alpha/2}$. For what concerns the time-scale in which the phase segregation phenomenon should be observed, the situation is hence very similar to the one which we already discussed in the case without a conservation law, since the perturbation is $O(\gamma^{d/2})$ and the instability will produce a growth of the type $\exp(t\alpha^2/2)$: the exit time, meant as time at which the linear approximation breaks down, will then be given, on the time scale λ^2 , by $t_c = d/\alpha^2$. The situation is however completely changed with respect to the spatial patterns that we observe (already in the linear analysis): in first approximation they will be a (random) superpositions of waves with wave numbers k with $|k| \approx \sqrt{\alpha/2}$ (the Cahn field [12]). In particular the spatial patterns will vary on the space scale 1 (since $\sqrt{\alpha/2}$ is a fixed number) and not on the very large scale λ (which diverges as $\gamma \rightarrow 0$), as we found in the Glauber case. The nonlinear analysis of the last stages of the escape is then much harder: as an extra difficulty, the conservation law makes the behavior of (8.24) and (8.23) essentially nonlocal and most of the tools available for RD and integral equations (8.15) and (6.16), notably domination techniques and barrier lemmas, do not apply to this case.

We remark that in some works in applied sciences (see e.g. [12], see also [97] and [89]) a procedure completely analogous to the one proven for systems without a conservation law is applied to the conserved case at a heuristic level: the random domain patterns after the appearances of the phases are taken to be distributed according to the sign of the random-wave field that one observes before the time of the escape, i.e. the Cahn field. This ansatz appears to give random domain patterns visually similar to those observed in experiments, computer simulations and numerical solutions of the Cahn–Hilliard equation. Detailed work in the direction we outlined has been done in $d = 1$ [61] and some recent work has been done also in higher dimensions [82].

Remark 3: The spatial patterns observed after the exit time evolve. As already remarked, in the case without a conservation law this evolution is by mean curvature [69],[68],[9],[26]. In the case with a conservation law, when the domain patterns are very large (they are not very large right after the exit time), the evolution is by the so called Mullins–Sekerka law (see [85],[87],[1],[57]). How the distribution of spatial patterns change under these laws is a very popular subject in physics and there are several conjectures, mostly based on the assumption that at very large time the patterns will be statistically self-similar (see [45] and references therein). A mathematical understanding of these problems is missing (and some of the physicists' conclusions are still controversial).

9. The Dynamics on Very Long Times: a Brief Look at Large Deviations

As we have seen in the previous sections, the hydrodynamic equations may fail completely to catch the behavior of an IPS beyond the time scale on which they are derived. While in the previous section we looked at situations in which the hydrodynamic solution is unstable, and thus particularly sensitive to the small perturbations arising from the discrete and stochastic nature of the system, we will now make some remarks about deviations from deterministic hydrodynamic behavior when the solution of the hydrodynamic equation is stable under small perturbations. We will restrict our analysis to cases in which the hydrodynamics is derived on the diffusive scaling. It is intuitively clear that the time that we have to wait to observe a *significant* (i.e. macroscopic) deviations from hydrodynamic behavior are much longer than the ones we have to wait in the presence of instabilities.

Several large deviation results for IPS, mostly associated with hydrodynamic limits, have been recently derived. We list here only a few references. The large deviation result for the simple exclusion process has been derived in [70]; many other results followed, dealing with more and more complex systems (see e.g. [63],[76],[75],[90]).

For the models introduced in Section 6, considered in finite macroscopic volumes (i.e. on boxes of diameter $\propto \gamma^{-1}$ lattice sites), the large deviation statement associated with Theorem 6.1 can be found in [20], for the Glauber case, and [3] for the Kawasaki case. Large deviations for systems with long range interaction (Section 6) is of particular interest, since the hydrodynamic equations derived for these systems have for, $\beta > 1$, more than one stationary stable state. The standard approach of Freidlin and Wentzel [46], extended to this infinite dimensional setting, can then be applied to understand how the system can make a noise-induced transition from one of these stationary states to another (see for details [20],[3]). This analysis presents however several difficulties, mainly due to the great complexity of these IPS. An essential source of simplification comes from the fact that these two models are reversible with respect to a Gibbs measure and one can then obtain an expression for the *quasi-potential* (the infimum of the large deviation functional over the trajectories which lead to the escape from the basin of attraction of a stable stationary point of the macroscopic law) in terms of the equilibrium free energy and, moreover, something is known a priori about the *minimizing* trajectories. Even in very simple non reversible models, the situation is much less clear. We refer to [63] [51] for results in the Glauber+Kawasaki case generated by (8.21) and to [52] and [80] for another particular case.

The most relevant question, from the viewpoint taken in this chapter, is whether understanding the structure of the large deviations for an IPS can give some further insights toward understanding which SPDE better approximates an IPS. We believe that this is indeed the case. For example the results in [70] seem to suggest that the stochastic equation (4.1) is, also from the large deviations viewpoint, the *correct* stochastic improvement of the Burgers equation, when we are concerned with approximating the WASEP. We will not elaborate further on this here and we refer to [40], where a similar viewpoint is set forth, and to [95] where an informal, but very clear introduction to large deviations from hydrodynamic limits is given.

Acknowledgments

We would like to thank Lorenzo Bertini, Fraydoun Rezakhanlou and Herbert Spohn for discussions and comments. E.P. thanks the Courant Institute and Rutgers University for their hospitality. G.G. and J.L.L. thank the IHES (Bures-sur-Yvette, France) for its hospitality.

Bibliography

- [1] N. Alikakos, P. W. Bates and X. Chen: Convergence of the Cahn–Hilliard equation to the Hele–Shaw model *Arch. Rat. Mech. Anal.* **128** (1994) 165–205.
- [2] S. Allen and J. Cahn: A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening *Acta Metall.* **27** (1979) 1084–1095.
- [3] A. Asselah and G. Giacomin: Metastability for a long range conservative dynamics *in preparation* (1997)
- [4] L. Bertini and N. Cancrini: The stochastic heat equation: Feynman–Kac formula and intermittency *J. Stat. Phys.* **78** (1995) 1131–1138.
- [5] L. Bertini and G. Giacomin: Stochastic Burgers and KPZ equations from particle systems *Comm. Math. Phys.* **183** (1997) 571–607.
- [6] L. Bertini and G. Giacomin: On the long time behavior of the stochastic heat equation *Preprint* (1997).
- [7] L. Bertini, E. Presutti, B. Rüdiger and E. Saada: Dynamical critical fluctuations and convergence to a stochastic non linear PDE in one dimension *Th. Prob. App.* **38** (1993) 689–741.
- [8] P. Billingsley: Weak convergence of probability measures *Wiley series in Probability and Mathematical Statistics* (1968) New York.
- [9] L. Bonaventura: Interface dynamics in an interacting spin system *Nonlinear Analysis TMA* **25** (1995) 799–815
- [10] A. Bovier and M. Zahradník: The low temperature phase of Kac–Ising models *J. Stat. Phys.* (to appear)
- [11] P. Buttà, I. Merola, and E. Presutti: On the validity of the van der Waals in Ising systems with long range interactions *Markov Processes and Related Fields* (to appear)
- [12] J.W. Cahn: Phase separation by spinodal decomposition in isotropic systems *J. Chem. Phys.* **42**(1) 93–99 (1965).
- [13] J. W. Cahn and J. E. Hilliard: Free energy of a nonuniform system.I. inter-facial free energy *Acta Metall.* **28** (1958) 258–267.
- [14] R. Caflisch: The fluid dynamic limit of the nonlinear Boltzmann equation *Comm. Pure Appl. Math.* **33** (1990) 651–666.
- [15] R.A. Carmona and S.A. Molchanov: Parabolic Anderson model and intermittency *Mem. Amer. Math. Soc.* n. **108**, (1994).
- [16] M. Cassandro, R. Marra, and E. Presutti: Upper bounds on the critical temperature for Kac potentials *J. Stat. Phys.* (to appear)
- [17] M. Cassandro, E. Orlandi and E. Presutti: Interfaces and typical Gibbs configurations for one dimensional Kac potentials *Theory Prob. Rel. Fields* **96** (1993) 57–96.
- [18] M. Cassandro and E. Presutti: Phase transitions in Ising systems with long but finite range interactions *Markov Processes and Related Fields* **2**, (1996) 241–262.
- [19] C.C. Chang and H.T. Yau: Fluctuations of one-dimensional Ginzburg–Landau models in non equilibrium *Comm. Math. Phys.* **145** (1992) 209–234.
- [20] F. Comets: Nucleation for a long range magnetic model *Ann. Inst. Henri Poincaré, Probabilités et Statistiques* **23** (1987) 135–178.
- [21] H.E. Cook: Brownian motion in spinodal decomposition *Acta Metal.* **18** (1970) 297–305.
- [22] G. Da Prato and A. Debussche: Stochastic Cahn–Hilliard equation *Nonlinear analysis TMA* **26** (1996) 241–263.
- [23] A. De Masi, R. Esposito and J.L. Lebowitz: Incompressible Navier–Stokes and Euler limits of the Boltzmann equation *Comm. Pure Appl. Math.* **42** (1989) 1189–1214.
- [24] A. De Masi, P. A. Ferrari, and J. L. Lebowitz: Reaction diffusion equations for interacting particle systems *J. Stat. Phys.* **44** (1986) 589–644.
- [25] A. De Masi, C. Kipnis, E. Presutti and E. Saada: Microscopic structure at the shock in the asymmetric simple exclusion *Stochastics* **27** (1989) 151–165.

- [26] A. DeMasi, E. Orlandi, E. Presutti, and L. Triolo: Glauber evolution with Kac potentials. I. Mesoscopic and macroscopic limits, interface dynamics *Nonlinearity* **7** (1994) 1–67.
- [27] A. De Masi, E. Orlandi, E. Presutti and L. Triolo: Glauber evolution with Kac potentials. II. Fluctuation *Nonlinearity* **9** (1996) 27–51.
- [28] A. De Masi, E. Orlandi, E. Presutti and L. Triolo: Glauber evolution with Kac potentials. III. Spinodal decomposition *Nonlinearity* **9** (1996) 53–114.
- [29] A. De Masi, A. Pellegrinotti, E. Presutti, and M. E. Vares: Spatial patterns when phases separate in an interacting particle system *Ann. Prob.* **22** (1994) 334–371.
- [30] A. De Masi and E. Presutti: Mathematical methods for hydrodynamic limits *Lecture Notes in Mathematics* **1501** (1991) Springer–Verlag, Berlin.
- [31] A. De Masi, E. Presutti and E. Scacciatelli: The weakly asymmetric exclusion process *Ann. Inst. H. Poincaré, Prob. et Stat.* **25** (1989) 1–38.
- [32] P. De Mottoni and M. Schatzman: Development of interfaces in \mathbb{R}^n *Proc. Royal Soc. Edinburgh* **116A** (1990) 207–220.
- [33] B. Derrida, J.L. Lebowitz and E. Speer: Shock Profiles for the Asymmetric Simple Exclusion Process in One Dimension *J. Stat. Phys.*, to appear (1997).
- [34] R.J. Di Perna and A. Majda: The validity of non nonlinear geometric optics for weak solutions of conservation laws *Commun. Math. Phys.* **98** (1985) 313–347.
- [35] P. Dittrich and J. Gärtner: A central limit theorem for weakly asymmetric for weakly asymmetric exclusion process *Math. Nachr.* **151** (1991) 75–93.
- [36] R. Durrett: Ten lectures on particle systems *Ecole d'Été de probabilités de Saint-Flour XXIII–1993, in Lectures Notes in Math.* **1608** Springer–Verlag, (1995).
- [37] A.C.D. van Enter, R. Fernandez and A.D. Sokal: Regularity properties and pathologies of position–space renormalization–group transformations: scope and limitations of Gibbsian theories *J. Stat. Phys.* **72** (1993) 879–1167.
- [38] R. Esposito, R. Marra and H. T. Yau: Diffusive limit of asymmetric simple exclusion *Rev. Math. Phys.* **6** (1994) 1233–1267.
- [39] R. Esposito, R. Marra and H. T. Yau: Navier–Stokes equations for stochastic particle systems on a lattice *Comm. Math. Phys.* **182** (1996) 395–456.
- [40] G.L. Eyink: Dissipation and large thermodynamics fluctuations *J. Stat. Phys.* **61** (1990) 533–572.
- [41] L. C. Evans, H. M. Soner, and P. E. Souganidis: Phase transitions and generalized motion by mean curvature *Comm. Pure Appl. Math.* **45** (1992) 1097–1123.
- [42] W.G. Faris and G. Jona–Lasinio: Large fluctuations for a nonlinear heat equation with noise *J. Phys. A: Math. Gen.* **15** (1982) 3025–3055.
- [43] P. Ferrari, C. Kipnis and E. Saada: Microscopic structure of traveling waves *Ann. Probab.* **19** (1991) 226–244.
- [44] P. Ferrari: Shocks in one-dimensional processes with drift *Probability and Phase Transition, Series NATO Adv. Sci. Ins. Ser. C, Math. Phys. Sci. (Cambridge, 1993)* Editor G. Grimmett, **420** Kluwer Acad. Publ. Dordrecht (1994) 35–48.
- [45] P. Fratzl, J.L. Lebowitz, O. Penrose and J. Amar: Scaling functions, self–similarity, and the morphology of phase segregating systems *Phys. Rev. B* **44** (1991) 4794–4811.
- [46] M.I. Freidlin and A.D. Wentzell: Random perturbations of dynamical systems *Number 260 in Grundlehren der mathematischen Wissenschaften* (1994) Springer–Verlag, Berlin.
- [47] J. Fritz: On the diffusive nature of entropy flow in infinite systems: remarks to a paper by Guo–Papanicolaou–Varadhan *Comm. Math. Phys.* **133** (1990) 331–352.
- [48] J. Fritz, R. Marra and E. Presutti: *in preparation*
- [49] J. Fritz and B. Rüdiger: Time dependent critical fluctuations of a one dimensional local mean field model *Prob. Th. Rel. Fields* **103** (1996) 381–407.
- [50] T. Funaki: Derivation of the hydrodynamical equation for one–dimensional Ginzburg–Landau model *Prob. Th. Rel. Fields* **82** (1989) 39–93.
- [51] D. Gabrielli, G. Jona–Lasinio, C. Landim and M.E. Vares: Microscopic reversibility and thermodynamics fluctuations *Proceedings of the conference: Boltzmann's legacy*, Rome (1995).
- [52] D. Gabrielli, G. Jona–Lasinio, C. Landim: Onsager reciprocity relations without microscopic reversibility *Phys. Rev. Lett.* **77** (1996), 1202–1205.
- [53] J. Gärtner: Convergence toward Burgers equation and propagation of chaos for weakly asymmetric exclusion process *Stoch. Proc. Appl.* **27** (1988) 233–260.

- [54] J. Gärtner and E. Presutti: Shock fluctuations in a particle system *Ann. Inst. H. Poincaré* **53** (1990) 1–14.
- [55] H.O. Georgii: Gibbs measures and phase transitions *de Gruyter* (1988) Berlin.
- [56] G. Giacomin: Phase separation and random domain patterns in a stochastic particle model *Stoch. Proc. Appl.* **51** (1994) 25–62.
- [57] G. Giacomin: Onset and structure of interfaces in Kawasaki+Glauber interacting particle system. *Prob. Th. Rel. Fields* **103** (1995) 1–24.
- [58] G. Giacomin: Van der Waals limit and phase separation in a particle model with Kawasaki dynamics *J. Stat. Phys.* **65** (1991) 217–234.
- [59] G. Giacomin and J.L. Lebowitz: Phase segregation dynamics in particle systems with long range interactions I: macroscopic limits *J. Stat. Phys.* **87** (1997) 37–61.
- [60] G. Giacomin and J.L. Lebowitz: Phase segregation dynamics in particle systems with long range interactions II : interface motion *SIAM J. Appl. Math.* to appear.
- [61] C.P. Grant: Spinodal decomposition for the Cahn–Hilliard equation *Comm. in PDEs* **18** (1993) 453–490.
- [62] M. Z. Guo, G. C. Papanicolaou and S. R. S. Varadhan: Nonlinear diffusion limit for a system with nearest neighbor interaction *Comm. Math. Phys.* **118** (1988) 31–59.
- [63] G. Jona Lasinio, C. Landim and M.E. Vares: Large deviations for a reaction diffusion model *Prob. Th. Rel. Fields* **97** (1993) 339–361.
- [64] R. Holley and D. W. Stroock: Central limit phenomena for various interacting systems *Ann. Math.* **110** (1979) 333–393.
- [65] K. Iwata: An infinite dimensional stochastic differential equation with state space $C(\mathbb{R})$ *Prob. Th. Rel. Fields* **74** (1987) 141–159.
- [66] M. Kardar : Replica Bethe ansatz studies of two-dimensional interfaces with quenched random impurities *Nucl. Phys. B* **290**, 582–602 (1987).
- [67] M. Kardar, G. Parisi and Y.-C. Zhang: Dynamical scaling of growing interfaces *Phys. Rev. Lett.* **56** (1986) 889–892.
- [68] M. A. Katsoulakis and P. Souganidis: Generalized motion by mean curvature as a macroscopic limit of stochastic Ising models with long range interaction and Glauber dynamics *Comm. Math. Phys.* **169** (1995) 61–97.
- [69] M. A. Katsoulakis and P. Souganidis: Interacting particle systems and generalized motion by mean curvature *Arch. Rat. Mech. Anal.* **127** (1994)
- [70] C. Kipnis, S. Olla and S.R.S. Varadhan: Hydrodynamic and large deviations for simple exclusion process *Comm. Pure Appl. Math.* **42** (1989) 115–137.
- [71] A. J. Koch and H. Meinhardt: Biological pattern formation: from basic mechanisms to complex structures *Rev. Mod. Phys.* **66** (1994) 1481–1507.
- [72] J. Krug and H. Spohn: Kinetic roughening of growing surfaces in *Solids far from equilibrium* (1994) C. Godrèche editor.
- [73] L.D. Landau and E.M. Lifschitz: Fluid Mechanics, (1959) Pergamon Press Oxford, London, Paris Frankfurt
- [74] C. Landim, S. Olla and H. T. Yau: First-order correction for the hydrodynamic limit of asymmetric simple exclusion processes in dimension $d \geq 3$ *Comm. Pure Appl. Math.* **50** (1997) 149–203.
- [75] C. Landim and M.E. Vares: Exponential estimate for reaction diffusion models *Prob. Th. Rel. Fields* **106** (1996) 151–186.
- [76] C. Landim and H. T. Yau: Large deviations of interacting particle systems in infinite volume *Comm. Pure Appl. Math.* **48** (1995) 339–379.
- [77] J.S. Langer: An Introduction to the Kinetics of first-order phase transitions *Solids far from equilibrium* (1991) C. Godrèche ed., Cambridge Univ.
- [78] J.L. Lebowitz, E. Orlandi and E. Presutti: A particle model for spinodal decomposition *J. Stat. Phys.* **63** (1991) 933–974.
- [79] J. L. Lebowitz and O. Penrose: Rigorous treatment of the Van der Walls–Maxwell theory of the liquid–vapor transition *J. Math. Phys.* **7** (1966) 98.
- [80] J. L. Lebowitz and H. Spohn: Comment on “Onsager Reciprocity Relations without Microscopic Reversibility” by Gabrielli et. al. *Phys. Rev. Lett.* **78** (1997) 394–395.
- [81] T.M. Liggett: Interacting particle systems **276** (1985) Grundlehren der mathematischen Wissenschaften, Springer–Verlag, Berlin.

- [82] S. Maier-Paape and T. Wanner: Spinodal decomposition for the Cahn–Hilliard equation in higher dimensions. Part I: probability and wavelength estimate *Preprint* (1997).
- [83] S.A. Molchanov: Ideas in the theory of random media *Acta Appl. Math.* **22**, 139–282 (1991) and references therein.
- [84] C. Müller: On the support of solutions to the heat equation with noise *Stochastics and Stochastics Reports* **37** (1991) 225–245.
- [85] W. W. Mullins and R. F. Sekerka: Morphological stability of a particle growing by diffusion or heat flow *J. Appl. Phys.* **34** (1963) 323–329.
- [86] National Research Council Mathematical research in materials science: opportunities and perspectives *Board on Mathematical Sciences. National Academy Press, Washington D.C.* (1993).
- [87] R.L. Pego: Front migration in the nonlinear Cahn–Hilliard equation *Proc. R. Soc. Lond. A* **422** (1989) 261–278.
- [88] O. Penrose and J. L. Lebowitz: Rigorous treatment of metastable states in the Van der Walls–Maxwell theory *J. Stat. Phys.* **3** (1971) 235–247.
- [89] P. Pieruschka and S. Marcelja: Statistical mechanics of random bicontinuous phases *J. Phys. II France* **2** (1992) 211–236.
- [90] J. Quastel: Large deviations from a hydrodynamic scaling limit for a non-gradient system *Ann. Prob.* **23** (1995) 724–742.
- [91] F. Rezakhanlou: Hydrodynamic limit for attractive particle systems in \mathbb{Z}^d *Comm. Math. Phys.* **140** (1991) 417–448.
- [92] F. Rezakhanlou: Microscopic structure of shocks in one conservation laws *Ann. Inst. Henri Poincaré* **12** (1995) 119–153.
- [93] J. Smoller: Shock waves and reaction–diffusion equations **258** (1983) Grundlehren der mathematischen Wissenschaften. Springer–Verlag, Berlin.
- [94] H. Spohn: Interface dynamics in particle models *J. Stat. Phys.* **71** (1993) 1081–1131.
- [95] H. Spohn: Large scale dynamics of interacting particles *Texts and Monographs in Physics* (1991) Springer–Verlag, Berlin.
- [96] D.W. Stroock and S.R.S. Varadhan: Multidimensional diffusion processes (1979) Springer–Verlag (New York).
- [97] M. Teubner: Level surfaces of Gaussian random fields and micro-emulsions *Europhys. Lett.* **14**(5) (1991) 403–408.
- [98] J. Walsh: An introduction to stochastic partial differential equations *École d'Été de probabilités de Saint-Flour XIV, Lect. Notes in Math.* **1180** (1986).
- [99] H. T. Yau: Metastability of Ginzburg–Landau model with conservation law *J. Stat. Phys.* **74** (1994) 63–88.
- [100] H. T. Yau: Relative entropy and hydrodynamics of Ginzburg–Landau model *Lett. Math. Phys.* **22** (1991) 63–88.

CHAPTER 4

Transport by Incompressible Random Velocity Fields: Simulations & Mathematical Conjectures

RENE A. CARMONA
Statistics & Operations Research,
C.E.O.R. Princeton University
Princeton, NJ 08544
U.S.A.
rcarmona@princeton.edu

FREDERIC CEROU
INRIA & LATP
CMI – 39 rue Joliot-Curie,
13453 Marseille cedex 13
France
Frederic.Cerou@inria.fr

1. Introduction

Transport properties of stochastic flows have attracted the attention of investigators from various branches of science: physics of plasma, fluids mechanics and turbulence, porous media, functional and stochastic (mathematical) analysis, One of our objectives is to review recent works on the numerical simulations of the transport properties of 2D incompressible random velocity fields, to report on investigations in progress and to formulate and discuss mathematical conjectures suggested by the results of our numerical simulations.

Our motivation comes from the analysis of mathematical models of physical oceanography. If we consider for example the case of the temperature or the salinity of the ocean, these concentrations appear as solutions of linear parabolic equations driven by a velocity field and/or an external forcing term. These equations become stochastic partial differential equations (SPDE's for short) when the velocity field and/or the forcing term are regarded as random fields. Indeed, we do not attempt to solve the nonlinear equation giving the velocity field and instead we use the known chaotic behavior of the solutions of the Navier-Stokes equation and real data from observations of drifters at the surface of the ocean to justify and illustrate our choices of statistics for random models for the flows.

From the numerical point of view, the concentration SPDE's can be solved numerically by fixing a sample realization of the random coefficients (velocity and/or forcing term) and using standard deterministic methods such as finite elements, multi-grid, . . . and the like. Such an approach is very involved and can be cumbersome. See [27] and [46] for recent attempts in the one-dimensional case with white

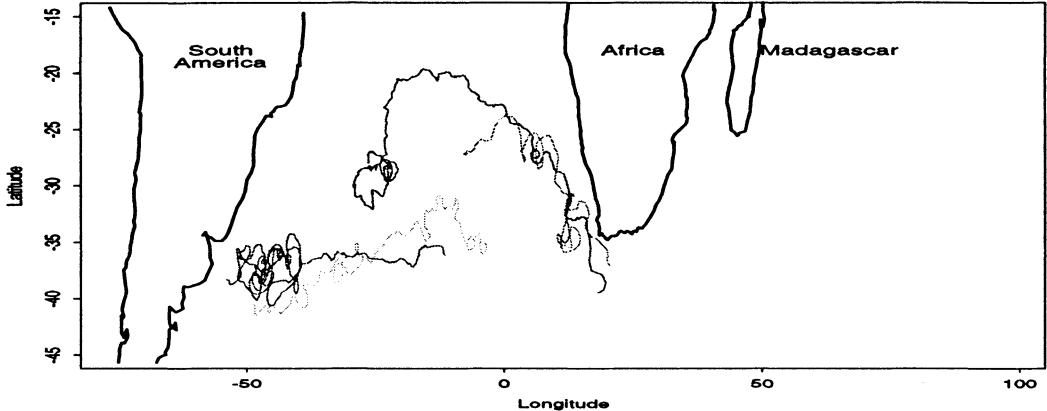


FIGURE 1. Example of drifter trajectories.

noise models. Also, it lacks the insight provided by a fine analysis of the properties of the randomness in the system given by the Lagrangian analysis of the flow. This is the point of view taken in this chapter. Since the numerical simulations of the transport properties of random velocity fields is one of our goals, this Lagrangian approach is more natural than the Eulerian approach tied to the SPDE.

Figure 1 shows the trajectories of several drifters at the surface of the ocean. The positions were recorded daily. They are real life instances of passive tracer transport. The sequences:

$$\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots$$

formed by the successive positions of a drifter are the result of the sampling of continuous time trajectories $t \mapsto \mathbf{x}(t)$, i.e.

$$\mathbf{x}_0 = \mathbf{x}(t_0), \mathbf{x}_1 = \mathbf{x}(t_1), \mathbf{x}_2 = \mathbf{x}(t_2), \dots$$

for times $t_0 = 0, t_1 = \Delta t, t_2 = 2\Delta t, \dots$ for some sampling interval $\Delta t > 0$. Saying that we are observing a passive tracer is merely saying that a given trajectory $t \mapsto \mathbf{x}(t)$ is a solution of the equation of motion:

$$(1.1) \quad d\mathbf{x}(t) = \vec{\mathbf{v}}(t, \mathbf{x}(t)) dt$$

where $\vec{\mathbf{v}}(t, \mathbf{x})$ denotes the velocity of the ocean at time t at the point \mathbf{x} . Note that for the sake of simplicity we did not include a viscosity term in the equation of motion. This will be done later.

Some of the statistical properties of these Lagrangian trajectories (i.e. the solutions of the equation of motion) have been investigated in [15] as bivariate time series and the consequences of the assumptions of time stationarity and space homogeneity of the models on the spectra of these series were tested.

We shall review part of the theoretical framework of the theory of random velocity fields with Gaussian statistics, we formulate precisely the mathematical assumptions of the physical models most widely accepted (incompressibility, Kolmogorov spectrum, ...) and we address the simulations issues. The present study

is intended for models where the external velocity field is *time dependent*. This excludes the stationary models for which the random velocity field is independent of time. Nevertheless we shall have many asides to discuss some of the recent results of Majda *et al* reported in [6], [7], [40], [33] and especially the papers [26] and [25] reporting on numerical computations. Even though most of these digressions are devoted to emphasize the differences between the time independent case and the models considered here, it is nevertheless clear that many of the ideas used in these works to deal with the space correlations and the spectral singularity are relevant to our considerations. In the time dependent case we review the massively parallel simulations reported in [13], the conjectures formulated and illustrated there and the theoretical results proved subsequently. In particular, we discuss the proof of the positivity of the upper Lyapunov exponent of the Jacobian flow given in [14] and the proof of the homogenization of the Lagrangian trajectories given in [17]. Finally, we discuss numerical estimates of the time evolution of the fractal dimension (Minkowski dimension to be specific) of the boundary of blobs of passive tracers transported by the flow and on the attempts to make the tracers feel the spectral singularity in the spectrum. Random velocity fields with non-Gaussian statistics have been considered in the literature. Most of the models which we know of are based on Poisson point processes. For the definition and the first properties of some of these models, the interested reader may consult for example [8] for the time independent case and [21] and [22] for the time dependent case. For these models, the time structure is still of the Ornstein-Uhlenbeck type and it is expected that most of the results proven for the Gaussian fields discussed in this chapter (strict positivity of the upper Lyapunov exponent of the Jacobian flow, exponential mixing, ...) will still hold. We end this chapter with a short discussion of the rationale behind these Poisson models.

All the numerical results reported in this chapter can be reproduced using a C - program available on the Web at the URL [19].

2. Gaussian Velocity Fields with Kolmogorov Spectra

We present the assumptions of a commonly used mathematical model of a random velocity field. Since the general considerations of this section do not depend upon the dimension of the space we shall consider velocity fields on the general Euclidean space \mathbb{R}^d . We restrict ourselves to the dimension $d = 2$ from Section 4 below on.

We assume that $\{\vec{v}(t, \mathbf{x}); t \geq 0, \mathbf{x} \in \mathbb{R}^d\}$ is a mean zero Gaussian vector field with values in \mathbb{R}^d . The hypothesis on the mean is chosen in order to concentrate our efforts on the understanding of the random fluctuations around a mean deterministic motion.

We assume that the velocity field is homogeneous in space. By this we mean that its distribution is invariant under the shifts of the space variable \mathbf{x} . In other words that, for each fixed $\mathbf{y} \in \mathbb{R}^d$:

$$\{\vec{v}(t, \mathbf{x} + \mathbf{y}); t \geq 0, \mathbf{x} \in \mathbb{R}^d\} \quad \text{and} \quad \{\vec{v}(t, \mathbf{x}); t \geq 0, \mathbf{x} \in \mathbb{R}^d\}$$

have identical distributions. We also assume that the random velocity field is stationary, i.e. its distribution is invariant under the shifts of the time variable t . This means that, for each fixed $s \geq 0$:

$$\{\vec{v}(s+t, \mathbf{x}); t \geq 0, \mathbf{x} \in \mathbb{R}^d\} \quad \text{and} \quad \{\vec{v}(t, \mathbf{x}); t \geq 0, \mathbf{x} \in \mathbb{R}^d\}$$

have identical distributions. These three assumptions (Gaussian statistics, homogeneity and stationarity) are quite common in the random mathematical models of turbulent fluids. See for example [47] or [9].

The distribution of the random field is completely characterized by its covariance tensor:

$$(2.1) \quad \mathbb{E}\{\vec{v}(s, \mathbf{x}) \otimes \vec{v}(t, \mathbf{y})\} = \Gamma(t-s, \mathbf{y}-\mathbf{x}), \quad s, t \geq 0, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^d.$$

If Φ is a random variable (or more generally a random vector) over the probability space of the velocity field, we use the notation $\mathbb{E}\{\Phi\}$ for its expectation. Notice that $\Gamma(t, \mathbf{x})$ is a $d \times d$ matrix. Its entries are given by the cross-covariances of the components of \vec{v} :

$$\Gamma_{j,\ell}(t, \mathbf{x}) = \mathbb{E}\{v_j(0, 0)v_\ell(t, \mathbf{x})\}, \quad j, \ell = 1, \dots, d.$$

2.1. Spectral Representation. The matrix $\Gamma(t, \mathbf{x})$ is the Fourier transform of a nonnegative definite matrix measure $\nu(d\omega, dk)$. As usual, we use ω to denote the frequency variable associated to the time variable t and we use the variable \mathbf{k} for the wave number variable associated to the space variable \mathbf{x} . We also assume that the entries of the covariance matrix are integrable. This guarantees the existence of densities $E_{j,\ell}(\omega, k)$ for the entries $\nu_{j,\ell}(d\omega, dk)$ of the matrix measure $\nu(d\omega, dk)$. In other words we have:

$$(2.2) \quad \Gamma(t, \mathbf{x}) = \int_{\mathbb{R}} \int_{\mathbb{R}^d} e^{i(\omega t + \mathbf{k} \cdot \mathbf{x})} E(\omega, \mathbf{k}) d\omega dk$$

for some matrix $E(\omega, \mathbf{k}) = [E_{j,\ell}(\omega, \mathbf{k})]_{j,\ell=1,\dots,d}$. Further assumptions on the spatial characteristics of the velocity field, such as *incompressibility* and *isotropy* can be used to specify further the covariance $\Gamma(t, \mathbf{x})$. This is usually done in terms of the form of its Fourier transform $E(\omega, \mathbf{k})$. See for example [9], [47] and/or [6]. The assumptions of isotropy and incompressibility essentially determine the form of the spatial part of the covariance. But it might also be desirable to build the structure of the time dependence and the time correlation into the model. In particular one may want to make the velocity field Markovian in time. In this case we get:

$$(2.3) \quad E_{j,\ell}(\omega, \mathbf{k}) = \frac{\beta(|\mathbf{k}|)}{\omega^2 + \beta(|\mathbf{k}|)^2} \mathcal{E}(|\mathbf{k}|) \frac{1}{|\mathbf{k}|^{d-1}} \left(\delta_{j,\ell} - \frac{k_j k_\ell}{|\mathbf{k}|^2} \right),$$

for some scalar functions $r \mapsto \beta(r)$ and $r \mapsto \mathcal{E}(r)$ of $r > 0$. The type of function $\mathcal{E}(r)$ used in fluid mechanics varies depending on the applications. Kolmogorov's theory of turbulence suggests a scale invariance which is brought into the model by the choice of a power function $\mathcal{E}(r) = r^{1-\epsilon}$ for some real parameter ϵ . We shall restrict ourselves to the interesting case $\epsilon > 0$. The classical Kolmogorov spectrum is associated to the value $\epsilon = 5/3$ in $d = 2$ dimensions. Since a power function cannot be integrable both at the origin and at infinity, this choice will lead to singularities which play a crucial role in the properties of the system and especially in the transport properties of such fluids.

Mathematical regularization of the singularities is usually achieved by the introduction of cut-off functions. In other words, the original function $\mathcal{E}(r)$ is replaced by a regularized form:

$$(2.4) \quad \mathcal{E}^{(reg)}(r) = \mathcal{E}(r)\chi_0(r)\chi_\infty(r)$$

where the infrared cut-off function $\chi_0(r)$ vanishes near the origin and the ultra-violet cut-off function $\chi_\infty(r)$ vanishes near ∞ . When the function $\mathcal{E}(r)$ is chosen to be zero near the origin (i.e. for $r \leq r_0$ for some $r_0 > 0$) and near infinity (i.e. for $r \geq r_1$ for some $r_1 < \infty$) we shall say that the velocity field is regular. The value of the parameter ϵ controls the spatial correlation in the velocity field. In the classical Kolmogorov's theory, the numbers r_0 and r_1 are related to the so-called integral and dissipation scales respectively.

The function $\beta(r)$ is usually chosen of the form:

$$\beta(r) = r^z$$

for some $z > 0$. In this way, the parameter z controls the time correlation in the velocity field. As a consequence, a Kolmogorov's spectrum can be characterized by the two parameters z and ϵ . Finally we recall the spectral representation formula:

$$(2.5) \quad \vec{v}(t, \mathbf{x}) = \int_{\mathbb{R}} \int_{\mathbb{R}^d} e^{i(\omega t + \mathbf{k} \cdot \mathbf{x})} E(\omega, \mathbf{k})^{1/2} W(d\omega, d\mathbf{k}).$$

which writes the random velocity field in terms of a complex vector white noise measure (in the L^2 -sense) $W(d\omega, d\mathbf{k})$.

3. Abstract Ornstein Uhlenbeck Velocity Fields

In order to motivate the introduction of Ornstein Uhlenbeck (O-U for short) processes we revisit formula (2.5) giving the spectral representation of the velocity field. Performing the integration with respect to ω in (2.2) gives:

$$\Gamma(t, \mathbf{x}) = \int_{\mathbb{R}^d} e^{i\mathbf{k} \cdot \mathbf{x}} E(t, \mathbf{k}) d\mathbf{k}$$

with:

$$(3.1) \quad E_{j,\ell}(t, \mathbf{k}) = e^{-|\mathbf{k}|^z|t|} \mathcal{E}(|\mathbf{k}|) \frac{1}{|\mathbf{k}|^{d-1}} \left(\delta_{j,\ell} - \frac{k_j k_\ell}{|\mathbf{k}|^2} \right).$$

In other words, the covariance is in the Fourier domain the superposition of the covariances of O-U processes and it is natural to conjecture that the spectral representation (2.5) can be rewritten in the form:

$$(3.2) \quad \vec{v}(t, \mathbf{x}) = \int_{\mathbb{R}^d} e^{i\mathbf{k} \cdot \mathbf{x}} \tilde{E}(\mathbf{k})^{1/2} \xi_t(d\mathbf{k}).$$

where the $\xi_t(d\mathbf{k})$ are, for each fixed t , orthogonal (independent) increment L^2 -measures in the Fourier domain and for each fixed infinitesimal $d\mathbf{k}$, OU processes in the time variable. This fact can be proven rigorously using the theory of L^2 martingale measures as developed in [46]. But most importantly, the very form of the representation (3.2) shows that, if we look at $\vec{v}(t, \mathbf{x})$ as a function of \mathbf{x} , say $\vec{v}(t) = \vec{v}(t, \cdot)$ for each fixed time t , then the process $\{\vec{v}(t); t \geq 0\}$ should be regarded as a form of Ornstein Uhlenbeck process with values in a space of

functions of \mathbf{x} , possibly of infinite dimensions since it appears as the result of a linear superposition (i.e. an integral) of many such processes.

3.1. The Abstract Set-Up. Let A and B be two nonnegative self-adjoint operators on a separable Hilbert space H . For the sake of simplicity we shall assume that the operators A and B commute and for each $t \geq 0$ we consider the self-adjoint operator:

$$\Gamma(t) = B(2A)^{-1}e^{-tA}.$$

Even though e^{-tA} is a contraction, the operator $\Gamma(t)$ may be unbounded and delicate domain problems may occur. Nevertheless, we shall keep the present discussion to a formal level and we shall ignore these domain questions until we get to the specific examples discussed in the next subsection. The nonnegative operator $\Gamma(0) = B(2A)^{-1}$ can be viewed as the covariance operator of a mean zero (possibly only additive) Gaussian measure on H . Enlarging H if necessary, this measure can be extended into a σ -additive bona-fide measure, say ν . Let us denote by \tilde{H} such an enlargement of H . The operators $\Gamma(t)$ can then be used to construct a \tilde{H} -valued mean zero Gaussian process, say $\{\mathbf{v}(t); t \geq 0\}$, such that the covariance of $\mathbf{v}(s)$ and $\mathbf{v}(t)$ is given by the operator $\Gamma(|t - s|)$. This statement will be made more precise in the following subsection when we discuss specific examples. Specific assumptions and gory details of this abstract construction are given in [4] where it is also shown that it is possible to view the process as the solution of a stochastic differential equation (in an infinite dimensional space) of the standard form:

$$(3.3) \quad d\mathbf{v}(t) = -A\mathbf{v}(t)dt + dW_B(t)$$

where $\{W_B(t); t \geq 0\}$ is a process of Brownian motion with covariance B in the sense that the operator B is the covariance operator of the (infinite dimensional) random variable $W_B(1)$. This process of Brownian motion does not need to be taking its values in the space \tilde{H} as long as its values are mapped back into this space by the operator semigroup e^{-tA} as formula (3.4) shows. In any case, this infinite dimensional O-U process also has the following integral representation:

$$(3.4) \quad \mathbf{v}(t) = \mathbf{v}(0) + \int_0^t e^{-(t-s)A}dW_B(s)$$

where the integral is given the usual sense of the classical theory of the Wiener integral except that we are dealing here with an operator valued integrand and a function valued integrator.

3.2. O-U Velocity Fields. We now consider specific instances of the general model described above and we show how this framework includes the incompressible Gaussian velocity fields with Kolmogorov spectra.

The first example we have in mind is the following. $H = L^2(\mathbb{R}^d, dx)$ is the classical Hilbert space of (equivalent classes) of square integrable complex functions on the Euclidean space \mathbb{R}^d , $A = (-\Delta)^{z/2}$ is a power of the negative Laplacian (at this stage z can be any nonnegative real number) and $B = \mathcal{E}((-(-\Delta)^{1/2}))$ is a specific nonnegative function of the same self-adjoint operator $-\Delta$. The fact that both A and B appears as functions of the same self-adjoint operator ensures that A and B commute. Specific properties of the nonnegative function $r \mapsto \mathcal{E}(r)$ on the half-axis $(0, \infty)$ will be discussed later but we use the same notation $\mathcal{E}(r)$ as earlier for

obvious reasons: this function is very much related to the functions \mathcal{E} and $\mathcal{E}^{(reg)}$ introduced earlier and to the functions \mathcal{E}_ϕ and $\mathcal{E}_\phi^{(reg)}$ discussed below.

To identify further the infinite dimensional O-U process $\mathbf{v}(t)$ constructed above to the velocity fields with Kolmogorov spectra, we assume that the spectral function $\mathcal{E}(r)$ is bounded, has bounded support and that this support is bounded away from 0. This is the case for the function $\mathcal{E}_\phi^{(reg)}(r)$ introduced in the next section to analyze the stream function in dimension $d = 2$. In this case, B is a bounded operator on H with a spectrum bounded away from 0 (and in particular with a bounded inverse.) More importantly, the compactness of the support makes it possible to choose \tilde{H} as a space of analytic functions. If we then denote by $\mathbf{v}(t, \mathbf{x})$ the value of the function $\mathbf{v}(t)$ resulting from the construction outlined in the previous subsection, at the point $\mathbf{x} \in \mathbb{R}^d$, then one has:

$$\mathbb{E}\{\mathbf{v}(s, \mathbf{x})\mathbf{v}(t, \mathbf{y})\} = \Gamma(t - s, \mathbf{y} - \mathbf{x}) = [\mathcal{E}((- \Delta)^{1/2})e^{-|t-s|\Delta^{z/2}}](x - y)$$

if we use the notation $\Gamma(t, \cdot)$ for the integral kernel of the operator $\Gamma(t)$. If we now go to the Fourier domain to express the kernel of this function of the negative Laplacian we get:

$$\mathbb{E}\{\mathbf{v}(s, \mathbf{x})\mathbf{v}(t, \mathbf{y})\} = \int_{\mathbb{R}^d} e^{i\mathbf{k} \cdot (\mathbf{y} - \mathbf{x})} \mathcal{E}(|\mathbf{k}|) e^{-|t-s||\mathbf{k}|^z} d\mathbf{k}$$

which is of the desired form (with $\beta(r) = r^z$) if we use the Fourier transform in time to express the exponential $e^{-|t-s||\mathbf{k}|^z}$. The field $\mathbf{v}(t, \mathbf{x})$ so defined (as well as its covariance function $\Gamma(t, \mathbf{x})$) is scalar so strictly speaking, this construction applies only to the two dimensional case of the stream function $\phi(t, \mathbf{x})$ introduced and analyzed in the next section. In the general case, one needs to consider a Hilbert space H of vector valued functions (by tensor products for example) and have the operators A and B act appropriately. These operators can be chosen as functions of the negative Laplacian as well and the construction of the infinite dimensional O-U process $\mathbf{v}(t)$ can be done in this case also and one can then recover a vector random field $\mathbf{v}(t, \mathbf{x})$ which has a spectrum of the same structure as the velocity field $\tilde{\mathbf{v}}(t, \mathbf{x})$ we started from in the introduction.

4. Simulation of the Velocity Field

From now on we restrict ourselves to the 2 - dimensional case. Because of the incompressibility assumption, this restriction implies that the velocity field can be derived from a *stream function* $\phi(t, \mathbf{x})$ via the formula:

$$\tilde{\mathbf{v}}(t, \mathbf{x}) = \begin{bmatrix} \frac{\partial \phi(t, \mathbf{x})}{\partial x_2} \\ -\frac{\partial \phi(t, \mathbf{x})}{\partial x_1} \end{bmatrix}.$$

The scalar random field $\{\phi(t, \mathbf{x}); t \geq 0, \mathbf{x} \in \mathbb{R}^2\}$ can be chosen to be stationary in time and homogeneous in space. In fact a formula similar to (3.2) can be derived and the stream random field can be used in the form (at least in distribution):

$$(4.1) \quad \phi(t, \mathbf{x}) = \int_{\mathbb{R}^2} e^{i\mathbf{k} \cdot \mathbf{x}} \sqrt{\tilde{E}_\phi(\mathbf{k})} \xi_t(d\mathbf{k}).$$

where the function $\tilde{E}_\phi(\mathbf{k})$ is now a scalar function instead of being a matrix valued function. This "special spectral representation" can be undone to recover the

standard spectral representation of the stream function. We get:

$$(4.2) \quad \phi(t, \mathbf{x}) = \int_{\mathbb{R}^2} e^{i\mathbf{k} \cdot \mathbf{x}} \sqrt{E_\phi(\omega, \mathbf{k})} W' d\omega d\mathbf{k}$$

where $W'(d\omega, d\mathbf{k})$ is a scalar valued orthogonal L^2 white noise measure and where the (scalar) spectral density $E_\phi(\omega, \mathbf{k})$ is of the form:

$$(4.3) \quad E_\phi(\omega, \mathbf{k}) = \frac{\beta(|\mathbf{k}|)}{\omega^2 + \beta(|\mathbf{k}|)^2} \mathcal{E}_\phi(|\mathbf{k}|)$$

in full analogy with (2.3). Simple arithmetic gives:

$$(4.4) \quad \mathcal{E}_\phi(r) = r^{-3} \mathcal{E}(r)$$

if $\mathcal{E}(r)$ stands for the spectral function introduced in (3.1).

4.1. Abstract Formalism. We first present the general form of the simulation algorithm. Because we plan to simulate the Ornstein Ulhenbeck process at successive times, we take advantage of the Markov property and we derive a formula to update the process. Coming back to the integral representation (3.4) we see that, whenever t and h are nonnegative we have:

$$\begin{aligned} \phi(t+h) &= e^{-hA} \left(e^{-tA} \phi(0) + \int_0^t e^{-(t-s)A} W_B(ds) \right) + \int_t^{t+h} e^{-(t+h-s)A} W_B(ds) \\ &= e^{-hA} \phi(t) + \int_t^{t+h} e^{-(t+h-s)A} W_B(ds). \end{aligned}$$

Not only this formula proves the Markov property (in time) of the process but it also provides a simple formula for the transition probability of the process. Indeed, the conditional distribution of $\phi(t+h)$ given the past \mathcal{F}_t of the process up to time t is the Gaussian distribution with mean $e^{-hA} \phi(t)$ and covariance equal to the covariance of the stochastic (Wiener) integral:

$$(4.5) \quad \int_t^{t+h} e^{-(t+h-s)A} W_B(ds).$$

Given the definition of the process W_B , the latter is given by:

$$(4.6) \quad \Gamma_h(f, g) = \int_0^h \langle B e^{-sA} f, g \rangle ds$$

where the notation $\langle \cdot, \cdot \rangle$ is used for the H -inner product. Hence the update procedure can be summarized in the following way:

- generate $\phi(0)$ with distribution ν (i.e. as a mean zero Gaussian random variable with covariance:

$$(4.7) \quad \Gamma_0(f, g) = \langle B(2A)^{-1} f, g \rangle.$$

- given a realization of $\phi(t)$ generate a realization of $\phi(t+h)$ as a Gaussian random variable with mean

$$(4.8) \quad \mu_{t+h} = e^{-hA} \phi(t)$$

and covariance Γ_h given in (4.6).

REMARK 4.1. The second bullet deals with the update of the stream function (and hence of the velocity field) as time evolves. Obviously, this step is not present in the analysis of time independent velocity fields. But as we shall see later, this is not the only reason simulations are much easier to set up in the time independent case.

4.2. Time Discretization and Expansion in a Basis. The idea is to discretize the time into a sequence $(t_n)_{n=0,1,\dots}$ and to use the Markov property to construct the next value $\vec{v}(t_{n+1})$ from the current value $\vec{v}(t_n)$ ignoring the way the field "got there" i.e. ignoring the values of $\vec{v}(t_0), \vec{v}(t_1), \dots, \vec{v}(t_{n-1})$. Moreover, since each $\vec{v}(t)$ is itself an infinite dimensional object, we shall need to use other finite dimensional approximations to construct the configurations $\vec{v}(t)$.

The discretization of time is straightforward. We choose a time interval Δt and we sample the continuous time by restricting its values to:

$$t_0 = 0, t_1 = \Delta t, t_2 = 2\Delta t, \dots, t_n = n\Delta t, \dots \dots$$

Because of the Markov property and because of our choice for the discretization of the time, at each instant t_j , we will use the knowledge of the characteristics of the field $\vec{v}(t_{j-1})$ at the preceding time to generate the samples of the velocity field $\vec{v}(t_j, \mathbf{x})$ at the locations \mathbf{x} where we need them. Notice that the choice of the time interval Δt has to depend upon the spectral characteristics, and especially of the radial lower bound r_0 on the support of the spectral density. This remark will be especially useful when we simulate velocity fields for different values of the spectral parameters ϵ and z . See the discussion of the renormalization problems in the last section below for details.

We now consider a more precise form of the simulation algorithm, still without paying much attention to the actual convergence of the formal expansions we use. Since we restrict ourselves to dimension $d = 2$ we find it more convenient to work with the stream function.

Let $(\varphi_n)_{n \geq 1}$ be a complete orthonormal system in H and let us expand the O-U process $\{\phi(t); t \geq 0\}$ as if it were an element of the Hilbert space. At each time t we want to take advantage of the expansion:

$$(4.9) \quad \phi(t) = \sum_{n=0}^{\infty} \langle \phi(t), \varphi_n \rangle \varphi_n.$$

Our goal is to replace the sum of the infinite series (4.9) is replaced by a finite sum. In particular we simulate the initial value $\phi(0)$ as:

$$(4.10) \quad \tilde{\phi}(0) = \sum_{n=0}^{N_0} \langle \phi(0), \varphi_n \rangle \varphi_n = \sum_{n=0}^{N_0} \xi_n^{(0)} \varphi_n$$

provided we set $\xi_n^{(0)} = \langle \phi(0), \varphi_n \rangle$. These random variables form a Gaussian sequence of mean zero random variables. Their covariance matrix is given by the

matrix elements:

$$(4.11) \quad \gamma^{(0)}(m, n) = \mathbb{E}\{\xi_m^{(0)} \xi_n^{(0)}\} = \langle B\varphi_m, \varphi_n \rangle.$$

In particular, the sequence $\xi^{(0)} = \{\xi_n^{(0)}; n = 0, \dots, N_0\}$ is independent if the orthonormal basis $\{\varphi_n\}_n$ is formed of eigenvectors of the self-adjoint operator B . In this case the variance of $\xi_n^{(0)}$ is merely the n -th eigenvalue (as usual counting multiplicity) i.e. the eigenvalue corresponding to the eigenfunction φ_n , and the generation of the $\xi_n^{(0)}$'s is very simple if these eigenvalues are known. Otherwise, the *matrix elements* $\langle B\varphi_m, \varphi_n \rangle$ have to be computed, the square root of the matrix $\gamma^{(0)} = [\gamma^{(0)}(m, n)]_{m, n=1, \dots, N_0}$ has to be computed and the vector $\xi^{(0)}$ of dimension $(N_0 + 1)$ can be generated as:

$$\xi^{(0)} = \sqrt{\gamma^{(0)}} \eta^{(0)}$$

where $\eta^{(0)}$ is a $(N_0 + 1)$ -dimensional vector of i.i.d. $N(0, 1)$ random variables.

Let us now consider the problem of the update of the process. Let us assume that the generation of the process at time $t = j\Delta t$ has been done according to the formula:

$$\tilde{\phi}(j) = \sum_{n=0}^{N_j} \xi_n^{(j)} \varphi_n$$

for some sequence $\{\xi_n^{(j)}; n = 0, \dots, N_j\}$ of real numbers. Obviously, this sequence forms a realization of a Gaussian sequence with a specific mean $\mu^{(j)}$ and covariance matrix $\gamma^{(j)}$. According to the discussion of the previous section, the next value $\tilde{\phi}(j+1)$ of the stream function at time $t_{j+1} = (j+1)\Delta t$ is constructed as the sum of a (conditional) mean $\mu^{(j+1)}$ and a fluctuating part. First we compute its mean. We choose the approximation of $e^{-(\Delta t)A} \phi(n\Delta t)$ obtained by replacing $\phi(n\Delta t)$ by the element of H whose components in the $\{\varphi_n\}_n$ basis are given by the components of the vector $\mu^{(j)}$. In other words we use the mean:

$$(4.12) \quad \begin{aligned} \mu^{(j+1)} &= e^{-(\Delta t)A} \sum_{n=0}^{N_j} \xi_n^{(j)} \varphi_n \\ &= \sum_{m=0}^{N_{j+1}} \left(\sum_{n=0}^{N_j} \langle e^{-(\Delta t)A} \varphi_n, \varphi_m \rangle \xi_n^{(j)} \right) \varphi_m \end{aligned}$$

In order to compute the fluctuating part we compute the matrix elements of the covariance operator $\Gamma_{\Delta t}$ on the basis elements φ_n (recall formula (4.6)).

$$\gamma^{(j+1)}(m, n) = \int_0^{\Delta t} \langle Be^{-sA} \varphi_m, \varphi_n \rangle ds$$

and we proceed as when we generated $\xi^{(0)}$: we first compute the square root of the matrix $\gamma^{(j+1)}$, then we generate i.i.d. standard Gaussian random variables and we multiply the vector so-obtained by the square root of $\gamma^{(j+1)}$.

A naive approach to the discretization of the space variable could be to use expansions into "bases of delta functions." In the same way we replaced the continuous time t by a grid of discrete values, it is possible to replace the space variable $\mathbf{x} \in \mathbb{R}^2$ by a discrete variable taking finitely many values in a lattice $h\mathbb{Z}^2$ for some mesh $h > 0$. This strategy has several drawbacks, especially from the point of view

of Monte-Carlo simulation. Indeed, for each fixed time t , there is a strong dependence between the values of $\vec{v}(t, \mathbf{x})$ and $\vec{v}(t, \mathbf{y})$ even when the points \mathbf{x} and \mathbf{y} are far apart, making the generation of the configuration $\vec{v}(t_j)$ from $\vec{v}(t_{j-1})$ extremely costly in memory storage and in computing time: this approach is not reasonable for our application.

Most importantly, the transport simulations which we discuss below require updates of passive tracers at each time step. These updates force us to compute the velocity field at the positions of the tracers and the latter do not have the good taste of staying on the points of the lattice $h\mathbb{Z}^2$. An artificial extrapolation would be needed in order to get the values of the velocity field necessary for the updates and we shall refrain from doing that.

4.3. Wavelet Expansions. As we have already emphasized, the initialization of the simulation process for the velocity field is the computation of the partial expansion (4.9):

$$\tilde{\phi}(0) = \sum_{n=0}^{N_0} \xi_n^{(0)} \varphi_n$$

where the φ_n 's form a basis of the space and the $\xi_n^{(0)}$'s form a mean zero Gaussian sequence with a covariance matrix given by the matrix elements of the operator B in the basis $\{\varphi_n\}_n$. At this stage, any orthonormal system would do essentially as well. In the time - independent case considered in [26] and [25] wavelet like bases were used. This last paper contains blunt propaganda for the choice of Meyer's wavelet. Given the fact that the authors consider a shear flow model (and hence a scalar field parameterized by a one-dimensional space variable) this choice should be obvious if one keeps in mind the fact that the mean square error $\mathbb{E}\{|v(0) - \tilde{v}(0)|^2\}$ is essentially the integral over the wave number space of the square difference between the spectral density and the Fourier transform of the partial expansion in the basis. Because of the special form of Meyer's wavelet and the fact that one can explicitly control the effects of scaling and translation in the Fourier domain, it is reasonable to consider paying the price of the construction of the correlation structure in the coefficients in the decomposition. But most importantly, the attractiveness of the wavelet decomposition is due to the presence of the spectral singularity of the model. Indeed, since all the elements of a wavelet basis *vanish* at the origin in the Fourier domain, replacing the infinite sum of (4.9) by the finite sum (4.10) kills two birds with one stone. Not only does the wavelet decomposition provide a *spatial discretization* suitable for numerical implementation but it also provides an algorithmic way to implement the spectral cut - off: this seems to be the ideal way to feel the effect of the spectral singularity. Majda and his students have used this idea to compute the structure function:

$$(4.13) \quad \mathbb{E}\{[v(x) - v(y)]^2\}$$

in the case of a scalar time independent Gaussian process $\{v(x); x \in \mathbb{R}\}$ of the fractional Brownian type which they use for the nontrivial component of a shear flow velocity field. A simple Gaussian computation shows that the structure function can be computed without the cut-off and that it is a specific power of $|x - y|$, namely:

$$(4.14) \quad \mathbb{E}\{[v(x) - v(y)]^2\} = c|x - y|^{2h-1},$$

if the spectral density of the scalar mean zero Gaussian (generalized) process is $|k|^{-3h}$. These authors use an expansion in the Meyer's wavelet basis to recover numerically this asymptotic with Monte Carlo computations of the expectation appearing in the definition (4.13) of the structure function. Another very appealing feature of the structure function computation is that *one knows ahead of time where the velocity field will have to be computed* and this can be taken advantage of in the setting up of the computations. Indeed, knowing x and y (by homogeneity, the computation is in fact done for 0 and $y - x$) one can determine, for each resolution the set of wavelet coefficients which will not be zero and limit the generation of the Gaussian random variable to these coefficients. The effect of this innocent looking remark is to increase the efficiency and the precision of the computations.

It is conceivable to consider the generalization of the numerical computation of the structure function to the more general case of isotropic flows. The Meyer's wavelet basis will have to be replaced by a basis of functions on the two-dimensional plane, and using tensor products is a natural way to construct such functions. Even though the isotropy will have to be lost in such an expansion, we believe that this generalization is possible. But our main concern is the implementation of these ideas in the time - dependent case.

The level of complexity of the generation of samples from the velocity field in the time - independent case is the same as the level of complexity of the generation of samples of $\phi(0)$ in the time dependent case. But in the time - dependent case, there is a enormous computing overhead due to the update of the field as time evolves and especially for building the time and space correlations. The general discussion of the previous section gives the general framework and it seems reasonable to try to implement it as long as the matrix elements of the operators B and $e^{-(\Delta t)A}$ can be computed in the *wavelet* basis. Formula (4.12) giving the update of the mean of the stream function is amenable to wavelet expansions because since (Δt) is small, one can use the approximation $e^{-(\Delta t)A} \approx I - (\Delta t)A$ to argue that the matrix representing the operator $e^{-(\Delta t)A}$ in the wavelet basis will be sparse (recall that both A and B are functions of the negative Laplacian $-\Delta$.)

We refrained from doing so because we are not only interested in the asymptotics of the Eulerian structure function, but we are mostly interested in the transport properties of the flow. As we will see shortly, it is impossible to know exactly in advance where the velocity field will have to be computed (i.e. where the tracer particles will be.) Given a finite time horizon, it is possible to estimate *with high probability* where *most* of the particles will be, but it is impossible to do that with certainty. Only in the case of a shear flow can this be done when the diffusivity κ is zero. This is the particular situation considered in [25] to compute the mean square displacement of two particles. But unfortunately, short of being able to control this set it is not possible to limit the set of wavelet coefficients which will enter significantly in the computation and the latter become infeasible.

In contrast, the Fourier expansions described below have enough built in independence for the computational burden to be minimal for regular models, and in particular when one does not attempt to feel the effect of the singularity. But before we give up on wavelet expansions, we would like to identify one more of the subtle pitfalls of this arena. Given a finite set of tracers to transport and a finite time horizon, it seems reasonable to work out the update of the statistical correlations of the velocity field in the following way: at each time step one uses

the knowledge of the positions of the tracers to determine which of the coefficients of the wavelet expansion will be significant and one restricts the random generation of the coefficients to these basis components. It seems reasonable to expect that a very meticulous book-keeping should make it possible to track these basis coefficient labels. But unfortunately, running the random generation from them would lead to the Monte Carlo simulation of some sort of Lagrangian distribution of the velocity field in lieu of the desired Eulerian statistics!

4.4. Fourier Expansions. This section is devoted to the description of the simulation of a random velocity field of the Kolmogorov type when cut-off functions have been used to remove the infrared and ultraviolet singularities of the spectrum. In this case the infinite dimensional Ornstein Uhlenbeck process is regular in the sense that the operator B is bounded, its spectrum is bounded away from zero and as a consequence, the operator A can be treated as bounded. In the same way we considered expansions in "bases of delta functions" in the space domain, one can contemplate expansions in "bases of delta functions" in the Fourier domain. This approach is not restrictive for regular fields and it is very much in the spirit of the spectral theory of stationary and/or homogeneous random fields, the simplest approach is to approximate the spectral measure $\nu(d\omega, dk)$ by simpler measures more amenable to random simulations on a computer. In particular, it is natural to approximate ν by a finite sum of point masses $\delta_{(\omega_i, k_i)}$ in the frequency-wave number domain. This approach is especially useful because it makes it possible to simulate directly the white noise without any concern with the possible dependence of the random variates to generate. Nevertheless, we refrain from using this approach directly, especially in the time variable. Indeed, in order to save memory (and computing time) we want to use an adaptive approach generating new values of the velocity field as we need them. More precisely, we want an approximation scheme which uses the fact that $\bar{v}(t)$ is a time homogeneous Markov process. For this reason we shall separate the problems of the discretization of the time and the space variables.

Instead we shall use a strategy based on the "nonstandard spectral representation formula" (4.1). We choose a finite partition of the support of the function $\mathcal{E}_\phi(|\mathbf{k}|)$, say $\mathcal{D} = \{D_\mathbf{k}; \mathbf{k} \in \mathcal{K}\}$ where the parameter \mathcal{K} is viewed as a set of wave vectors in \mathbb{R}^2 (typically we shall make sure that $\mathbf{k} \in D_\mathbf{k}$.) We then construct an approximation in the form:

$$(4.15) \quad \phi_K(t, \mathbf{x}) = \sum_{\mathbf{k} \in \mathcal{K}} e^{i\mathbf{k} \cdot \mathbf{x}} e(\mathbf{k}) \xi_t(D_\mathbf{k})$$

where the $e(\mathbf{k})$'s are coefficients obtained by taking the square root of the integral of the function $\mathcal{E}_\phi(|\mathbf{k}|)$ over the patch $D_\mathbf{k}$. Notice that the $\{\xi_t(D_\mathbf{k}); t \geq 0\}$ are independent Ornstein Uhlenbeck processes. The choice of the partition \mathcal{D} , of the set \mathcal{K} of wave vectors and of the corresponding approximate spectral coefficients $e(\mathbf{k})$ have a strong influence on the qualities of the approximation of the stream function and consequently on the velocity field itself and the corresponding transport properties of the flow. Nevertheless, except for the particular case of a set \mathcal{K} contained in a line, most of the qualitative properties which we want to illustrate by numerical simulations are stable when one varies these choices.

Formula (4.15) is written with the continuous variable t , but as we explained earlier the time variable will be sampled on a regular grid. The effect of the sampling

will just be to turn the continuous time Ornstein Uhlenbeck processes into discrete time autoregressive processes of order 1. See below for details.

Simple arithmetic on formula (4.15) leads to the real representation:

$$(4.16) \quad \phi(t, \mathbf{x}) = \sum_{\mathbf{k} \in \mathcal{K}} [a_{\mathbf{k}}(t) \cos(\mathbf{k} \cdot \mathbf{x}) + b_{\mathbf{k}}(t) \sin(\mathbf{k} \cdot \mathbf{x})]$$

where the $\{a_{\mathbf{k}}(t); t \geq 0\}$ and the $\{b_{\mathbf{k}}(t); t \geq 0\}$ are independent scalar Ornstein Uhlenbeck processes. As explained below, the latter become auto regressive processes after discretization of the time using the simplest Euler scheme. The components $v_1(t, \mathbf{x})$ and $v_2(t, \mathbf{x})$ of the velocity vector $\vec{\mathbf{v}}(t, \mathbf{x})$ can then be computed according to the formulas:

$$(4.17) \quad v_1(t, \mathbf{x}) = \sum_{\mathbf{k} \in \mathcal{K}} k_2 [a_{\mathbf{k}}(t) \sin(\mathbf{k} \cdot \mathbf{x}) - b_{\mathbf{k}}(t) \cos(\mathbf{k} \cdot \mathbf{x})]$$

and:

$$(4.18) \quad v_2(t, \mathbf{x}) = \sum_{\mathbf{k} \in \mathcal{K}} k_1 [-a_{\mathbf{k}}(t) \sin(\mathbf{k} \cdot \mathbf{x}) + b_{\mathbf{k}}(t) \cos(\mathbf{k} \cdot \mathbf{x})].$$

As formula (5.2) shows, we shall need to compute the velocity field at the instants t_j at the positions \mathbf{x}_{j-1} of the tracer particle. But notice that, once the program has knowledge of the auto-regressive processes $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$ labeled by the set \mathcal{K} of wave vectors, the velocity field can be computed at **any** point \mathbf{x} by mere evaluation of the trigonometric functions sin and cos for the real argument $\mathbf{k} \cdot \mathbf{x}$. The updates of the auto regressive coefficients are done via the formulas:

$$(4.19) \quad a_{\mathbf{k}}(t_j) = \alpha_{\mathbf{k}} a_{\mathbf{k}}(t_{j-1}) + \sigma_{\mathbf{k}} \epsilon_{\mathbf{k}}^{(a)}(j)$$

and:

$$(4.20) \quad b_{\mathbf{k}}(t_j) = \beta_{\mathbf{k}} b_{\mathbf{k}}(t_{j-1}) + \sigma_{\mathbf{k}} \epsilon_{\mathbf{k}}^{(b)}(j)$$

where the $\epsilon_{\mathbf{k}}^{(a)}(j)$ and the $\epsilon_{\mathbf{k}}^{(b)}(j)$ are independent standard normal random variables and where the constants $\alpha_{\mathbf{k}}$, $\beta_{\mathbf{k}}$ and $\sigma_{\mathbf{k}}$ can be computed from the spectral density of the stream function:

$$\sigma_{\mathbf{k}}^2 = |\mathbf{k}|^{-2-\epsilon} \Delta t, \quad \alpha_{\mathbf{k}} = \beta_{\mathbf{k}} = 1 - |\mathbf{k}|^z \Delta t \approx e^{-|\mathbf{k}|^z \Delta t}.$$

The auto-regressive approximation based on the Euler scheme was used in [13]. In many applications we need to compute very long trajectories and stability problems arise when using Euler schemes. In the case of OU processes, we prefer the use of explicit formulas (see expression (4.5) and the text above) for the mean and covariance of the increments, which do not require more computation than Euler scheme. This alternative was used to produce the numerical results reported in this chapter.

REMARK 4.2. Let us write $\mathcal{K} = \cup_{j=1}^M \mathcal{K}_j$ where \mathcal{K}_j is a subset of a line and let us denote by \mathbf{k}_j a unit vector on this line. One can then rewrite the formulae (4.17)

and (4.18) by summing first over j and one gets an expression of the form:

$$(4.21) \quad \vec{v}(t, \mathbf{x}) = \sum_{j=1}^M v_j(t, \mathbf{k}_j \cdot \mathbf{x}) \mathbf{k}_j^\perp$$

where the scalar random fields $\{v_j(t, x) : t \geq 0, x \in \mathbb{R}\}$ are defined by:

$$(4.22) \quad v_j(t, x) = \sum_{k \in \mathbb{R}, k \mathbf{k}_j \in \mathcal{K}_j} k [a_{k \mathbf{k}_j}(t) \sin(kx) - b_{k \mathbf{k}_j}(t) \cos(kx)].$$

In other words, formula (4.21) shows that the velocity field $\vec{v}(t, \mathbf{x})$ can be viewed as an additive superposition of random shears in the directions of the \mathbf{k}_j 's. This is the point of view advocated in [40] for the time independent case.

4.5. Simulations Reported in this Chapter. As explained earlier, we shall restrict ourselves to transport simulations for flows generated by time dependent velocity fields with finitely Fourier modes as given in formulae (4.17) and (4.18) above. Some illustrations obtained from expansions in a wavelet basis are reported in [29] but they are too limited to be used here.

None of the phenomena which we illustrate in this chapter depend upon the particular choice of the set \mathcal{K} of Fourier modes used in the definition of the velocity field as long as this set is not contained in a line in the plane (in which case we are dealing with a plain shear flow). Consequently, we shall refrain from using too large a set \mathcal{K} . Strictly speaking, our velocity fields will not satisfy the scaling and the isotropy properties which are fundamental in most theories. But because of their simplicity, the simulations can be set up very efficiently. By this we mean that the velocity field can be calculated very fast anywhere a passive tracer is found. In particular, since the results we want to illustrate show as well for these simple models, the simulations can be run on small workstations without having to use the parallel simulations described in [13].

A snapshot (a picture taken at a fixed value of the time t) of a sample realization of the stream function ϕ is given in Figure 2. Notice the smoothness of the surface and the similarities with some of the plots of Figure 13 of [34].

Figure 3 shows several streamlines (i.e. level curves of the stream function) together with small arrows representing the values of the velocity field at the same time and selected locations. These streamlines are closed curves going around the extrema of the stream function, the direction depending on whether the extremum is a maximum or a minimum. Because of the smoothness of the stream function already remarked and the Gaussian nature of the statistics, these extrema are essentially parabolic and the closed streamlines near an extremum are essentially circular. The analysis of the topological properties of the streamlines has been of crucial importance of the understanding of the transport properties of time independent velocity fields. See for example [34] for an extensive review of the results in the physics literature, [5] and [8] for examples of applications to specific models and [2] and [3] for examples of mathematical results on the boundedness of these lines. Indeed, in this case the passive tracers have to travel along the streamlines. So, once a realization of the stream function is obtained and once the streamlines are identified the transport properties are usually deduced easily. Unfortunately, this approach is extremely difficult in the time dependent case because of the time evolution of the streamlines. Letting the time evolve shows that some of these close

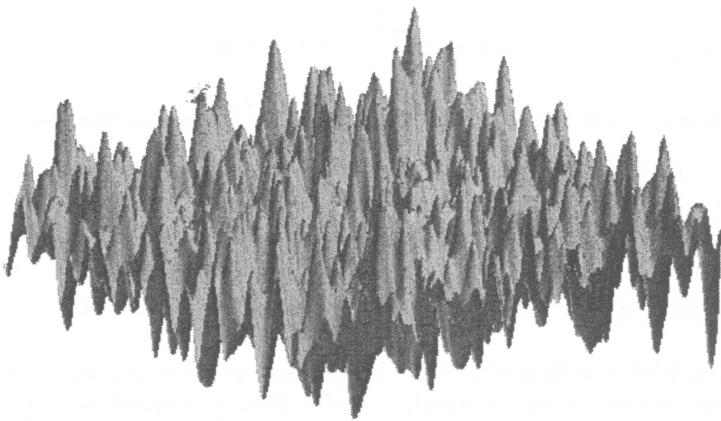


FIGURE 2. Snapshot of a realization of the stream function.

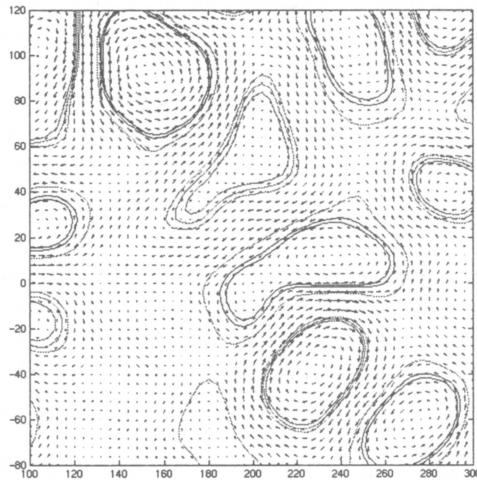


FIGURE 3. Several streamlines (i.e. level curves of the stream function) at a fixed time. The arrows represent the values of the velocity field at this time at locations selected for readability.

streamlines persist for quite some time, giving rise to vorticity eddies. We shall come back to that point when we discuss the Poisson models later in Section 7.

Even though finite Fourier mode velocity fields are not strictly speaking homogeneous and isotropic, Figure 4 shows that the sample realizations of these velocity fields almost possess these properties. Also, note the similarities between the images of the modulus and the angle of the realization of the velocity field contained in Figure 4 with Figure 10 in [34].

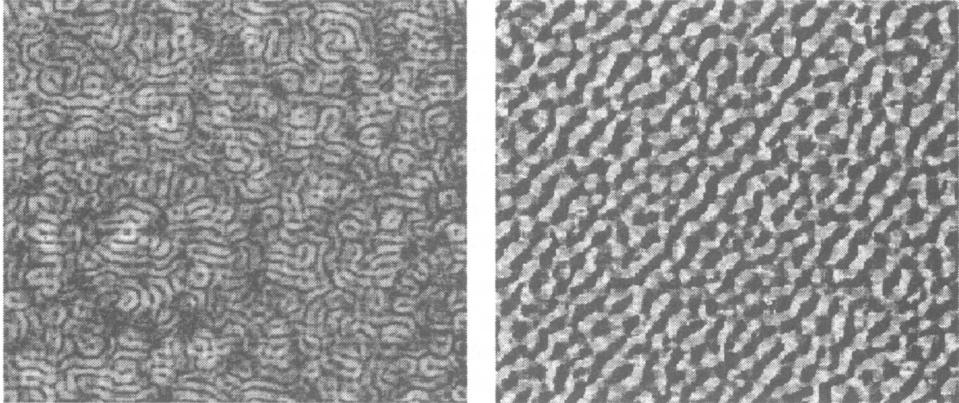


FIGURE 4. Left: raster image of the length of $\vec{v}(t, \cdot)$ taken from a snapshot of a realization of the velocity field. Right: raster image of the corresponding angle with a fixed direction. The color code used to visualize the various directions is irrelevant at this stage, we merely intend to illustrate the near homogeneity and isotropy of the velocity field.

The figures of this section have been taken from sequences of images produced to illustrate the time evolution of the realizations of the various scalar random fields (stream function, velocity modulus, velocity angle, \cdot). These images have been put together into animations which can be visualized with a web browser. They can be found at the URL [18]. We find it extremely instructive to observe these time evolutions and we shall make a case based on evidence they produce when we discuss the Poisson models later in Section 7. Unfortunately, a book chapter is not the appropriate medium for this type of illustration.

4.6. Mean Square Error and Choice of the Parameters. The estimation of the error is one of the many important theoretical problems which does not seem to have a satisfactory answer. If the goal is to understand the effects of the various approximations on the equation of motion and the transport properties of the flow, then the relevant quantity to control is:

$$(4.23) \quad \mathbb{E}\left\{\sup_{0 \leq t \leq T} \|\vec{v}(t, \mathbf{x}) - \vec{v}_K(t, \mathbf{x})\|^2\right\}$$

where we use the notation \vec{v}_K for the velocity field computed from the approximating stream function ϕ_K , whether the approximation is based on a wavelet basis expansion or a Fourier expansion as described above. Unfortunately, because of the presence of the supremum inside the expectation, this error is difficult to estimate. On the other hand, the mean square error:

$$(4.24) \quad \mathbb{E}\{\|\vec{v}(t, \mathbf{x}) - \vec{v}_K(t, \mathbf{x})\|^2\}$$

computed for a fixed (single) time t is much easier to estimate, especially when the approximate decomposition is computed on an orthonormal system. This was done in the case of a time independent shear flow for an approximate decomposition in

a wavelet basis both in [26] and [25]. A rough upper bound for (4.24) was derived in [41] for an approximate decomposition using a discrete set \mathcal{K} of Fourier modes. This bound was used to discuss the choice of the set \mathcal{K} . In particular it was argued that the set \mathcal{K} should be chosen as a finite subset of a small set of concentric circles with radii in the inertial range $[r_0, r_1]$. Moreover \mathcal{K} should be chosen appropriately if the isotropy condition is desired, the optimal choices of the number of circles, of the values of the radii and of the numbers of wave vectors \mathbf{k} per circle depending upon the values of the spectral parameters.

Unfortunately, the control of the mean square error (4.24) is of no use when we consider the large time properties of the flow. In particular, it is not sufficient to guarantee the convergence of the Lyapunov exponents of the (finite dimensional) approximate systems toward the Lyapunov exponent of the (infinite dimensional) full system. It is known that the approximation of Lyapunov exponents is already a touchy business in the case of ordinary stochastic differential equations (see for example [44, 32]) when the approximation is based merely on time discretization. The present situation is even more complex because of the approximation based on both time and space discretizations.

Besides, even though no one doubts the strict positivity of the upper Lyapunov exponent of the Jacobian flow in its full generality, no complete proof exists and the attempts based on the control of finite dimensional approximations have failed so far.

5. Transport Simulations

We are concerned with the simulation of the trajectories of passive individual tracers as well as collective properties of the time evolution of sets of tracers. As explained in the introduction, the trajectories of the tracers are the solutions of the equation of motion:

$$(5.1) \quad dX_t = \bar{\mathbf{v}}(t, X_t)dt + \sqrt{2\kappa}dW_t,$$

where $\{W_t; t \geq 0\}$ is a Wiener process which is statistically independent of the velocity field $\bar{\mathbf{v}}$ and the constant $\kappa \geq 0$ represents the molecular diffusivity. For the sake of simplicity we shall assume that $\kappa = 0$ when we discuss numerical simulations and we shall reinstall κ (and hence the viscosity term) when we review the theoretical results on homogenization and more generally on transport renormalization. For each realization of the velocity field, the equation of motion is an ordinary differential equation and we denote by $\varphi_{s,t}$ the corresponding flow. The existence of such a flow is quite a difficult result in the presence of the Wiener process. It is one of the fundamental results of the modern theory of stochastic differential equations. See for example [37]. In any case, $\varphi_{0,t}(\mathbf{x})$ is the solution X_t of (5.1) with initial condition $X_0 = \mathbf{x}$.

Once the choice of a sampling rate Δt has been made, the positions of the tracer particles are computed at the times $t_j = t_0 + j\Delta t$. In order to update the position of a particle one uses the simplest Euler scheme:

$$(5.2) \quad \mathbf{x}(t_j) = \mathbf{x}(t_{j-1}) + \bar{\mathbf{v}}(t_j, \mathbf{x}(t_{j-1}))\Delta t.$$

Figure 6 shows a typical trajectory of a passive tracer. Using the notations introduced below, we used a set \mathcal{K} of Fourier modes with 400 elements to simulate

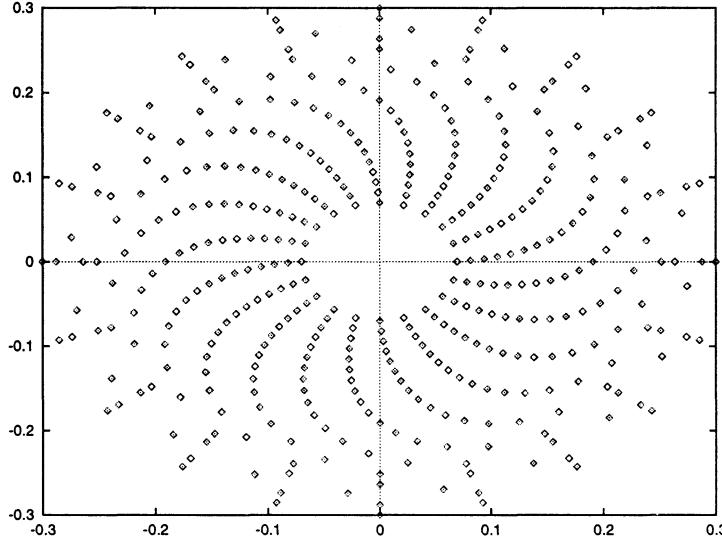


FIGURE 5. Fourier modes used for the simulations.

the values of the velocity field. These Fourier modes are plotted on figure 5. The components of these Fourier modes and the values of the corresponding α and σ parameters entering in the definitions of the scalar O-U processes are given by:

$$(5.3) \quad \alpha_{\mathbf{k}} = |\mathbf{k}|^z \quad \text{and} \quad \sigma_{\mathbf{k}} = |\mathbf{k}|^{\frac{2-\epsilon}{2}}$$

with $\epsilon = 0.3$ and $z = 1.2$.

The values of the α and σ parameters correspond to the values $\epsilon = 1.8$ and $z = 0.5$ of the spectral parameters of the model. The initial position \mathbf{x} of the tracer shown in Figure 6 is the origin, the time increment is $\Delta t = 0.01$ and we computed the position $\mathbf{x}(t_j)$ for $J = 120$ time samples t_j .

The simulation of the time evolution trajectories reduces to the problem of the generation of the velocity field, which in turn reduces to the generation of the stream function because of formula (4).

5.1. Liquid Curves and Interface Fractralization. The purpose of this section is to pursue the analysis initiated in [12] of the time evolution of curves in the plane. All the results discussed below apply to general curves but for the sake of illustration we shall restrict ourselves to simple closed curves which we view as the boundaries of simply connected domains in the plane. They are called *liquid curves* in [34].

Several authors have considered this problem in the case of Brownian flows where the analysis of the Lyapunov exponent is easier (see for example [38, 10, 39, 48].) Results in the case of Ornstein Uhlenbeck velocity fields are reported in [12] and [14]. Mathematically speaking we are interested in the time evolution of a smooth curve:

$$[0, 1] \ni \alpha \mapsto \gamma_0(\alpha) \in \mathbb{R}^d.$$

The latter can be viewed as the boundary at time $t = 0$ of a cloud of pollutant dust particles which are light enough to be passively transported by the velocity

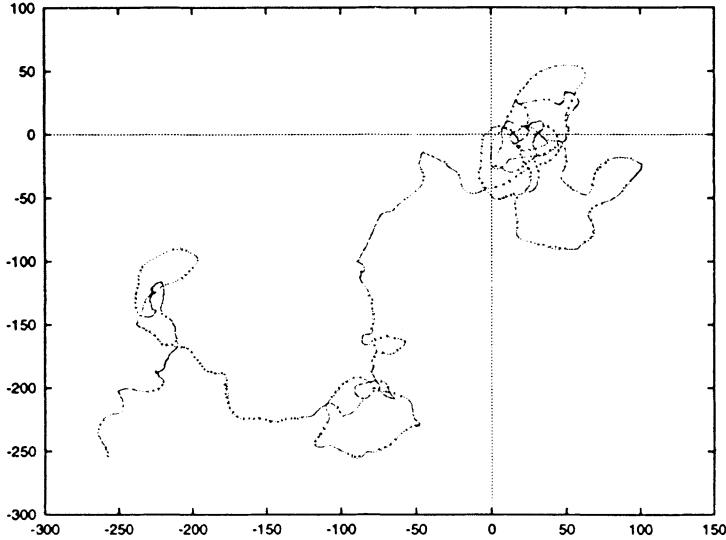


FIGURE 6. Example of a passive tracer trajectory.

field and we also assume that these dust particles do not interact. We analyze the time evolution of the shape of the cloud through the deformations over time of its boundary. We denote by γ_t the boundary of the cloud at time t .

$$[0, 1] \ni \alpha \mapsto \gamma_t(\alpha) = \varphi_{0,t}(\gamma_0(\alpha)) \in \mathbb{R}^d.$$

The results of the numerical simulations reported in [13] (and reproduced in [12]) confirmed what was known from the numerical simulations reported in [34] and from proof existing in the case of Brownian flows: the length of γ_t increases in all cases and the increase is more dramatic in the case of isotropic flows than in the case of shear flows. In fact it was shown in [12] that the length of γ_t is growing at most polynomially in the case of a shear flow and a simple argument was given to show that in general, the exponential rate of growth of the length of γ_t is bounded from below by the Lyapunov exponent of the Jacobian flow. Also, numerical evidence was given for the strict positivity of this Lyapunov exponent in the case of flows with finitely many Fourier modes. This fact was subsequently proven in [14]: the positivity of the Lyapunov exponent holds true as long as the Fourier modes are not concentrated on a line (i.e. as long as we are not dealing with a shear flow.)

Since the flow $\varphi_{s,t}$ is a flow of diffeomorphisms, the curve γ_t remains smooth at all times, but as its length grows, its shape becomes very complex and intricate, filling up more and more space in the plane. After all, the length of γ_t grows exponentially while its diameter grows only polynomially (presumably linearly as explained below in Subsection 5.2.)

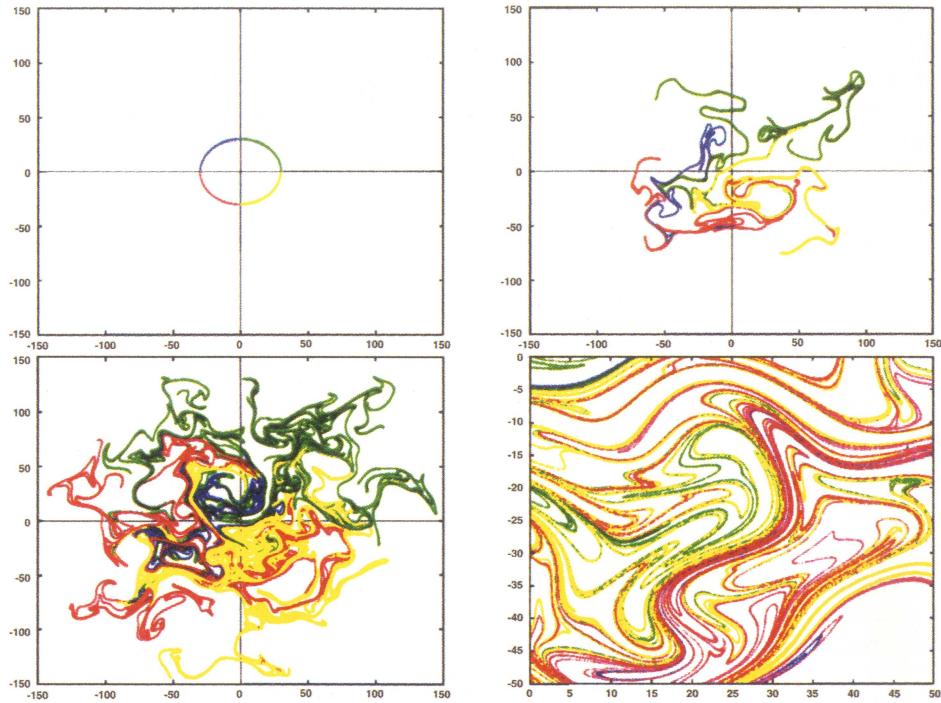


FIGURE 7. Example of a time evolution of the unit circle.

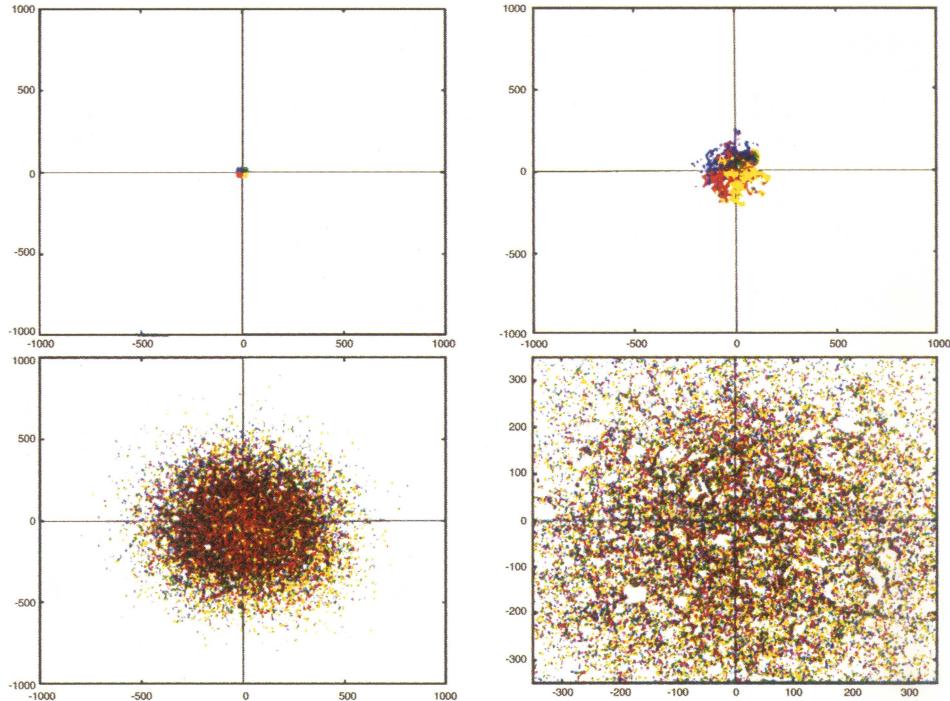


FIGURE 8. Illustration of the mixing properties of the OU velocity fields.

Figure 7 shows a typical example of such a time evolution. The parameters of the velocity field were the same as before. From left to right and top to bottom the plots give the sets γ_t for $t = 0, 19.2$, and $t = 36$ for which we plotted the whole curve and a zoom on the central part. Notice that the circle is sampled with $N = 500$ points at time $t = 0$ and that, in order to keep the density constant along γ_t we had to sample γ_t with $N = 500$, $N = 17084$ and $N = 954582$ at the instants $t = 0$ $t = 16.8$ and $t = 36$. Striking animations showing the time evolution of the curve γ_t are given on the Web at the URL [18]. The exponential increase of the length of γ_t imposes on a corresponding exponential increase of the number of points to plot at each given time and this is the main limitation in the production of the animation files: memory capacity is reached very fast.

In any case, these numerical results suggest that the dimension of the subset of the plane formed by the points $\gamma_t(s)$ may, at least in an appropriate asymptotic regime, become strictly larger than 1.

Since there are many notions of fractal dimension, for the sake of simplicity, we concentrate on two specific choices. They are related to the Minkowski dimension and the so-called correlation dimension of γ_t .

We first discuss the Minkowski dimension. For each fixed time t and each $\epsilon > 0$ we compute the area $|\gamma_t(\epsilon)|$ of the sausage of width ϵ around γ_t i.e. the area of the ϵ neighborhood:

$$\gamma_t(\epsilon) = \{\mathbf{x}; \exists \alpha \in [0, 1], |\gamma_t(\alpha) - \mathbf{x}| \leq \epsilon\}$$

of the curve γ_t and we look for a small ϵ asymptotic behavior of the form $|\gamma_t(\epsilon)| \approx C\epsilon^d$ so we could claim that the dimension of the curve is d . The existence of such an asymptotic behavior should be evidenced by a plot of the logarithm of $|\gamma_t(\epsilon)|$ versus the logarithm of ϵ . Such a scatterplot should show that $|\gamma_t(\epsilon)|$ can be explained by a linear regression on ϵ and the Minkowski dimension should be given by the estimate of the slope. See Figure 9. Because of the very intricate shape of the curve γ_t , the computation of the area $|\gamma_t(\epsilon)|$ could be problematic. We used a simple Monte Carlo simulation. We generated a large number of points in a rectangle containing the curve γ_t with a uniform distribution and the area $|\gamma_t(\epsilon)|$ was approximated by the proportion of points within a distance ϵ of any of the points forming the curve γ_t .

We now explain the computation of the correlation dimension of γ_t , as proposed by Grassberger and Procaccia [30, 31] to characterize strange attractors. For each fixed time t and each $\epsilon > 0$ we consider

$$C(\epsilon) = \frac{1}{N^2} \sum_{i,j=1}^N 1_{\{\|X_i - X_j\| < \epsilon\}}$$

where N is the number of points of the discretized curve $\{X_i, 1 \leq i \leq N\}$. $C(\epsilon)$ is expected to behave like ϵ^ν for ϵ small, and ν is called the correlation dimension. As before we try to measure the small ϵ behavior of the logarithm of this number $C(\epsilon)$ as a function of the logarithm of ϵ and we derive an estimate of the dimension of γ_t as before. Experiments carried out in [30] on some strange attractors show that ν is a good estimate for the Hausdorff dimension.

Figure 9 shows numerical results attempting to determine the Minkowski and the correlation dimensions of γ_t in the particular case of the unit circle γ_0 evolved under a velocity field with the same characteristics as above (section 5.3).

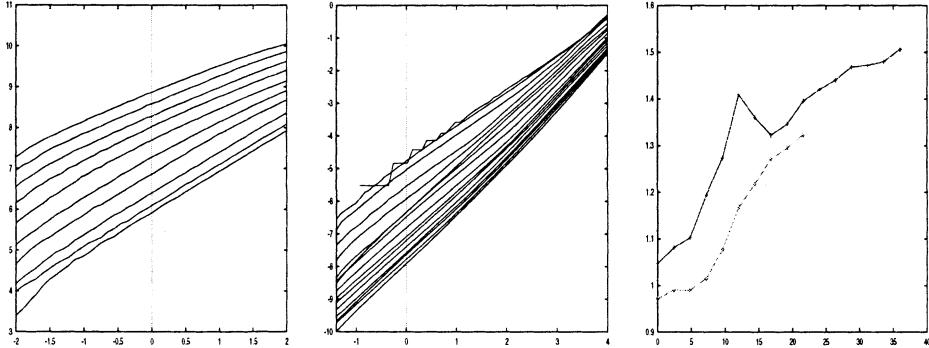


FIGURE 9. Numerical evidence of the existence of an effective (fractal) dimension. The Minkowski and the correlation exponents are treated in the left and the middle plots respectively. The estimations of the dimension by estimating the slopes are shown on the right plot. The upper curve is related to the correlation dimension.

As explained earlier, the fractalization phenomenon is only a large time asymptotic effect. For each fixed finite $t > 0$, the curve γ_t is C^∞ and consequently, the small ϵ behavior should be typical of a one-dimensional set, i.e. we should find a slope equal to 1 when we plot the logarithm of $N_t(\epsilon)$ versus the logarithm of ϵ . In fact, when ϵ is too small for t , the zooming effect of $N_t(\epsilon)$ should show that γ_t is a C^∞ curve and the effective dimension should be 1. On the other hand, there should be an intermediate regime where ϵ is not small enough and t is large enough to fool $N_t(\epsilon)$ into believing that it is dealing with a curve of dimension greater than 1. The numerical simulations reported in Figure 9 were intended to exhibit by varying both ϵ and t , this intermediate regime.

REMARK 5.1. 1. Related questions have been considered in [7] where a slightly different problem involving the asymptotic dimension of interfaces was considered in a renormalization regime for time independent shear flows.

2. Another interesting challenge is the numerical computation of the singularity spectrum of γ_t in the multifractal formalism of [11] and the analysis of its time evolution. We believe that an intermediate regime with fractal characteristics should be evidenced as we did above.

We also discuss a few conjectures for which we could not find numerical evidence (one way or another).

5.2. Diameter Asymptotics. At time $t = 0$ we consider a closed set D_0 and we consider its evolution under the flow. In other words we consider the time evolution of the random set $D_t = \varphi_{0,t}(D_0)$. This subsection is concerned with the analysis of the diameter d_t of this set. Notice that this diameter is equal to the diameter of the boundary ∂D_t which in turn is the result of the transport by the flow of the boundary of the initial set D_0 since $\partial D_t = \varphi_{0,t}(\partial D_0)$. In other words, we could start from ∂D_0 instead of D_0 without changing anything!

The goal is to find a (deterministic) function $t \mapsto r_t$ for which:

$$\lim_{t \rightarrow \infty} \frac{d_t}{a_t} = 1$$

almost surely, and possibly "tests" to determine upper and lower classes for this limit theorem. This problem is more difficult than it looks at first and no answer is known, even in the case of the Brownian flows. It is easy to show that:

$$\lim_{t \rightarrow \infty} \frac{d_t}{\sqrt{t}} = \infty$$

when the cardinality of D_0 is infinite, but the exact rate a_t is elusive. The simplicity of the formulation of the problem is deceptive in the sense that it is not indicative of its difficulty. Y. Sinai (private conversation) suggested that a_t could be a linear function of time. This conjecture is more or less explicit in [34]. What is even more frustrating is that we were not able to set up simulations to validate (or invalidate) this conjecture. Indeed, any finite number of points of D_0 will escape at the diffusive rate of \sqrt{t} and the rate a_t can only be achieved by means of elongated fingers sticking out of ∂D_t and surviving for a long time before being absorbed by the bulk of D_t , at which time new long finger will emerge. The sampling of the plane (and/or ∂D_t) would have to be refined in an adaptive way to illustrate such a phenomenon and we were not able to set up such simulations. This problem is one of the very many challenging questions which are left open.

5.3. Mixing. The asymptotic transport properties discussed in this chapter are strongly dependent upon the mixing properties of the flow. The latter have been studied in the case of Brownian flows in [38], [10] and [39] and it was shown there that the positivity of the upper Lyapunov exponent of the Jacobian flow was responsible for some of this mixing. Figure 8 illustrate the strong mixing properties of the Ornstein Uhlenbeck flows, even with a small number of Fourier modes as long as these modes are not aligned in the Fourier plane (recall that it was proven in [14] that this condition guarantees the strict positivity of the Lyapunov exponent). The plots of The color plate Figure 8 was produced from flows generated by velocity fields of the type (4.17) and (4.18) with again the same parameters. Initially, we distribute the drifters (about 50,000) uniformly in a square. We gave the same color to all the points contained in the same quadrant, and we plot the positions of these points over time (more precisely to times $t = 0$, $t = 48$ and $t = 360$) with the same colors. The four colors being applied in a specific order by the printer (green, blue, red and finally yellow), the misleading impression that the numbers of points of the different colors are not the same, does result from this printing artifact. The plot on the bottom right was obtained by zooming in the central region of the plot on the bottom left.

5.4. Upper Lyapunov Exponent of the Jacobian Flow. As a final remark on the properties of the upper Lyapunov exponent of the Jacobian flow we mention the numerical analysis of the dependence of this exponent upon the spectral parameters ϵ and z which was performed in [12] using parallel computations on the MASPAR II. These numerical experiments suggest that this exponent is a smooth convex function of ϵ and z but no rigorous result in this direction is known.

6. Homogenization & Spectral Singularity Renormalization

Physical systems with oscillating and/or random coefficients are difficult to analyze and many ad-hoc techniques have been proposed to average out the small scale variations of the coefficients and isolate the features of the solutions which are characteristics of scales larger than those of the typical variations of the coefficients.

Among all the mathematical procedures designed to derive an effective equation governing the large time / large scale behavior of the solutions, homogenization is certainly the most popular. In a nutshell, this procedure consists in studying the limit when the *small parameter* δ tends to 0 of the solutions after the rescalings

$$(6.1) \quad t \rightarrow t/\delta^2 \quad \text{and} \quad \mathbf{x} \rightarrow \mathbf{x}/\delta$$

of the time and the space variables. It can be applied to *ordinary* (stochastic) differential equations such as the equation of motion as well as to (stochastic) partial differential equations such as the stochastic parabolic equation of the concentrations of passive tracers.

Because this problem has been thoroughly investigated, the literature on the subject is enormous. Most of the published works concern the time independent case. We shall only quote the lecture notes [43] for the probabilistic approach and [7] for a more analytic perspective. Both are well suited for the type of equations considered here and we suggest to the interested reader to consult the many references therein.

For the time dependent case we refer to [16] and to the references therein. The proof of [16] is different from most of the proofs in the time independent case in that it takes full advantage of the Markov property of the Eulerian velocity field. The main thrust of this proof is to identify the distribution of the Lagrangian motions and to analyze its ergodic properties. This proof was originally restricted to the finite Fourier mode case. It has been extended to the general case in [17]. One of its interesting characteristics is that it does not rely on the assumption $\kappa > 0$ and according to [34], no rigorous proof was known that this assumption could be dispensed with.

In its strongest probabilistic formulation homogenization can be viewed as a functional central limit theorem for the rescaled solutions of the equation of motion and the result states that in the large time / large scale regime given by the diffusive scaling (6.1) the Lagrangian motions are essentially Brownian motions with a variance, the so-called effective diffusivity, which is a complicated function of the original diffusivity constant κ and the spectral parameters of the model. As explained above this result was proved in [16] for the finite Fourier mode models and we believed at that time that the result should hold in full generality as long as there is no infrared spectral singularity. In fact we proved in [17] that it is valid for general O-U models as long as:

$$(6.2) \quad \epsilon + z < 2.$$

This result was proved by Avellaneda and Majda in [6] for the shear flow models. In their path-breaking analysis of the shear flow models these authors showed that depending on the nature of the spectral singularity (more precisely depending on which region of the (ϵ, z) -plane of spectral parameters the couple (ϵ, z) belongs to) it is possible to find a large time / large scale rescaling

$$(6.3) \quad t \rightarrow t/\rho(\delta)^2 \quad \text{and} \quad \mathbf{x} \rightarrow \mathbf{x}/\delta$$

for which the solutions of the rescaled parabolic equation for the concentrations converges to the solution of a limiting equation. According to these results, in each region of the spectral parameter plane, the nature of the time scaling $\rho(\delta)$ and the nature of the limiting equation give a sub-diffusive, diffusive or super-diffusive large time / large scale motion. This partition of the (ϵ, z) plane identified in [6] for shear

flow models was claimed to be universal. Attempts have been made in [17] to show what remains of this partition in the general case but the results remain speculative and partial at best.

A different set of renormalization procedures was proposed in [28] for the same shear flow models studied in [6]. A heated debate followed the publication of [23] and the controversy remains mostly unresolved.

The results of the renormalization procedures are striking mathematical properties of the asymptotics of the models. Our experience is that they are extremely difficult if not impossible to illustrate numerically.

As we already pointed out, the computation of the structure function requires the evaluation of the velocity field at one (single) point in space which does not change with time or the realization of the velocity field. Moreover this computation does not see much of the spectral singularity. It is conceivable that computations of this type are feasible, especially in the time-independent case. Computations of the structure function are described in [26] where the authors use wavelet - like expansions to confirm numerically some of the theoretical findings of the renormalization theory of time independent shear flows. We were able to reproduce these results with Fourier - like expansions but because the use of wavelets is well adapted to the approach to the singularity, at equal approximation error, the computational burden of the Fourier - like expansions is bigger. A similar wavelet - like expansion was used in [25] to compute numerically the mean square displacement $\mathbb{E}\{\|\varphi_{0,t}(\mathbf{x}_1) - \varphi_{0,t}(\mathbf{x}_2)\|^2\}$ as a function of the time t and the initial separation $\|\mathbf{x}_1 - \mathbf{x}_2\|$ of the starting points of the Lagrangian trajectories. These computations can be done explicitly in the case of time independent shear flow. The authors use the numerical computation of this mean square displacement by Monte Carlo simulation of the velocity field and solution of the equation of motion as a test for the method chosen to simulate the velocity field. These simulations are of the transport type but, because the authors of [25] are using a time independent shear flow model, given the knowledge of t and $\|\mathbf{x}_1 - \mathbf{x}_2\|$, the points at which the velocity will have to be computed form a deterministic (linear) set which can be computed up-front. With this knowledge it become easy to determine which wavelet coefficients will be needed and which Gaussian random variables need to be generated in order to set up the simulation. This is not possible in the general case of a *non shear flow* model for which one cannot predict with certainty where the velocity field will have to be computed. As we explained earlier in Subsection 4.3 the number of Gaussian random variables to generate and the number of wavelet coefficients to compute are so big that the computations become prohibitive, especially in the time-dependent case.

Notice also that simulations attempting to illustrate the homogenization regime are reported in [34], but no means is provided to assess the nature of the approximations and to quantify the corresponding errors.

We believe that the numerical analysis of the renormalization limits of the transport properties of singular velocity fields of a general nature is beyond the scope of the current computer and mathematical technology.

7. Poisson Models

Gaussian fields are not the only models of stationary and homogeneous random velocity fields which have been studied for their transport properties. A popular

model consists in the linear superposition of elementary localized velocity fields scattered in space as a Poisson point process. These elementary velocity fields are chosen to mimic the behavior of vortices. These vortices do not interact and are not transported by the flow, so these models are different from the computational algorithms proposed by Chorin to compute numerically solutions of the 2-D Navier Stokes equation [20]. A time independent model of this type has been considered in [8] but we shall concentrate our discussion on the time dependent case of shot noise velocity fields introduced by Cinlar in [21] and [22]. The "*O-U nature*" of these models makes us believe that the results presented here for O-U velocity fields with Gaussian statistics should still hold for these shot noise Poisson velocity fields. We even believe that the strong mixing properties (consequences of the uniform spectral gap and in particular the lack of infrared spectral singularity) of the velocity field introduced in [21] and [22] should make some of the proofs (positivity of the upper Lyapunov of the Jacobian flow, homogenization, ...) easier. But these claims are mere speculations since we do not know of any proof of these results.

There is no doubt in our mind that the shot noise Poisson models form a class of interesting mathematical models which deserve to be studied for their own sake. But we would like to re-battle one of the common arguments used to justify their introduction: "*they possess vorticity features which cannot be produced by the Gaussian models*". In fact, the theory of extremes of random fields tells us that the very high peaks of a homogeneous Gaussian stream function form (at least asymptotically) a Poisson process in the plane. See for example [1]. Moreover, as we saw from Figure 3 each peak produces a vortex and if a peak remains a peak for a significant amount of time, this vortex will create a spinning motion of the passive tracers up until the peak of the stream function is absorbed. This phenomenon is quite obvious from the animations available on the Web page [18]. We are convinced that these long lived vortices do exists in Gaussian models and one may legitimately ask :

why create artificially and by brute force features which are already present in the Gaussian models?

Acknowledgments: The first named author (RC) would like to thank Jean Pierre Fouque for pointing out the review article [34] of Isichenko. The second named author (FC) would like to thank the Department of Civil Engineering & Operations Research of Princeton University for its warm hospitality and the Office of Naval Research for partially supporting the visits during which most of this work was done.

Bibliography

- [1] R.J. Adler (1980): Geometry of Random Fields. Wiley, New York, N.Y.
- [2] K.S. Alexander (1996): Boundedness of level lines for two-dimensional random fields. *Ann. Probab.* **24**, 1653 - 1674.
- [3] K.S. Alexander and S.A. Molchanov (1994): Percolation of level sets for two dimensional random fields with lattice symmetry. *J. Stat. Phys.* **77**, 627 - 643.
- [4] A. Antoniadis and R. Carmona (1985): Infinite Dimensional Ornstein Ulhenbeck Processes *Probab. Th. Rel. Fields* **74**, 31-54.
- [5] M. Avellaneda, F. Elliott and A. Appelian (1993): Trapping, Percolation and Anomalous Diffusion of Particles in a Two-Dimensional Random Field. *Journal of Statistical Physics* **72**, 1227-1304.
- [6] M. Avellaneda and A. Majda (1990): Mathematical models with exact renormalization for turbulent transport. *Commun. Math. Phys.*, **131**, 381-429.

- [7] M. Avellaneda and A. Majda (1990): Mathematical Models with Exact Renormalization for Turbulent Transport. *Commun. Math. Phys.*, **131**, 381-429.
- [8] M. Avellaneda and S. Torquato (1991): Diffusion and Geometric Effects in Passive Advection by Random Arrays of Vortices. *Phys. Fluids A*, **3**, 1880-1891.
- [9] G.K. Batchelor (1982): The theory of homogeneous turbulence. *Cambridge University Press*, Cambridge.
- [10] P.H. Baxendale (1986): Asymptotic Behavior of Stochastic Flows of Diffeomorphisms: Two Case Studies. *Proba. Th. Rel. Fields* **73**, 51-85.
- [11] J.F. Buzy, E. Bacry and A. Arneodo (1992): Multifractal formalism for fractal signals: the structure function approach versus the wavelet transform modulus maxima method. (preprint)
- [12] R. Carmona (1996): Transport Properties of Gaussian Velocity Fields. in *Real and Stochastic Analysis: Recent Advances* ed. M.M. Rao, CRC Press
- [13] R. Carmona, S. Grishin and S.A. Molchanov (1994): Massively Parallel Simulations of the Transport Properties of Gaussian Velocity Fields. to appear in *Mathematical Models for Oceanography*, eds R. Adler, P. Muller and B. Rozovskii. Birkhäuser.
- [14] R. Carmona, S. Grishin, S.A. Molchanov and L. Xu (1995): Exponential Stretching for Ornstein Ulhenbeck Flows with Finitely Many Frequency Modes. *Electr. Comm. Probab.* (to appear).
- [15] R. Carmona and A. Wang (1994): Comparison Tests for the Spectra of Dependent Multivariate Time Series. in *Mathematical Models for Oceanography*, eds R. Adler, P. Muller and B. Rozovskii. Birkhäuser.
- [16] R. Carmona and L. Xu (1997): Homogenization for Time Dependent 2-D Incompressible Gaussian Flows. *Ann. Appl. Probab.* (to appear)
- [17] R. Carmona and L. Xu (1997): Renormalization Results for Time Dependent Incompressible OU Velocity Fields. (in preparation)
- [18] URL <http://chelsea.princeton.edu/~rcarmona/simul/transport.html>
- [19] URL <http://soil.princeton.edu/~rcarmona/researchinterests.html>
- [20] A.J. Chorin (1994): Vorticity and Turbulence. Springer Verlag, New York, N.Y.
- [21] E. Cinlar (1994): On a Stochastic Velocity Field. (preprint)
- [22] E. Cinlar (1994): Poisson Shot Noise Velocity Fields. (preprint)
- [23] E. Cinlar and C.L. Zirbel (1996): Dispersion of Particle Systems in Brownian Flows. *Adv. Appl. Probab.* **28**, 53-74.
- [24] I. Daubechies (1992): Ten Lectures on Wavelets. SIAM
- [25] F. Elliott, D. Horntrop and , A. Majda (1996): A Fourier - Wavelet Monte Carlo Method for Fractal Random Fields. (preprint)
- [26] F. Elliott and A. Majda (1994): A Wavelet Monte Carlo Method for Turbulent Diffusion with Many Spatial Scales. *J. Comput. Phys.* **113**, 82-111.
- [27] J. Gaines (1995): Numerical Experiments with S(P)DE's. in *Stochastic Partial Differential Equations* ed. A. Etheridge. London Math. Soc. Lect. Note Series **216**, Cambridge Univ. Press.
- [28] J. Glimm and Q. Zhang (1992): Inertial Range Scaling of Laminar Shear Flow as a Model of Turbulent Transport. *Commun. Math. Phys.*, **146**, 217-229.
- [29] S. Grishin (1997): PhD CEOR, Princeton University
- [30] P. Grassberger and I. Procaccia (1983): Characterization of Strange Attractors. *Phys. Rev. Lett.* **50**(5), 346-349.
- [31] P. Grassberger and I. Procaccia (1983): Estimation of the Kolmogorov Entropy from a Chaotic Signal. *Phys. Rev. A* **28**(4), 2591-2593.
- [32] A. Grorud and D. Talay (1995): Approximation of Upper Lyapunov Exponents of Bilinear Stochastic Differential Systems. *SIAM J. Numer. Anal.* **28**(4), 1141-1164.
- [33] D.J. Horntrop and A.J. Majda (1994): Subtle Statistical Behavior in Simple Models for Random Advection-Diffusion. *J. Math. Sci. Univ. Tokyo* **1**, 23-70.
- [34] M.B. Isichenko (1992): Percolation, statistical topography and transport in random media. *Rev. Modern Phys.* **64**, 961-1043.
- [35] R.M. Kraichnan (1959): The structure of isotropic turbulence at very high Reynolds numbers. *J. Fluids Mech.* **5**, 497-543.
- [36] R.M. Kraichnan (1974): Convection of a passive scalar by a quasi uniform random straining field. *J. Fluids Mech.* **64**, 737-762.
- [37] H. Kunita (1990): Stochastic Flows and Differential Equations. Cambridge University Press.

- [38] Y. Le Jan (1984): On isotropic Brownian motions. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **70**, 609-620.
- [39] Y. Le Jan (1991): Asymptotic Properties of Isotropic Brownian Flows. in *Spatial Stochastic Processes*, eds K.S. Alexander, J.C. Watkins, pp 219-232, Birkhäuser, Boston.
- [40] A.J. Majda (1994): Random Shearing Direction Model for Isotropic Turbulent Diffusion. *J. Statist. Phys.* **75**, 1153-1165.
- [41] S.A. Molchanov (1994): Topics in Statistical Oceanography. in *Mathematical Models for Oceanography*, pp. 343-380, eds R. Adler, P. Muller and B. Rozovskii. Birkhäuser.
- [42] A.S. Monin and A.M. Yaglom (1971): Statistical Fluid Mechanics: Mechanics of Turbulence. MIT Press, Cambridge, Mass.
- [43] S. Olla (1994): Homogenization of Diffusion Processes in Random Fields. (preprint, Ecole Polytechnique Paris.)
- [44] D. Talay (1991): Lyapunov Exponents of Nonlinear Stochastic Differential Equations. *SIAM J. Appl. Mat.* (to appear).
- [45] G.I. Taylor (1953): Dispersion of Soluble Matter in Solvent flowing Slowly through a Tube. *Proc. Royal Soc. A* **219**, 186-203.
- [46] J.B. Walsh (1986): An Introduction to Stochastic Partial Differential Equations in *Ecole d'Eté de Probabilités de Saint Flour 1984* Lect. Note in Math. **1180**, 265-439.
- [47] A.M. Yaglom (1987): Correlation Theory of Stationary and Related Random Functions. vol. I: Basic Results. Springer Verlag, New York, N.Y.
- [48] C.L. Zirbel (1993): Stochastic Flows: Dispersion of a Mass Distribution and Lagrangian Observations of a Random Field. Ph. D. Princeton.

This page intentionally left blank

Part 2

MATHEMATICAL THEORY

This page intentionally left blank

CHAPTER 5

An analytic approach to SPDEs

N.V. KRYLOV
 127 Vincent Hall
 University of Minnesota
 Minneapolis, MN, 55455
 krylov@math.umn.edu

1. Introduction

Evolutional stochastic partial differential equations (SPDEs) arise in many applications of probability theory and have been treated since long ago (see [30]). An example of a linear second-order SPDE is given by the following equation in \mathbb{R}^d :

$$(1.1) \quad du = (a^{ij}u_{x^i x^j} + b^i u_{x^i} + cu + f) dt + (\sigma^{ik}u_{x^i} + \nu^k u + g^k) dw_t^k, \quad t > 0,$$

where summation with respect to the repeated indices i, j, k is assumed as usual, i and j go from 1 to d , and k may run through $1, 2, \dots$.

The main purpose of this publication is to present a theory of solvability of the Cauchy problem for linear and some quasi-linear equations like (1.1) in spaces of summable functions with exponent of summability $p \geq 2$. If $p = 2$, so that we are concerned with solutions belonging to the Sobolev spaces $W_2^n(\mathbb{R}^d)$, such a theory does exist and is rather complete and satisfactory (see, for instance, [30]). Recall that $W_2^n(\mathbb{R}^d)$ is the set of all generalized functions on \mathbb{R}^d whose derivatives up to and including the n th order belong to $L_2(\mathbb{R}^d)$. Some results concerning the solvability of the first boundary-value problem in spaces like $W_2^n(D)$, where D is a smooth domain, can be found in [2], [6], [20], and [32]. Roughly speaking, the main tool in W_2^n -theory is integration by parts. There are also approaches based on semigroup methods [5], [6], which work well for the equations with nonrandom leading coefficients a^{ij} and again in the Hilbert-space framework.

The necessity of the L_p -theory arises, for instance, when one wants to find the solutions numerically. The convergence rate and the way the finite difference should be chosen depend on smoothness properties of solutions. Here by the smoothness properties we mean continuity, Hölder continuity, differentiability, continuity of derivatives, and so on.

One of inconveniences of W_2^n -theory is that $W_2^n(\mathbb{R}^d) \subset C^{n-d/2}(\mathbb{R}^d)$ only if $2n > d$, and one can prove that the solutions belong to $W_2^n(\mathbb{R}^d)$ only if the coefficients are $n - 2$ times continuously differentiable. Therefore, if we want to get the solutions m times continuously differentiable with respect to $x \in \mathbb{R}^d$, we have to suppose that the coefficients of the equation are more than $m + d/2 - 2$ times continuously differentiable even if the free terms are of class $C_0^\infty(\mathbb{R}^d)$. At the same

time, $W_p^n(\mathbb{R}^d) \subset C^{n-d/p}(\mathbb{R}^d)$ if $pn > d$, and, by taking p sufficiently large, we see that the solutions have almost as many usual derivatives as generalized ones. Actually, exactly for this purpose the spaces $W_p^n(\mathbb{R}^d)$ with $p \geq 2$ have already been used in SPDE theory (see, for instance, [30]), but the corresponding results, obtained again by integration by parts, were not sharp. It is worth mentioning that sharp results concerning $C^{2+\alpha}(\mathbb{R}^d)$ -theory are recently obtained in [27] for equations like (1.1) but with $\sigma \equiv 0$.

Another advantage of the W_p^n setting with $p \geq 2$ can be seen in the case of very popular equations with so-called cylindrical white noise (see, for instance, [8], [26], [28], [35], and references therein). Although these equations are covered by the general W_p^n -theory for any $p \geq 2$ (see Section 8.3), for $p = 2$ we get only the solutions summable to any degreee, and the solutions become continuous only for $p > 2$. By the way, as in [30], we consider n positive and negative, but in contrast with [30] we allow n to have non-integer values. For general n we are working in the spaces of Bessel potentials $H_p^n(\mathbb{R}^d)$, and, in the case of equations with cylindrical white noise, we take n slightly less than $(-3/2)$.

Our main tool is the theory of spaces $H_p^n(\mathbb{R}^d)$, borrowed from [33], together with a result from [17] or [18] which is an analog of the so-called maximal regularity property of stochastic convolutions in Hilbert spaces obtained by Da Prato (cf. [4]). We also use some results from the theory of parabolic equations and follow a general scheme of proving the solvability of PDEs adopted in this theory. We discuss this general scheme in Sec. 2.

The main source of these notes is author's article [20]. However, it underwent a major revision and restructuring here, so that, in author's opinion, the whole subject looks more natural now. Also, the proofs are given with much more details, and several mistakes have been corrected. In this regard the contribution of S. Lototsky, H. Yoo, and A. Zatezalo is greatly appreciated. The author has delivered four lectures on the subject at Workshop/School on SPDEs, Theory and Applications, Los Angeles, January, 3–7, 1996. This was a very good opportunity to think all over again. It is a pleasure to thank the organizers of the Workshop for the invitation and hospitality.

2. Generalities

We want to explain here some basic ideas guiding the investigation we are going to present further.

Rewrite equation (1.1) in the following form

$$(2.1) \quad du = (Lu + f) dt + (\Lambda^k u + g^k) dw_t^k, \quad t > 0,$$

where

$$Lu = a^{ij} u_{x^i x^j} + b^i u_{x^i} + cu, \quad \Lambda^k u = \sigma^{ik} u_{x^i} + \nu^k u.$$

Recall that we are using the summation convention. Of course, w_t^k are independent Wiener processes and du is Itô's stochastic differential with respect to t . The equation is considered for all $x \in \mathbb{R}^d$ and $t > 0$. The coefficients σ, ν , and the “free” term g may vanish, and then (2.1) becomes a usual parabolic partial differential equation. Therefore, it is natural to recall some ideas and results from the L_p -theory of parabolic partial differential equations.

Let \mathbb{R}^d be a d -dimensional Euclidean space of points $x = (x^1, \dots, x^d)$. By a distribution or a generalized function on \mathbb{R}^d we mean an element of the space \mathcal{D} of

real-valued Schwarz distributions defined on C_0^∞ , where $C_0^\infty = C_0^\infty(\mathbb{R}^d)$ is the set of all infinitely differentiable functions with compact support.

For given $p \in (1, \infty)$ and $n \in (-\infty, \infty)$, define the space $H_p^n = H_p^n(\mathbb{R}^d)$ (called the space of Bessel potentials or the Sobolev space with fractional derivatives) as the space of all generalized functions u such that $(1 - \Delta)^{n/2}u \in L_p = L_p(\mathbb{R}^d)$. To explain the meaning of this, let us first introduce $(1 - \Delta)^{n/2}$ in the following way (see [14]). If $\alpha \in (0, 1)$, then, for a constant $c(\alpha)$ and all $z \leq 0$,

$$(1 - z)^\alpha = c(\alpha) \int_0^\infty \frac{e^{-t} e^{zt} - 1}{t^\alpha} \frac{dt}{t}.$$

If we formally substitute here Δ instead of z , then we get the following definition of $(1 - \Delta)^\alpha$

$$(2.2) \quad (1 - \Delta)^\alpha u = c(\alpha) \int_0^\infty \frac{e^{-t} T_t u - u}{t^\alpha} \frac{dt}{t},$$

where T_t is the semigroup associated with Δ , that is, a solution of $T'_t = \Delta T_t$ similarly to e^{zt} which is a solution of $f' = zf$. The following is an explicit formula for T_t :

$$(2.3) \quad T_t u(x) := \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d} u(y) e^{-\frac{1}{4t}|x-y|^2} dy.$$

In the same way for any $\alpha > 0$ we define

$$(2.4) \quad (1 - \Delta)^{-\alpha} u = d(\alpha) \int_0^\infty t^\alpha e^{-t} T_t u \frac{dt}{t},$$

where $d(\alpha)$ is an appropriate constant. It turns out (see [14]) that formulas (2.2) and (2.4) are sufficient to consistently define $(1 - \Delta)^{n/2}$ for any $n \in (-\infty, \infty)$. The result of application of $(1 - \Delta)^{-n/2}$ to an $f \in L_p$ is defined as a limit of truncated integrals in (2.2) or (2.4). For a distribution u , we say that $u \in H_p^n$ if there is $f \in L_p$ such that $u = (1 - \Delta)^{-n/2}f$ in the sense of distributions, that is, if u and $(1 - \Delta)^{-n/2}f$ coincide as elements in the space \mathcal{D} . In this case we also write $(1 - \Delta)^{n/2}u = f$.

For $u \in H_p^n$ one introduces the norm

$$\|u\|_{n,p} := \|(1 - \Delta)^{n/2}u\|_p,$$

where $\|\cdot\|_p$ is the norm in L_p . It is known (see, for instance, [33], [34]) that H_p^n is a Banach space with norm $\|\cdot\|_{n,p}$ and the set C_0^∞ is dense in H_p^n .

Next, for fixed T one introduces the space $H_p^{1,2}(T) = H_p^{1,2}((0, T) \times \mathbb{R}^d)$ as

$$\{u = u(t, x) : \|u\|_{1,2,p}^p := \int_0^T \left\| \frac{\partial u}{\partial t}(t, \cdot) \right\|_p^p dt + \int_0^T \|u(t, \cdot)\|_{2,p}^p dt < \infty\}.$$

The norm $\|\cdot\|_{1,2,p}$ makes $H_p^{1,2}(T)$ a Banach space.

The investigation of the deterministic counterpart of equation (2.1)

$$(2.5) \quad \frac{\partial u}{\partial t} = Lu + f$$

with zero initial condition goes in the following way (see [31]). First, for the simplest equation

$$(2.6) \quad \frac{\partial u}{\partial t} = \Delta u + f,$$

its solvability in $H_p^{1,2}(T)$ is proved by means of explicit formulas and some estimates of heat potentials, provided that $f \in L_p((0, T) \times \mathbb{R}^d)$. In Subsec. 4.1 we use the following theorem which is proved in [31] (also see Remark 2.3.2 in [34]).

THEOREM 2.1. *For any $f \in L_p((0, T) \times \mathbb{R}^d)$ and $u_0 \in H_p^{2-2/p}$ there exists a unique solution $u \in H_p^{1,2}(T)$ of the heat equation (2.6) with initial data $u(0) = u_0$. In addition,*

(2.7)

$$\|u_{xx}\|_{L_p((0,T)\times\mathbb{R}^d)} + \left\| \frac{\partial u}{\partial t} \right\|_{L_p((0,T)\times\mathbb{R}^d)} \leq N(d, p)(\|f\|_{L_p((0,T)\times\mathbb{R}^d)} + \|u_0\|_{2-2/p,p}),$$

$$\|u\|_{1,2,p} \leq N(d, p, T)(\|f\|_{L_p((0,T)\times\mathbb{R}^d)} + \|u_0\|_{2-2/p,p}),$$

where u_{xx} is the matrix of second-order derivatives of u with respect to x .

This theorem yields a bounded operator \mathcal{R}_1 which maps any $f \in L_p((0, T) \times \mathbb{R}^d)$ into the solution $u \in H_p^{1,2}(T)$ of the heat equation (2.6) with zero initial data.

Then, the so-called *a priori estimate* is obtained for (2.5). One assumes that there is a solution $u \in H_p^{1,2}(T)$ of (2.5) with zero initial condition and inequality (2.7) is proved, where N is a constant probably depending on T and some characteristics of L (below, N , usually without indices or arguments, denotes various constants, and writing $N(d, p)$ is just to indicate that N depends only on d and p). At this point, one of the most fruitful general ideas in the theory of partial differential equations, linear or not, is used. The idea says that obtaining a priori estimates for solutions in a class of functions implies that the equation is solvable in the same class. We will see how this works below.

Of course, this idea only works “as usual”. For instance, there is no trouble to estimate the sup norm of the second order derivative for possible solutions of the equation $(u'')^2 = f$ on $[0, 1]$ satisfying $u(0) = 0$ and $u(1) = 1$, where f is a given bounded function. So, one has an a priori estimate, but nevertheless the set of solutions is empty if $f < 0$. Also, the statement “if (2.5) has a solution $u \in H_p^{1,2}(T)$, then (2.7) holds” looks bad from the point of view of formal logic (in which any statement with a false assumption is true). In fact, instead of (2.7), one proves

$$(2.8) \quad \|u\|_{1,2,p} \leq N \|Lu - \frac{\partial u}{\partial t}\|_{L_p((0,T)\times\mathbb{R}^d)}$$

for any $u \in H_p^{1,2}(T)$ such that in a certain sense $u(0, \cdot) = 0$. A usual way to prove (2.8) consists of observing that (2.8) is true for $L = \Delta$, then applying perturbation methods to get (2.8) for L that are close enough to Δ , and finally, replacing Δ with other operators with constant coefficients and using partitions of unity in order to get small regions in which L is close to an operator with constant coefficients (we will see this in details in the proof of Theorem 5.1).

The last step is to use the *method of continuity*. Instead of (2.5) one considers the following family of equations

$$(2.9) \quad \frac{\partial u}{\partial t} = L_\lambda u + f$$

with $\lambda \in [0, 1]$, where

$$L_\lambda u = \lambda \Delta + (1 - \lambda)L.$$

Assume that the a priori estimate (2.8) holds with the same constant N for all L_λ in place of L . Also assume that, for a $\lambda = \lambda_0 \in [0, 1]$, equation (2.9) with zero

initial data has a unique solution $u \in H_p^{1,2}(T)$ for any $f \in L_p((0, T) \times \mathbb{R}^d)$. Then we have an operator \mathcal{R}_{λ_0} such that $\mathcal{R}_{\lambda_0} f = u$. By the way, this assumption is satisfied for $\lambda_0 = 1$ by Theorem 2.1. From (2.8) we get that

$$(2.10) \quad \|\mathcal{R}_{\lambda_0} f\|_{1,2,p} \leq N \|f\|_{L_p((0,T) \times \mathbb{R}^d)}.$$

For other $\lambda \in [0, 1]$ we rewrite (2.9) as

$$\frac{\partial u}{\partial t} = L_{\lambda_0} u + \{(\lambda - \lambda_0)(\Delta - L)u + f\}, \quad u = \mathcal{R}_{\lambda_0}\{(\lambda - \lambda_0)(\Delta - L)u + f\}$$

and we solve the last equation by iterations. Define $u_0 = 0$ and

$$u_{n+1} = \mathcal{R}_{\lambda_0}\{(\lambda - \lambda_0)(\Delta - L)u_n + f\}.$$

Then by (2.10)

$$\begin{aligned} \|u_{n+1} - u_n\|_{1,2,p} &\leq N |\lambda - \lambda_0| \|(\Delta - L)(u_n - u_{n-1})\|_{L_p((0,T) \times \mathbb{R}^d)} \\ &\leq N_1 |\lambda - \lambda_0| \|u_n - u_{n-1}\|_{1,2,p}, \end{aligned}$$

and we have a strongly convergent sequence u_n in $H_p^{1,2}(T)$ if $N_1 |\lambda - \lambda_0| \leq 1/2$. Consequently, one can solve (2.9) for such λ . Starting from $\lambda = 1$, one reaches $\lambda = 0$ in finitely many steps, and this finishes the proof of solvability of (2.5).

There are two central objects in the above argument. These are the Banach space $H_p^{1,2}(T)$ and the operator $L - \partial/\partial t : H_p^{1,2}(T) \rightarrow L_p((0, T) \times \mathbb{R}^d)$. Since we want to implement the same kind of argument for equations like (2.1), the first thing to do is to find an appropriate counterpart of $H_p^{1,2}(T)$. This was a problem for some time, since one cannot expect any differentiability property with respect to t for solutions u of (2.1). Then an observation appeared that $H_p^{1,2}(T)$ can also be defined without using $\partial u / \partial t$. Indeed, almost obviously,

$$H_p^{1,2}(T) = \{u : u(t, x) = u(0, x) + \int_0^t f(s, x) ds, u, u_x, u_{xx}, f \in L_p((0, T) \times \mathbb{R}^d)\}.$$

Somewhat unusual about this definition is that in the formula

$$u(t, x) = u(0, x) + \int_0^t f(s, x) ds$$

the function f only belongs to $L_p((0, T) \times \mathbb{R}^d)$ but u, u_x, u_{xx} also belong to $L_p((0, T) \times \mathbb{R}^d)$ (by definition).

Now the guess is natural that a stochastic counterpart $\mathcal{H}_p^2(T)$ of the spaces $H_p^{1,2}(T)$ could be the space of functions $u = u(\omega, t, x)$ such that

$$(2.11) \quad u(t, x) = u(0, x) + \int_0^t f(s, x) ds + \int_0^t g^k(s, x) dw_s^k,$$

$$E \int_0^T \int_{\mathbb{R}^d} \{|u| + |u_x| + |u_{xx}| + |f|\}^p dx dt < \infty,$$

and something of the same type is satisfied for $g = \{g^k\}$. It may look a little bit surprising that one needs $p \geq 2$ and

$$E \int_0^T \int_{\mathbb{R}^d} \{|g| + |g_x|\}^p dx dt < \infty,$$

which involves both g and g_x , where

$$|g|^2 := \sum_{k=1}^{\infty} |g^k|^2, \quad |g_x|^2 := \sum_{k=1}^{\infty} |g_x^k|^2.$$

We will explain later why one needs $p \geq 2$ (see Remark 4.3). The need for some conditions on g_x may be explained by the fact that in parabolic equations one derivative in t is worth two derivatives in x . The stochastic integral in (2.11) has, so to speak, one half derivative in time, so one can expect that it also has one derivative in x . The second derivative should be provided by differentiability of g itself. This is a kind of phenomenological explanation. The real explanation, however, is that with such definition of stochastic $\mathcal{H}_p^2(T)$ spaces one can construct a theory completely analogous to the theory discussed above.

For $p = 2$ we can also see the necessity of imposing conditions on g_x from the model one-dimensional equation

$$du = u_{xx} dt + g dw_t,$$

where $x \in \mathbb{R}$ and w_t is a one-dimensional Wiener process. Forgetting for a while about rigor, by Itô's formula we get

$$\begin{aligned} d(u^2) &= (2uu_{xx} + g^2) dt + 2ug dw_t, \quad \int_{\mathbb{R}} u^2(t, x) dx = \int_{\mathbb{R}} u^2(0, x) dx \\ &+ 2 \int_0^t \int_{\mathbb{R}} uu_{xx}(s, x) dx ds + \int_0^t \int_{\mathbb{R}} g^2(s, x) dx ds + \int_0^t \int_{\mathbb{R}} ug(s, x) dx dw_s. \end{aligned}$$

Integrating by parts we see that

$$E\|u(t, \cdot)\|_2^2 + 2E \int_0^t \|u_x(s, \cdot)\|_2^2 ds = E\|u(0, \cdot)\|_2^2 + E \int_0^t \|g(s, \cdot)\|_2^2 ds.$$

This shows that without assumptions on g_x we *only* get an estimate of the first derivative of u .

With the spaces $\mathcal{H}_p^2(T)$ at hand, we write (2.1) in an operator form by introducing the operator (L, Λ) which can be applied to *any* element $u \in \mathcal{H}_p^2(T)$. Namely for a $u \in \mathcal{H}_p^2(T)$ we write $(L, \Lambda)u = -(f, g)$ if and only if

$$u(t) = u(0) + \int_0^t [Lu + f](s) ds + \int_0^t [\Lambda^k u + g^k](s) dw_s^k.$$

The expression $(L, \Lambda)u$ does make sense for any $u \in \mathcal{H}_p^2(T)$ since, if we have (2.11), then $(L, \Lambda)u = (Lu - f, \Lambda u - g)$. Now, instead of equation (2.1) we have the equation $(L, \Lambda)u = -(f, g)$ in the Banach space $\mathcal{H}_p^2(T)$, which makes the situation completely analogous to the one in the theory of parabolic equations.

This finishes our explanation of the basics of the theory, and now we are ready to present the details.

3. The Stochastic Banach Spaces

Let (Ω, \mathcal{F}, P) be a complete probability space, $(\mathcal{F}_t, t \geq 0)$ be an increasing filtration of σ -fields $\mathcal{F}_t \subset \mathcal{F}$ containing all P -null subsets of Ω , and \mathcal{P} be the predictable σ -field generated by $(\mathcal{F}_t, t \geq 0)$. Let $\{w_t^k; k = 1, 2, \dots\}$ be a family of independent one-dimensional \mathcal{F}_t -adapted Wiener processes defined on (Ω, \mathcal{F}, P) .

We fix a $p \geq 2$ and an integer $d \geq 1$ and use the notation C_0^∞ , \mathcal{D} , L_p , H_p^n , $\|\cdot\|_p$, and $\|\cdot\|_{n,p}$ introduced in Sec. 2. We are dealing with distributions, and, for

a distribution u and $\phi \in C_0^\infty$, by (u, ϕ) or (ϕ, u) we mean the result of application of u to ϕ . Observe that, for $u \in H_p^n$ and $\phi \in C_0^\infty$, by definition

$$(3.1) \quad (u, \phi) = ((1 - \Delta)^{n/2}u, (1 - \Delta)^{-n/2}\phi) = \int_{\mathbb{R}^d} [(1 - \Delta)^{n/2}u](x)(1 - \Delta)^{-n/2}\phi(x) dx,$$

where the last integral is a usual Lebesgue integral. Since $(1 - \Delta)^{n/2}u \in L_p$, one can define (u, ϕ) by (3.1) for any ϕ whose derivatives vanish sufficiently fast at infinity, say exponentially fast.

Sometimes it is useful to notice that $\|\cdot\|_{n,p} \leq \|\cdot\|_{m,p}$ for $m \geq n$. Recall that the set C_0^∞ is dense in H_p^n and the latter is a subset of \mathcal{D} (see, for instance, Theorem 2.3 (ii) of [33]). Also recall (see, for instance [33]) that, for integers $n \geq 0$, the space H_p^n coincides with the Sobolev space $W_p^n = W_p^n(\mathbb{R}^d)$.

We apply the same definitions to l_2 -valued functions g , where l_2 is the set of all real-valued sequences $g = \{g^k; k = 1, 2, \dots\}$ with the norm defined by $\|g\|_{l_2}^2 := \sum_k |g^k|^2$. Specifically,

$$\|g\|_p := \||g|_{l_2}\|_p, \quad \|g\|_{n,p} := \||(1 - \Delta)^{n/2}g|_{l_2}\|_p.$$

Finally, for stopping times τ , we denote $(0, \tau] = \{(\omega, t) : 0 < t \leq \tau(\omega)\}$

$$\mathbb{H}_p^n(\tau) = L_p((0, \tau], \mathcal{P}, H_p^n), \quad \mathbb{H}_p^n = \mathbb{H}_p^n(\infty),$$

$$\mathbb{H}_p^n(\tau, l_2) = L_p((0, \tau], \mathcal{P}, H_p^n(\mathbb{R}^d, l_2)), \quad \mathbb{L}_{\dots\dots} = \mathbb{H}_{\dots\dots}^0 \dots$$

The norms in these spaces are defined in an obvious way. By convention, elements of spaces like \mathbb{H}_p^n are treated as functions rather than distributions or classes of equivalent functions, and if we know that a function of this class has a modification with better properties, then we always consider this modification. For instance, if we take $u \in H_p^n$ and $n - d/p > 0$, then u has a bounded continuous modification, but we talk about $\sup_x u(x)$ instead of \sup of this modification. Also, elements of spaces $\mathbb{H}_p^n(\tau, l_2)$ need not be defined or belong to H_p^n for all $(\omega, t) \in (0, \tau]$. As usual, these properties are needed only for almost all (ω, t) .

For $n \in \mathbb{R}$ and

$$(f, g) \in \mathcal{F}_p^n(\tau) := \mathbb{H}_p^n(\tau) \times \mathbb{H}_p^{n+1}(\tau, l_2),$$

set

$$\|(f, g)\|_{\mathcal{F}_p^n(\tau)} := \|f\|_{\mathbb{H}_p^n(\tau)} + \|g\|_{\mathbb{H}_p^{n+1}(\tau, l_2)}.$$

DEFINITION 3.1. For a \mathcal{D} -valued function $u \in \bigcap_{T>0} \mathbb{H}_p^n(\tau \wedge T)$, we write $u \in \mathcal{H}_p^n(\tau)$ if $u_{xx} \in \mathbb{H}_p^{n-2}(\tau)$, $u(0, \cdot) \in L_p(\Omega, \mathcal{F}_0, H_p^{n-2/p})$, and there exists $(f, g) \in \mathcal{F}_p^{n-2}(\tau)$ such that, for any $\phi \in C_0^\infty$, the equality

$$(3.2) \quad (u(t, \cdot), \phi) = (u(0, \cdot), \phi) + \int_0^t (f(s, \cdot), \phi) ds + \sum_{k=1}^{\infty} \int_0^t (g^k(s, \cdot), \phi) dw_s^k$$

holds for all $t \leq \tau$ with probability 1. We also define $\mathcal{H}_{p,0}^n(\tau) = \mathcal{H}_p^n(\tau) \cap \{u : u(0, \cdot) = 0\}$,

$$(3.3) \quad \|u\|_{\mathcal{H}_p^n(\tau)} = \|u_{xx}\|_{\mathbb{H}_p^{n-2}(\tau)} + \|(f, g)\|_{\mathcal{F}_p^{n-2}(\tau)} + (E\|u(0, \cdot)\|_{n-2/p,p}^p)^{1/p}.$$

As always, we drop τ in $\mathcal{H}_p^n(\tau)$ and $\mathcal{F}_p^n(\tau)$ if $\tau = \infty$.

REMARK 3.2. It is worth noting that the elements of $\mathcal{H}_p^n(\tau)$ are assumed to be defined for all (ω, t) and take values in \mathcal{D} . Obviously, $\mathcal{H}_p^n(\tau)$ is a linear space. As usual, we identify two elements u_1 and u_2 of $\mathcal{H}_p^n(\tau)$ if $\|u_1 - u_2\|_{\mathcal{H}_p^n(\tau)} = 0$. Actually, as we will see later, in this way we only identify indistinguishable functions. Also, observe that the series of stochastic integrals in (3.2) converges uniformly in t in probability on $[0, \tau \wedge T]$ for any finite T , since the quadratic variations of these stochastic integrals satisfy (cf. (3.1))

$$\begin{aligned} \sum_{k=1}^{\infty} \int_0^{\tau \wedge T} (g^k(s, \cdot), \phi)^2 ds &= \sum_{k=1}^{\infty} \int_0^{\tau \wedge T} ((1 - \Delta)^{(n-1)/2} g^k(s, \cdot), (1 - \Delta)^{(1-n)/2} \phi)_{L_2}^2 ds \\ &\leq \|(1 - \Delta)^{(1-n)/2} \phi\|_1 \int_0^{\tau \wedge T} \sum_{k=1}^{\infty} (|(1 - \Delta)^{(n-1)/2} g^k(s, \cdot)|^2, |(1 - \Delta)^{(1-n)/2} \phi|)_{L_2} ds \\ &\leq N \int_0^{\tau \wedge T} \|(\sum_{k=1}^{\infty} |(1 - \Delta)^{(n-1)/2} g^k(s, \cdot)|^2)^{1/2}\|_p^2 ds < \infty \quad (\text{a.s.}), \end{aligned}$$

where $N = \|(1 - \Delta)^{(1-n)/2} \phi\|_1 \|(1 - \Delta)^{(1-n)/2} \phi\|_q$, $q = p/(p-2)$ and we have used that $p \geq 2$.

As a consequence of the uniform convergence, $(u(t, \cdot), \phi)$ is continuous in t on $[0, \tau \wedge T]$ for any finite T (a.s.).

REMARK 3.3. There can exist only one couple (f, g) for which (3.2) holds. Indeed, if there are two, then one can represent zero as a sum of a continuous process of bounded variation and a continuous local martingale. One knows that this is only possible if both processes vanish. Therefore, the couple (f, g) is uniquely determined by u , and notation $\|u\|_{\mathcal{H}_p^n(\tau)}$ in (3.3) makes sense.

REMARK 3.4. It is known that the operator $(1 - \Delta)^{m/2}$ maps isometrically H_p^n onto H_p^{n-m} for any n, m . Also, the inequalities from Remark 3.2 can be used to show that given $u \in \mathcal{H}_p^n(\tau)$, one can in (3.2) take any infinitely differentiable function ϕ whose derivatives vanish sufficiently fast at infinity, say exponentially fast. This allows us to substitute $(1 - \Delta)^{m/2} \phi$ in (3.2) instead of ϕ and shows that the operator $(1 - \Delta)^{m/2}$ maps isometrically $\mathcal{H}_p^n(\tau)$ onto $\mathcal{H}_p^{n-m}(\tau)$ for any n, m . The same is true for $\mathbb{H}_p^n(\tau)$.

DEFINITION 3.5. For $u \in \mathcal{H}_p^n(\tau)$, if (3.2) holds, then we write $f = \mathbb{D}u$, $g = \mathbb{S}u$ (for “deterministic” and “stochastic” parts of u) and we also write

$$u(t) = u(0) + \int_0^t \mathbb{D}u(s) ds + \int_0^t \mathbb{S}^k u(s) dw_s^k, \quad du = f dt + g^k dw_t^k \quad t \leq \tau.$$

REMARK 3.6. It follows from Definitions 3.1 and 3.5 that the operators \mathbb{D} and \mathbb{S} are continuous operators from $\mathcal{H}_p^n(\tau)$ to $\mathbb{H}_p^{n-2}(\tau)$ and $\mathbb{H}_p^{n-1}(\tau, l_2)$ respectively. From Theorem 4.2 and Remark 3.4 it follows that \mathbb{S} maps $\mathcal{H}_p^n(\tau)$ onto $\mathbb{H}_p^{n-1}(\tau, l_2)$. However, at this point we do not know how rich $\mathcal{H}_p^n(\tau)$ is. Nevertheless obviously $H_p^{1,2}(T) \subset \mathcal{H}_p^2(T)$.

THEOREM 3.7. The spaces $\mathcal{H}_p^n(\tau)$ and $\mathcal{H}_{p,0}^n(\tau)$ are Banach spaces with norm (3.3). In addition if $\tau \leq T$, where T is a finite constant, then for $u \in \mathcal{H}_p^n(\tau)$

$$(3.4) \quad \|u\|_{\mathbb{H}_p^n(\tau)} \leq N(d, T) \|u\|_{\mathcal{H}_p^n(\tau)}, \quad E \sup_{t \leq \tau} \|u(t, \cdot)\|_{n-2,p}^p \leq N(d, T) \|u\|_{\mathcal{H}_p^n(\tau)}^p.$$

Proof. We first deal with (3.4). Obviously

$$\|u\|_{\mathbb{H}_p^n(\tau)} = \|(1 - \Delta)u\|_{\mathbb{H}_p^{n-2}(\tau)} \leq \|u\|_{\mathbb{H}_p^{n-2}(\tau)} + \|u\|_{\mathcal{H}_p^n(\tau)},$$

so that to prove (3.4) we only need to prove that

$$(3.5) \quad E \sup_{t \leq \tau} \|u(t, \cdot)\|_{n-2,p}^p \leq N \|u\|_{\mathcal{H}_p^n(\tau)}^p.$$

Owing to Remark 3.4 we may and will assume that $n = 2$. Take a nonnegative function $\zeta \in C_0^\infty$ with unit integral, for $\varepsilon > 0$ define $\zeta_\varepsilon(x) = \varepsilon^{-d}\zeta(x/\varepsilon)$, and for generalized functions u let $u^{(\varepsilon)}(x) = u * \zeta_\varepsilon(x)$. Observe that $u^{(\varepsilon)}(x)$ is a continuous (infinitely differentiable) function of x for any distribution u . Plugging in $\zeta_\varepsilon(\cdot - x)$ instead of ϕ in (3.2), we get that for any x the equality

$$(3.6) \quad u^{(\varepsilon)}(t, x) = u^{(\varepsilon)}(0, x) + \int_0^t f^{(\varepsilon)}(s, x) ds + \sum_{k=1}^{\infty} \int_0^t g^{(\varepsilon)k}(s, x) dw_s^k$$

holds almost surely for all $t \leq \tau$. If necessary, we redefine the stochastic integrals in (3.6) in such a way that (3.6) would hold for all ω, t , and x such that $t \leq \tau$. Here

$$E\|u^{(\varepsilon)}(0, \cdot)\|_p^p \leq E\|u(0, \cdot)\|_p^p \leq E\|u(0, \cdot)\|_{n-2/p,p}^p \leq \|u\|_{\mathcal{H}_p^n(\tau)}^p,$$

where we use that, by Minkowski's inequality, $\|h^{(\varepsilon)}\|_p \leq \|\zeta_\varepsilon\|_1 \|h\|_p = \|h\|_p$. Similarly,

$$\begin{aligned} \left| \int_0^t f^{(\varepsilon)}(s, x) ds \right|^p &\leq T^{p-1} \int_0^\tau |f^{(\varepsilon)}(s, x)|^p ds, \\ E \sup_{t \leq \tau} \left\| \int_0^t f^{(\varepsilon)}(s, \cdot) ds \right\|_p^p &\leq T^{p-1} E \int_0^\tau \|f(s, \cdot)\|_p^p ds \leq T^{p-1} \|u\|_{\mathcal{H}_p^n(\tau)}^p. \end{aligned}$$

Finally, by Burkholder–Davis–Gundy inequalities

$$\begin{aligned} E \sup_{t \leq \tau} \left| \sum_{k=1}^{\infty} \int_0^t g^{(\varepsilon)k}(s, x) dw_s^k \right|^p &\leq NE \left| \int_0^\tau \sum_{k=1}^{\infty} |g^{(\varepsilon)k}(s, x)|^2 ds \right|^{p/2} \\ &= NE \left| \int_0^\tau |g^{(\varepsilon)}|_{l_2}^2(s, x) ds \right|^{p/2}, \end{aligned}$$

and as above (the first term below makes sense by virtue of (3.6))

$$\begin{aligned} E \sup_{t \leq \tau} \left\| \sum_{k=1}^{\infty} \int_0^t g^{(\varepsilon)k}(s, \cdot) dw_s^k \right\|_p^p &\leq \int_{\mathbb{R}^d} E \sup_{t \leq \tau} \left| \sum_{k=1}^{\infty} \int_0^t g^{(\varepsilon)k}(s, x) dw_s^k \right|^p dx \\ &\leq NE \int_{\mathbb{R}^d} \left| \int_0^\tau |g^{(\varepsilon)}|_{l_2}^2(s, x) ds \right|^{p/2} dx \leq NE \left(\int_0^\tau \| |g^{(\varepsilon)}|_{l_2}^2(s, \cdot) \|_{p/2} ds \right)^{p/2} \\ &= NE \left(\int_0^\tau \| |g^{(\varepsilon)}|_{l_2}(s, \cdot) \|_p^2 ds \right)^{p/2} \leq NE \int_0^\tau \| |g^{(\varepsilon)}|_{l_2}(s, \cdot) \|_p^p ds \\ &\leq N \|g\|_{\mathbb{L}_p(\tau, l_2)}^p \leq N \|u\|_{\mathcal{H}_p^n(\tau)}^p. \end{aligned}$$

This, along with (3.6), leads to

$$(3.7) \quad E \sup_{t \leq \tau} \|u^{(\varepsilon)}(t, \cdot)\|_p^p \leq N \|u\|_{\mathcal{H}_p^n(\tau)}^p.$$

Furthermore, by using the fact that $\|h^{(\varepsilon)} - h^{(\gamma)}\|_p \rightarrow 0$ whenever $h \in L_p$ and $\varepsilon, \gamma \rightarrow 0$ and by considering $u^{(1/m)} - u^{(1/k)}$ instead of $u^{(\varepsilon)}$, we easily see that that $u^{(1/m)}(t \wedge \tau, x)$ is a Cauchy sequence in $L_p(\Omega, B([0, T], L_p))$. Define \bar{u} as its limit

in this space. Then, for a subsequence m' , we have $u^{(1/m')}(t, \cdot) \rightarrow \bar{u}(t, \cdot)$ in L_p if $t \leq \tau$ with probability 1. On the other hand $u^{(1/m)}(t, \cdot) \rightarrow u(t, \cdot)$ in the sense of distributions for all ω and t such that $t \leq \tau(\omega)$. Therefore with probability one we have $u(t, \cdot) \in L_p$ for $t \leq \tau$. Now, (3.7) and Fatou's lemma yield (3.5) for $n = 2$. As explained above, this proves (3.4).

Next, we derive the first assertion of our theorem from (3.4). As usual, we have only to check the completeness of $\mathcal{H}_p^n(\tau)$. If $\{u_j\}$ is a Cauchy sequence in $\mathcal{H}_p^n(\tau)$, then it is a Cauchy sequence in $\mathbb{H}_p^n(\tau \wedge T)$ for any T , and there is $u \in \bigcap_{T>0} \mathbb{H}_p^n(\tau \wedge T)$ such that $\|u - u_j\|_{\mathbb{H}_p^n(\tau \wedge T)} \rightarrow 0$ as $j \rightarrow \infty$. Furthermore, u_{jxx} form a Cauchy sequence and therefore converge in $\mathbb{H}_p^{n-2}(\tau)$. It follows easily that $\|u_{xx} - u_{jxx}\|_{\mathbb{H}_p^{n-2}(\tau)} \rightarrow 0$.

Also, for $u_j(0), f_j, g_j$ corresponding to u_j , there is $u(0) \in L_p(\Omega, \mathcal{F}_0, H_p^{n-2/p})$ and $(f, g) \in \mathcal{F}_p^{n-2}(\tau)$ such that

$$E\|u(0) - u_j(0)\|_{n-2/p,p}^p \rightarrow 0, \quad \|f - f_j\|_{\mathbb{H}_p^{n-2}(\tau)} \rightarrow 0, \quad \|g - g_j\|_{\mathbb{H}_p^{n-1}(\tau, l_2)} \rightarrow 0.$$

By using the argument from Remark 3.2, one can show that for any $\phi \in C_0^\infty$ equality (3.2) holds in $(0, \tau]$ almost everywhere.

On the other hand, (3.4) also implies that for u (at least for a modification of u) it holds that

$$E \sup_{t \leq \tau \wedge T} \|u(t, \cdot) - u_j(t, \cdot)\|_{n-2,p}^p \rightarrow 0$$

for any constant $T < \infty$. Adding to this that the processes $(u_j(t, \cdot), \phi)$ are continuous (a.s.) (see Remark 3.2), we conclude that $(u(t, \cdot), \phi)$ is also continuous (a.s.). Thus, for any $\phi \in C_0^\infty$, equality (3.2) not only holds in $(0, \tau]$ almost everywhere but also for all $t \leq \tau$ almost surely. Hence, $u \in \mathcal{H}_p^n(\tau)$ and $u_j \rightarrow u$ in $\mathcal{H}_p^n(\tau)$. The theorem is proved.

REMARK 3.8. We could replace the first term on the right in (3.3) with $\|u\|_{\mathbb{H}_p^n(\tau)}$ and, for bounded τ , we would get an equivalent norm by virtue of (3.4). The form of (3.3) that we have chosen is convenient in the future when we need certain constants to be independent of T , see, for instance, Theorem 4.10.

REMARK 3.9. In Theorems 7.1 and 7.2 below, we prove much sharper estimates than (3.5).

We also need the following properties of the spaces $\mathcal{H}_p^n(\tau)$ and $\mathbb{H}_p^n(\tau)$.

THEOREM 3.10. Take $g \in \mathbb{H}_p^n(l_2)$. Then there exists a sequence $g_j \in \mathbb{H}_p^n(l_2)$, $j = 1, 2, \dots$, such that $\|g - g_j\|_{\mathbb{H}_p^n(l_2)} \rightarrow 0$ as $j \rightarrow \infty$ and

$$g_j^k = \begin{cases} \sum_{i=1}^j I_{(\tau_{i-1}^j, \tau_i^j]}(t) g_j^{ik}(x) & \text{if } k \leq j, \\ 0 & \text{if } k > j, \end{cases}$$

where τ_i^j are bounded stopping times, $\tau_{i-1}^j \leq \tau_i^j$, and $g_j^{ik} \in C_0^\infty$.

Proof. The argument in Remark 3.4 and the fact that C_0^∞ is dense in any H_p^n show that we only need to consider $n = 0$. Further, one can easily understand that the set of $g \in \mathbb{L}_p(l_2)$ for which the statement holds forms a linear closed subspace \mathbb{L} of $\mathbb{L}_p(l_2)$. We have to prove that $\mathbb{L} = \mathbb{L}_p(l_2)$. If this is not true, then, by Riesz's theorem, there is a nonzero $h \in \mathbb{L}_q(l_2)$ with $q = p/(p-1)$ such that

$$E \int_0^\infty \int_{\mathbb{R}^d} (h, g)_{l_2} dx dt = 0$$

for any $g \in \mathbb{L}$. In particular,

$$E \int_0^\infty I_{(0,\tau]} \left(\int_{\mathbb{R}^d} h^k g \, dx \right) dt = 0$$

for any bounded stopping time τ , $k \geq 1$, and $g \in C_0^\infty$. Since $\int_{\mathbb{R}^d} h^k g \, dx$ is equal to a predictable function (a.e.), it follows that $\int_{\mathbb{R}^d} h^k g \, dx = 0$ on $(0, \infty]$ (a.e.). By taking g from a countable subset \mathcal{G} in C_0^∞ that is dense in L_p , we get that, on a subset of $(0, \infty]$ of full measure,

$$\int_{\mathbb{R}^d} h^k g \, dx = 0 \quad \forall g \in \mathcal{G}, k \geq 1.$$

But then $h^k = 0$ on $(0, \infty] \times \mathbb{R}^d$ (a.e.). This contradicts that $h \neq 0$ and proves the theorem.

THEOREM 3.11. *Let $T \in (0, \infty)$. If $u_j \in \mathcal{H}_p^n(T)$, $j = 1, 2, \dots$, and $\|u_j\|_{\mathcal{H}_p^n(T)} \leq K$, where K is a finite constant, then there exists a subsequence j' and a function $u \in \mathcal{H}_p^n(T)$ such that*

- (i) $u_{j'}, u_{j'}(0, \cdot), \mathbb{D}u_{j'},$ and $\mathbb{S}u_{j'}$ converge weakly to u , $u(0, \cdot)$, $\mathbb{D}u$, and $\mathbb{S}u$ in $\mathbb{H}_p^n(T)$, $L_p(\Omega, H_p^{n-2/p})$, $\mathbb{H}_p^{n-2}(T)$, and $\mathbb{H}_p^{n-1}(T, l_2)$ respectively;
- (ii) $\|u\|_{\mathcal{H}_p^n(T)} \leq K$;
- (iii) for any $\phi \in C_0^\infty$ and any $t \in [0, T]$ we have $(\phi, u_{j'}(t, \cdot)) \rightarrow (\phi, u(t, \cdot))$ weakly in $L_p(\Omega)$.

Proof. From properties of L_p spaces, it follows that there exists a subsequence j' such that $u_{j'}, u_{j'}(0, \cdot), \mathbb{D}u_{j'}, \mathbb{S}u_{j'}$ converge weakly to some u, u_0, f, g in $\mathbb{H}_p^n(T)$, $L_p(\Omega, H_p^{n-2/p})$, $\mathbb{H}_p^{n-2}(T)$, and $\mathbb{H}_p^{n-1}(T, l_2)$ respectively. Then, for any $\phi \in C_0^\infty$, the expressions $(u_{j'}(t, \cdot), \phi)$, $(u_{j'}(0, \cdot), \phi)$, $(\mathbb{D}u_{j'}(s, \cdot), \phi)$, and $(\mathbb{S}^k u_{j'}(s, \cdot), \phi)$ in the formula

$$(u_{j'}(t, \cdot), \phi) = (u_{j'}(0, \cdot), \phi) + \int_0^t (\mathbb{D}u_{j'}(s, \cdot), \phi) \, ds + \sum_{k=1}^{\infty} \int_0^t (\mathbb{S}^k u_{j'}(s, \cdot), \phi) \, dw_s^k$$

converge weakly in corresponding spaces. Since integration and stochastic integration can be considered as continuous linear operators and any continuous linear operator is also weakly continuous, we have that, for any $\phi \in C_0^\infty$,

$$(3.8) \quad (u(t, \cdot), \phi) = (u_0, \phi) + \int_0^t (f(s, \cdot), \phi) \, ds + \sum_{k=1}^{\infty} \int_0^t (g^k(s, \cdot), \phi) \, dw_s^k$$

for almost all $(\omega, t) \in \Omega \times [0, T]$.

By Banach–Saks theorem, there is a sequence $(v_{j'}, \mathbb{D}v_{j'}, \mathbb{S}v_{j'})$ of convex combinations of $(u_{j'}, \mathbb{D}u_{j'}, \mathbb{S}u_{j'})$ which converges strongly to (u, f, g) in $\mathbb{H}_p^n(T) \times \mathbb{H}_p^{n-2}(T) \times \mathbb{H}_p^{n-1}(T, l_2)$. From (3.4), it follows that

$$E \sup_{t \leq T} \|v_j - v_i\|_{n-2,p}^p \rightarrow 0$$

as $i, j \rightarrow \infty$. Therefore, there is a H_p^{n-2} -valued function v on $\Omega \times [0, T]$ such that

$$E \sup_{t \leq T} \|v_j - v\|_{n-2,p}^p \rightarrow 0.$$

In particular, $(v_j(t, \cdot), \phi) \rightarrow (v(t, \cdot), \phi)$ uniformly on $[0, T]$ in probability for any $\phi \in C_0^\infty$. On the other hand, the strong convergence of v_j to v in $\mathbb{H}_p^n(T)$ implies

that $(v_j(t, \cdot), \phi) \rightarrow (u(t, \cdot), \phi)$ on $\Omega \times [0, T]$ in measure. This shows that $(v(t, \cdot), \phi) = (u(t, \cdot), \phi)$ on $\Omega \times [0, T]$ (a.e.). Because of arbitrariness of ϕ and the fact that C_0^∞ is dense in the separable spaces conjugate to H_p^m , $u = v$ (as generalized functions) on $\Omega \times [0, T]$ (a.e.).

Thus, $v \in \mathbb{H}_p^n(T)$. Also, $(v_j(t), \phi)$ are given by equations similar to (3.8), which implies that $(v_j(t), \phi)$ are continuous in t (a.s.). The uniform convergence of $(v_j(t), \phi)$ to $(v(t), \phi)$ in probability yields the continuity of $(v(t), \phi)$ (a.s.). By the above, (3.8) holds for almost all $(\omega, t) \in \Omega \times [0, T]$ if we replace $(u(t, \cdot), \phi)$ by $(v(t, \cdot), \phi)$. Since the latter is continuous and the right-hand side of (3.8) is continuous, $(v(t, \cdot), \phi)$ equals the right-hand side of (3.8) for all $t \in [0, T]$ (a.s.). Hence, $v \in \mathcal{H}_p^n(T)$ and we have proved assertion (i) for v instead u , which is irrelevant.

Assertion (ii) follows from the equality $u = v$ on $\Omega \times [0, T]$ (a.e.) and from the fact that the norm of a weak limit is less than the liminf of norms.

To prove (iii), take $\phi \in C_0^\infty$ and $\xi \in L_q(\Omega)$ with $q = p/(p-1)$ and write

$$E\xi(u_j(t, \cdot), \phi) = E\xi(u_j(0, \cdot), \phi) + E\xi \int_0^t (\mathbb{D}u_j(s, \cdot), \phi) ds + E\xi \int_0^t (\mathbb{S}^k u_j(s, \cdot), \phi) dw_s^k.$$

By what has been said about the properties of the operators of integration and by (i),

$$\begin{aligned} \lim_{j' \rightarrow \infty} E\xi(u_{j'}(t, \cdot), \phi) &= \lim_{j' \rightarrow \infty} [E\xi(u_{j'}(0, \cdot), \phi) + E\xi \int_0^t (\mathbb{D}u_{j'}(s, \cdot), \phi) ds \\ &\quad + E\xi \int_0^t (\mathbb{S}^k u_{j'}(s, \cdot), \phi) dw_s^k] = E\xi(u(0, \cdot), \phi) \\ &\quad + E\xi \int_0^t (\mathbb{D}u(s, \cdot), \phi) ds + E\xi \int_0^t (\mathbb{S}^k u(s, \cdot), \phi) dw_s^k = E\xi(u(t, \cdot), \phi), \end{aligned}$$

which proves (iii) and the theorem.

4. Model Equations

As explained in Sec. 2, to implement our general scheme we have to show that (at least) simple equations are solvable in $\mathcal{H}_{p,0}^n(\tau)$.

Except for Subsec. 4.2, we will always understand equations like (2.1) in the sense of Definition 3.5, which means that we will be looking for a function $u \in \mathcal{H}_{p,0}^n(\tau)$ such that

$$\mathbb{D}u = Lu + f, \quad \mathbb{S}u = \Lambda u + g.$$

In this section, we consider equation (2.1) when $b^i = c = \nu^k = 0$ and the coefficients a and σ do not depend on x . Throughout the section, we fix real-valued functions $a^{ij}(t)$ and l_2 -valued functions $\sigma^i(t) = \{\sigma^{ik}(t), k \geq 1\}$ defined for $i, j = 1, \dots, d$ on $\Omega \times (0, \infty)$. Define

$$\alpha^{ij}(t) = \frac{1}{2}(\sigma^i(t), \sigma^j(t))_{l_2}$$

and assume that a and σ are \mathcal{P} -measurable functions, and in the matrix sense

$$(a^{ij}) = (a^{ij})^*, \quad K(\delta^{ij}) \geq (a^{ij}) \geq (a^{ij} - \alpha^{ij}) \geq \delta(\delta^{ij}),$$

where K and δ are some fixed strictly positive constants. By the way, the assumption that $a > \alpha$ is necessary even to have L_2 -theory for SPDEs with constant coefficients (see [30] or Remark 4.8).

Equation (2.1) takes the following form:

(4.1)

$$du(t, x) = (a^{ij}(t)u_{x^i x^j}(t, x) + f(t, x)) dt + (\sigma^{ik}(t)u_{x^i}(t, x) + g^k(t, x)) dw_t^k, \quad t > 0.$$

Our plan is as follows. In Subsec. 4.1 we consider the case of the heat equation with random right-hand side and get basic a priori estimates. The results of Subsec. 4.2 show how to reduce general case of equation (4.1) to the case of the heat equation and get some representation formulas for solutions. At that point it turns out that the assumption that $u \in \mathcal{H}_p^n$ is not very convenient, and we consider larger spaces. In the final Subsec. 4.3, we present the basic a priori estimates and the existence and uniqueness theorem for equation (4.1).

4.1. Particular Case $a^{ij} = \delta^{ij}$, $\sigma = 0$. We start with the equation

$$(4.2) \quad du(t, x) = (\Delta u(t, x) + f(t, x)) dt + g^k(t, x) dw_t^k, \quad t > 0.$$

We need a lemma from [17] or [20]. Remember that the operators T_t are defined by (2.3) and, as always, $p \geq 2$.

LEMMA 4.1. *Let $-\infty \leq a < b \leq \infty$, $g \in L_p((a, b) \times \mathbb{R}^d, l_2)$. Then*

$$\int_{\mathbb{R}^d} \int_a^b \left[\int_a^t |\nabla T_{t-s}g(s, \cdot)(x)|_{l_2}^2 ds \right]^{p/2} dt dx \leq N(d, p) \int_{\mathbb{R}^d} \int_a^b |g(t, x)|_{l_2}^p dt dx.$$

THEOREM 4.2. *Take $f \in \mathbb{H}_p^{-1}$, $g \in \mathbb{L}_p(l_2)$. Then*

- (i) *equation (4.2) with zero initial condition has a unique solution $u \in \mathcal{H}_p^1$;*
- (ii) *for this solution, we have*

$$(4.3) \quad \|u_{xx}\|_{\mathbb{H}_p^{-1}} \leq N(d, p)(\|f\|_{\mathbb{H}_p^{-1}} + \|g\|_{\mathbb{L}_p(l_2)});$$

(iii) *for this solution, we have $u \in C_{loc}([0, \infty), L_p)$ almost surely, and, for any $\lambda, T > 0$,*

$$E \sup_{t \leq T} (e^{-p\lambda t} \|u(t, \cdot)\|_p^p) + E \int_0^T e^{-p\lambda t} \| |u|^{(p-2)/p} |u_x|^{2/p}(t, \cdot) \|_p^p dt \leq$$

$$(4.4) \quad N(d, p, \lambda)(\|e^{-\lambda t} f\|_{\mathbb{H}_p^{-1}(T)}^p + \|e^{-\lambda t} g\|_{\mathbb{L}_p(T, l_2)}^p).$$

Proof. It is well known that there exists a continuous linear operator

$$P : H_p^{-1} \rightarrow (L_p)^{d+1}$$

such that if $h \in H_p^{-1}$ and $Ph = (h_0, \tilde{h}^1, \dots, \tilde{h}^d)$, then $h = h_0 + \operatorname{div} \tilde{h}$ and

$$(4.5) \quad \|\tilde{h}\|_p + \|h_0\|_p \leq N(d, p)\|h\|_{-1, p}, \quad \|h\|_{-1, p} \leq N(d, p)\{\|\tilde{h}\|_p + \|h_0\|_p\}.$$

Actually, one can take $\tilde{h} = -\operatorname{grad}(1 - \Delta)^{-1}h$ and $h_0 = h - \operatorname{div} \tilde{h} = (1 - \Delta)^{-1}h$. Indeed, $\|h_0\|_p = \|h\|_{-2, p} \leq \|h\|_{-1, p}$. Also, the fact that $\partial/\partial x^i$ is a bounded operator from H_p^n to H_p^{n+1} for any n (see, for instance, [33]) means that $(\partial/\partial x^i)(1 - \Delta)^{-1/2}$ is a bounded operator from H_p^n to H_p^n and $(\partial/\partial x^i)(1 - \Delta)^{-1}$ is a bounded operator from H_p^n to H_p^{n-1} . This is why $\|\tilde{h}\|_p \leq N(d, p)\|h\|_{-1, p}$. This gives the first estimate in (4.5). On the other hand, $(1 - \Delta)^{-1/2}h = (1 - \Delta)^{-1/2}h_0 + \operatorname{div}(1 - \Delta)^{-1/2}\tilde{h}$, and both operators $(1 - \Delta)^{-1/2}$ and $\operatorname{div}(1 - \Delta)^{-1/2}$ are bounded on L_p .

Define $(f_0, \tilde{f}) = Pf$. Then equation (4.2) takes the form

$$(4.6) \quad du = (\Delta u + f_0 + \operatorname{div} \tilde{f}) dt + g^k dw_t^k,$$

and we supply it with zero initial condition. We will prove that, for arbitrary $f_0, \tilde{f}^i \in \mathbb{L}_p$, our assertions hold for (4.6) in place of (4.2). Of course, in (4.3) and (4.4), by f we mean $f_0 + \operatorname{div} \tilde{f}$.

A particular case. First we consider the case in which

$$(4.7) \quad \begin{aligned} f_0(t, x) &= \sum_{i=1}^m I_{(\tau_{i-1}, \tau_i]}(t) f_{0i}(x), \quad \tilde{f}(t, x) = \sum_{i=1}^m I_{(\tau_{i-1}, \tau_i]}(t) \tilde{f}_i(x), \\ g(t, x) &= \sum_{k=1}^m g^k(t, x) h_k, \quad g^k(t, x) = \sum_{i=1}^m I_{(\tau_{i-1}, \tau_i]}(t) g^{ik}(x), \end{aligned}$$

where (h_k) is the standard orthonormal basis in l_2 , $m < \infty$, τ_i are bounded stopping times, $\tau_{i-1} \leq \tau_i$, and $f_{0i}, \tilde{f}_i, g^{ik} \in C_0^\infty$.

Set

$$(4.8) \quad \begin{aligned} v(t, x) &= \int_0^t g^k(s, x) dw_s^k = \sum_{i,k=1}^m g^{ik}(x) (w_{t \wedge \tau_i}^k - w_{t \wedge \tau_{i-1}}^k), \\ u(t, x) &= v(t, x) + \int_0^t T_{t-s}[\Delta v + f](s, \cdot)(x) ds, \quad \forall t \geq 0. \end{aligned}$$

As easy to see, the function $u - v$ is infinitely differentiable in (t, x) and satisfies the equation

$$\frac{\partial z}{\partial t} = \Delta z + \Delta v + f.$$

It follows that, for any x , the function $u(t, x)$ satisfies almost surely the following form of (4.6):

$$(4.9) \quad u(t, x) = \int_0^t (\Delta u(s, x) + f(s, x)) ds + \sum_{k=1}^m \int_0^t g^k(s, x) dw_s^k.$$

Next, we want to obtain some bounds on norms of u . Let

$$u_1(t, x) = \int_0^t T_{t-s} f(s, x) ds.$$

According to Theorem 2.1 (for any ω),

$$(4.10) \quad \|u_{1xx}\|_{L_p(\mathbb{R}_+, H_p^{-1})} \leq N \|f\|_{L_p(\mathbb{R}_+, H_p^{-1})}.$$

Furthermore, use again that the operators $(\partial/\partial x^i)(1 - \Delta)^{-1/2}$ are bounded in L_p for any $p > 1$. Then

$$(4.11) \quad \|u_{xx} - u_{1xx}\|_{H_p^{-1}}^p \leq N \|u_x - u_{1x}\|_{L_p}^p = N \int_0^\infty \int_{\mathbb{R}^d} E |u_x - u_{1x}|^p(t, x) dx dt.$$

To make some further transformations of this formula, we note that if $z^k = z^k(x)$ are bounded Borel functions, then, by Itô's formula applied to the increment over $[0, t]$ of

$$\left(\int_r^t T_{t-s} z^k ds \right) (w_{r \wedge \tau_2}^k - w_{r \wedge \tau_1}^k)$$

(with no summation in k) as a function of r , we obtain (a.s.)

$$0 = - \int_0^t (w_{r \wedge \tau_2}^k - w_{r \wedge \tau_1}^k) T_{t-r} z^k dr + \int_0^t I_{(\tau_1, \tau_2]}(r) \left(\int_r^t T_{t-s} z^k ds \right) dw_r^k.$$

By using this for our particular g , or by using the stochastic version of the Fubini theorem and coming back to (4.8), for any $t \geq 0$ and $x \in \mathbb{R}^d$, we get

$$\begin{aligned} u_x(t, x) - u_{1x}(t, x) &= v_x(t, x) + \int_0^t T_{t-s} \sum_{k=1}^m \int_0^s \Delta g_x^k(r, x) dw_r^k dr \\ &= v_x(t, x) - \sum_{k=1}^m \int_0^t \int_r^t \frac{d}{ds} T_{t-s} g_x^k(r, x) ds dw_r^k = \sum_{k=1}^m \int_0^t T_{t-r} g_x^k(r, x) dw_r^k \quad (\text{a.s.}). \end{aligned}$$

Hence, by the Burkholder–Davis–Gundy inequality,

$$\begin{aligned} E|u_x - u_{1x}|^p(t, x) &\leq NE \left[\int_0^t \sum_{k=1}^m |T_{t-r} g_x^k(r, x)|^2 dr \right]^{p/2} \\ &= NE \left[\int_0^t |T_{t-r} g_x(r, x)|_{l_2}^2 dr \right]^{p/2}. \end{aligned}$$

By plugging this into (4.11) and by applying Lemma 4.1, we obtain

$$\|u_x - u_{1x}\|_{\mathbb{L}_p}^p \leq NE \int_0^\infty \int_{\mathbb{R}^d} \left[\int_0^t |\nabla T_{t-s} g(s, x)|_{l_2}^2 ds \right]^{p/2} dx dt \leq N \|g\|_{\mathbb{L}_p(l_2)}^p.$$

This along with (4.10) gives us (4.3). However, we do not know yet that $u \in \mathcal{H}_p^1$.

Our next step is to prove (4.4) for sufficiently large λ . We repeat briefly the corresponding arguments from [25] or [30]. From (4.9) and Itô's formula,

$$\begin{aligned} &|u(t, x)|^p e^{-\lambda t} \\ &= \int_0^t e^{-\lambda s} (p|u|^{p-2} u \Delta u + p|u|^{p-2} u f + \frac{1}{2} p(p-1)|u|^{p-2} |g|_{l_2}^2 - \lambda|u|^p)(s, x) ds \\ &\quad + p \sum_{k \leq m} \int_0^t e^{-\lambda s} |u|^{p-2} u g^k(s, x) dw_s^k. \end{aligned}$$

We integrate here with respect to x and use the stochastic Fubini theorem and the fact that $u(t, x)$, $g(t, x)$, and their derivatives decrease very fast when $|x| \rightarrow \infty$. Then we integrate by parts in $\int |u|^{p-2} u \Delta u dx$ and also notice that, for $q = p/(p-1)$,

$$\int_{\mathbb{R}^d} |u|^{p-2} u f(s, x) dx = -(p-1) \int_{\mathbb{R}^d} |u|^{p-2} u_x \cdot \tilde{f}(s, x) dx + \int_{\mathbb{R}^d} |u|^{p-2} u f_0(s, x) dx,$$

$$\begin{aligned} &|\int_{\mathbb{R}^d} |u|^{p-2} u_x \cdot \tilde{f}(s, x) dx| \leq \int_{\mathbb{R}^d} (|u|^{(p-2)/2} |u_x|)^q |u|^{q(p-2)/2} dx \\ &\quad + \|\tilde{f}(s, \cdot)\|_p^p \leq N \|f(s, \cdot)\|_{-1,p}^p + N_1 \|u(s, \cdot)\|_p^p + \frac{1}{2} \||u|^{(p-2)/p} |u_x|^{2/p}(s, \cdot)\|_p^p, \end{aligned}$$

$$\begin{aligned} \int_{\mathbb{R}^d} |u(s, x)|^{p-2} u(s, x) f_0(s, x) dx &\leq \|f_0(s, \cdot)\|_p^p + \|u(s, \cdot)\|_p^p \\ &\leq N \|f(s, \cdot)\|_{-1,p}^p + \|u(s, \cdot)\|_p^p. \end{aligned}$$

In this way, for $\lambda \geq p(p-1)N_1 + p + p(p-1)/2$, we get

$$\begin{aligned} & \|u(t, \cdot)\|_p^p e^{-\lambda t} + \frac{p(p-1)}{2} \int_0^t \| |u|^{(p-2)/p} |u_x|^{2/p}(s, \cdot) \|_p^p e^{-\lambda s} ds \\ & \leq N \int_0^t [\|f(s, \cdot)\|_{-1,p}^p + \|g(s, \cdot)\|_p^p] e^{-\lambda s} ds \\ & \quad + p \sum_{k \leq m} \int_0^t e^{-\lambda s} \left\{ \int_{\mathbb{R}^d} |u|^{p-2} u g^k(s, x) dx \right\} dw_s^k, \end{aligned}$$

where $N = N(p)$. After this, basically, one takes expectations and applies certain standard transformations based on the Burkholder–Davis–Gundy inequalities. For more detail we refer the reader to [25] or [30].

Finally, the assertion about the arbitrariness of λ in (4.4) can be easily justified by rescaling arguments when instead of f, g , and w one takes $(c^2 f, c g)(c^2 t, c x)$ and $c^{-1} w_{c^2 t}$ and gets $u(c^2 t, c x)$ instead of $u(t, x)$.

From our explicit formulas and from the particular choice of f and g , it also follows that $u \in C_{loc}([0, \infty), H_p^n)$ for any n (and for any ω). This proves (iii).

From estimates (4.3) and (4.4), we conclude that $u \in \bigcap_{T>0} \mathbb{H}_p^1(T)$. Furthermore, from the pointwise equation (4.9) by the stochastic Fubini theorem it follows easily that u solves (4.2) in the sense of Definition 3.1. Hence $u \in \mathcal{H}_p^1$, which proves a part of assertion (i). The uniqueness in (i) follows from the fact that for $f \equiv 0, g \equiv 0$ we have the heat equation and the uniqueness of its solution in our class of functions is a standard fact (we say more about this in the proof of Lemma 4.9). This completes the proof in the case of step functions f, g .

General case. In the case of general f, g , we observe that the uniqueness in (i) is proved as above. As far as other assertions are concerned we are going to use Theorem 3.10 and Remark 3.6.

If we consider all functions f_0, \tilde{f}^i, g^k as one sequence, then, by Theorem 3.10, we can approximate them by functions $f_{0j}, \tilde{f}_j^i, g_j^k$ of type (4.7). Let u_j be the corresponding solutions of (4.6). By the result for the particular case, $\{u_j\}$ is a Cauchy sequence in \mathcal{H}_p^1 , and, by Theorem 3.7, there is a $u \in \mathcal{H}_p^1$ to which u_j converges in \mathcal{H}_p^1 . Remark 3.6 and the convergence $\|u_{xx} - u_{jxx}\|_{\mathbb{H}_p^{-1}} \rightarrow 0$ prove that $\mathbb{D}u = \Delta u + f$ and $\mathbb{S}u = g$. In particular, this proves assertion (i).

Assertion (ii) follows from the construction of u . From assertion (iii) available in the particular case, we get that u_j is a Cauchy sequence in $L_p(\Omega, C([0, T], L_p))$ for any T . Therefore, it converges in this space to a function \bar{u} . It follows easily that, for any $\phi \in C_0^\infty$,

$$(\bar{u}(t, \cdot), \phi) = \int_0^t \{(\bar{u}(s, \cdot), \Delta \phi) + (f(s, \cdot), \phi)\} ds + \int_0^t (g^k(s, \cdot), \phi) dw_s^k$$

for all t (a.s.). Therefore, $u - \bar{u}$ is a generalized solution of the heat equation with zero initial condition and with bounded L_p -norm (a.s.). This implies that $\|(u - \bar{u})(t, \cdot)\|_p = 0$ for all t (a.s.), so that $u \in C([0, T], L_p)$ for all T (a.s.). Finally, we get (4.4) by Fatou's lemma taking into account that

$$\int_0^T \int_{\mathbb{R}^d} |\nabla(u - u_j)|^p dx dt = \int_0^T \int_{\mathbb{R}^d} |\nabla(1 - \Delta)^{-1/2}(1 - \Delta)^{1/2}(u - u_j)|^p dx dt$$

$$\leq N \int_0^T \int_{\mathbb{R}^d} |(1 - \Delta)^{1/2}(u - u_j)|^p dx dt \rightarrow 0$$

in probability for any T . The theorem is proved.

REMARK 4.3. Although (2.7) holds for all $p \in (1, \infty)$, it follows from [17] that Lemma 4.1 is false if $p < 2$.

4.2. Relation of the Solutions of (4.1) to the Solutions of the Heat Equation. It turns out that the investigation of general equation (4.1) with coefficients independent of x can be quite formally reduced to the particular case of the heat equation. First, we explain how to do this without caring about rigorousness, and then proceed with formal proofs.

The first observation consists of the following. Assume that we have

$$(4.12) \quad du(t, x) = f(t, x) dt + g^k(t, x) dw_t^k,$$

and we define a process x_t and a function v by

$$(4.13) \quad x_t^i = \int_0^t \sigma^{ik}(s) dw_s^k, \quad i = 1, \dots, d, \quad v(t, x) = u(t, x - x_t).$$

Then, by applying formally the Itô–Wentzell formula, we get

$$(4.14) \quad \begin{aligned} dv(t, x) = & [f(t, x - x_t) + \alpha^{ij}(t)v_{x^i x^j}(t, x) - (g_{x^i}(t, x - x_t), \sigma^i(t))_{l_2}] dt \\ & + [g^k(t, x - x_t) - v_{x^k}(t, x)\sigma^{ik}(t)] dw_t^k. \end{aligned}$$

This shows how to introduce the terms $v_{x^i}\sigma^{ik}$ in equation (4.12) and also shows again a kind of necessity for g to have the first derivatives in x .

This device alone is not sufficient, since, if we had $\Delta u + \bar{f}$ instead of f in (4.12), then, in (4.14), we would get the second order differential operator $(\delta^{ij} + \alpha^{ij})\partial^2/\partial x^i \partial x^j$ with coefficients strongly related to the coefficients of $v_{x^i}\sigma^{ik}(t)$. We could get around this difficulty if we manage to start with equations with more general operators L instead Δ . Here the second observation comes that if, instead of (4.12), we consider

$$du(t, x) = (\Delta u + \bar{f}) dt + g^k(t, x) dw_t^k,$$

and take expectations in the counterpart of (4.14) corresponding to this equation, then, assuming that σ is nonrandom, we get indeed an equation for $E v(t, x)$ with operator L different from Δ . By the way, this method of studying parabolic equations with coefficients independent of x was applied in [18] in order to show that “whatever” estimate is true for the heat equation, it is also true for any parabolic equation with coefficients independent of x . Of course, taking expectations “kills” all randomness in the equation, and therefore we use a conditional expectation.

DEFINITION 4.4. Denote by \mathfrak{D} the set of all \mathcal{D} -valued functions u (written as $u(t, x)$ in a common abuse of notation) on $\Omega \times [0, \infty)$ such that, for any $\phi \in C_0^\infty$,

- (i) the function (u, ϕ) is \mathcal{P} -measurable,
- (ii) for any $\omega \in \Omega$ and $T \in (0, \infty)$, we have

$$(4.15) \quad \int_0^T \sup_{x \in \mathbb{R}^d} |(u(t, \cdot), \phi(\cdot - x))|^2 dt < \infty.$$

In the same way, we define $\mathfrak{D}(l_2)$ by considering l_2 -valued linear functionals on C_0^∞ and replacing $|\cdot|$ in (4.15) by $|\cdot|_{l_2}$.

REMARK 4.5. Notice that $(u(t, \cdot), \phi(\cdot - x))$ is continuous in x and Borel in t so that (4.15) makes sense. Also, for $p \geq 2$, $q = p/(p-1)$, and any n ,

$$(4.16) \quad \int_0^T \sup_{x \in \mathbb{R}^d} |(u(t, \cdot), \phi(\cdot - x))|^2 dt \leq \int_0^T \|(u(t, \cdot))\|_{n,p}^2 \|\phi\|_{n,q}^2 dt \\ \leq \|\phi\|_{-n,q}^2 T^{(p-2)/p} \left(\int_0^T \|(u(t, \cdot))\|_{n,p}^p dt \right)^{2/p}.$$

This shows that if $u \in \mathcal{H}_p^n$, then condition (4.15) is satisfied at least for almost all ω . Also, if $u \in \mathcal{H}_p^n$, then (3.2) holds, which shows that $(u(t, \cdot), \phi)$ is indistinguishable from a predictable process. This is true for any $\phi \in C_0^\infty$. From separability of H_q^{-n} , it follows that we can modify u on a set of probability zero and after this we get a function belonging to \mathfrak{D} . This is the sense in which we write

$$(4.17) \quad \mathcal{H}_p^n \subset \mathfrak{D}.$$

DEFINITION 4.6. Let $f, u \in \mathfrak{D}$, $g \in \mathfrak{D}(l_2)$. We say that the equality

$$(4.18) \quad du(t, x) = f(t, x) dt + g(t, x) dw_t, \quad t > 0,$$

holds in the sense of distributions if for any $\phi \in C_0^\infty$, with probability 1 for all $t \geq 0$ we have

$$(4.19) \quad (u(t, \cdot), \phi) = (u(0, \cdot), \phi) + \int_0^t (f(s, \cdot), \phi) ds + \int_0^t (g^k(s, \cdot), \phi) dw_s^k.$$

Observe that, since $\|(g, \phi)(t)\|_{l_2}^2$ is locally summable in t , the last series in (4.19) converges uniformly in t on every finite interval of time in probability.

In this subsection, we always understand equation (4.1) in the sense of distributions. Notice that if $u \in \mathcal{H}_p^n$ and u satisfies (4.19), then, by (4.17), $u \in \mathfrak{D}$ and (4.18) holds in the sense of distributions. An advantage of Definition 4.6 is that one need not check summability of any derivative.

LEMMA 4.7. Let $f, u \in \mathfrak{D}$, $g \in \mathfrak{D}(l_2)$. Assume the definitions in (4.13). Then (4.12) holds (in the sense of distributions) if and only if (4.14) holds (in the sense of distributions).

Proof. First remember that, for a distribution $\alpha(x)$ and $y \in \mathbb{R}^d$, by $\alpha(x-y)$ we mean the distribution defined by $(\alpha, \phi(\cdot + y))$. Also from relations like (cf. (4.16))

$$\begin{aligned} \int_0^T \sup_{y \in \mathbb{R}^d} |(v_{xx}(t, \cdot), \phi(\cdot - y))|^2 dt &= \int_0^T \sup_{y \in \mathbb{R}^d} |(v(t, \cdot), \phi_{xx}(\cdot - y))|^2 dt \\ &= \int_0^T \sup_{y \in \mathbb{R}^d} |(u(t, \cdot), \phi_{xx}(\cdot + x_t - y))|^2 dt = \int_0^T \sup_{y \in \mathbb{R}^d} |(u(t, \cdot), \phi_{xx}(\cdot - y))|^2 dt < \infty, \\ &\quad \int_0^T \sup_{y \in \mathbb{R}^d} |((g_{x^i}(t, \cdot - x_t), \sigma^i(t))_{l_2}, \phi(\cdot - y))|^2 dt \\ &= \int_0^T \sup_{y \in \mathbb{R}^d} |((g_{x^i}(t, \cdot - x_t), \phi(\cdot - y)), \sigma^i(t))_{l_2}|^2 dt \\ &\leq \int_0^T |\sigma^i(t)|_{l_2}^2 dt \int_0^T \sup_{y \in \mathbb{R}^d} |(g_{x^i}(t, \cdot - x_t), \phi(\cdot - y))|_{l_2}^2 dt < \infty \end{aligned}$$

it follows that $v(t, x)$, $f(t, x - x_t)$ and $(g_{x^i}(t, x - x_t), \sigma^i(t))_{l_2}$ belong to \mathfrak{D} and $g(t, x - x_t)$ and $v_{x^i}(t, x)\sigma^i(t)$ belong to $\mathfrak{D}(l_2)$. Furthermore, for any $\phi \in C_0^\infty$, the function $F(t, x) := (u(t, \cdot - x), \phi)$ has a stochastic differential in t for any x and is infinitely differentiable with respect to x . Now our assertion immediately follows from the Itô–Wentzell formula for $F(t, x_t)$. By the way, one can easily memorize (and perhaps prove) this formula by considering the following computations:

$$\begin{aligned} d \int_{\mathbb{R}^d} u(t, x - x_t) \phi(x) dx &= d \int_{\mathbb{R}^d} u(t, x) \phi(x + x_t) dx = \int_{\mathbb{R}^d} \phi(x + x_t) du(t, x) dx \\ &\quad + \int_{\mathbb{R}^d} u(t, x) d\phi(x + x_t) dx + \int_{\mathbb{R}^d} (du(t, x)) d\phi(x + x_t) dx \\ &= \left\{ \int_{\mathbb{R}^d} \phi(x + x_t) f(t, x) dx + \int_{\mathbb{R}^d} u(t, x) \alpha^{ij}(t) \phi_{x^i x^j}(x + x_t) dx \right. \\ &\quad \left. + \int_{\mathbb{R}^d} g^k(t, x) \phi_{x^i}(x + x_t) \sigma^{ik}(t) dx \right\} dt \\ &\quad + \left\{ \int_{\mathbb{R}^d} \phi(x + x_t) g^k(t, x) dx + \int_{\mathbb{R}^d} u(t, x) \phi_{x^i}(x + x_t) \sigma^{ik}(t) dx \right\} dw_t^k \\ &= \int_{\mathbb{R}^d} \{f(t, x - x_t) + \alpha^{ij}(t) u_{x^i x^j}(t, x - x_t) - g_{x^i}^k(t, x - x_t) \sigma^{ik}(t)\} \phi(x) dx dt \\ &\quad + \int_{\mathbb{R}^d} \{g^k(t, x - x_t) - u_{x^i}(t, x - x_t) \sigma^{ik}(t)\} \phi(x) dx dw_t^k. \end{aligned}$$

The lemma is proved.

REMARK 4.8. If, instead of (4.12), u satisfies the equation

$$du(t, x) = (a^{ij}(t) u_{x^i x^j}(t, x) + h(t, x)) dt + \sigma^{ik}(t) u_{x^i}(t, x) dw_t^k,$$

then (4.14) takes the form

$$(4.20) \quad \frac{\partial}{\partial t} v(t, x) = ((a^{ij}(t) - \alpha^{ij}(t)) v_{x^i x^j}(t, x) + h(t, x - x_t)) \quad t > 0,$$

and can be considered on each ω separately. Observe that if $a(t) < \alpha(t)$, then the initial-value problem $v(0) = v_0$ for equation (4.20) is *ill posed*.

This shows that operators appearing in the stochastic term should be subordinated in a certain sense to the operator in the deterministic part of equation. This is needed in order to construct an L_p -theory. On the other hand, take $d = 1$ and a one-dimensional Wiener process w_t . Consider the following equation

$$du(t, x) = i u'_x(t, x) dw_t.$$

Surprisingly enough and somewhat in spite of what is said above, this equation has a very nice solution for each initial data $u_0 \in L_2$. One gets the solution after passing to Fourier transforms. It turns out that $\tilde{u}(t, \xi) = u_0(\xi) \exp\{\xi w_t - (1/2)|\xi|^2 t\}$. The function $\tilde{u}(t, \xi)$ decays very fast when $|\xi| \rightarrow \infty$, which shows that $u(t, x)$ is infinitely differentiable in x . Also notice that, taking expectations, we see that $E u(t, x) = u_0(x)$ if u_0 is nonrandom, and in this case we get a representation of any L_2 function as an integral over Ω of functions $u(\omega, 1, x)$ which are infinitely differentiable in x . However, the major drawback of such equations is that $E|u(t, 0)|^p = \infty$ for any $p > 1$ if, for example, $\tilde{u}_0(\xi) \geq \exp\{-\lambda|\xi|\}$, where λ is a constant.

LEMMA 4.9. *Let $f \in \mathfrak{D}$, $g \in \mathfrak{D}(l_2)$, u_0 be a \mathcal{D} -valued function on Ω . Then the following assertions hold:*

(i) *In \mathfrak{D} there can exist only one (up to evanescence) solution of equation (4.1) with the initial condition $u(0, \cdot) = u_0$.*

(ii) *Let $\mathcal{F}_t = \mathcal{W}_t \vee \mathcal{B}_t$ for $t \geq 0$, and assume that σ -fields \mathcal{W}_t and \mathcal{B}_t form independent increasing filtrations. Let W and B be sets such that $W \cup B = \{1, 2, \dots\}$. Assume that (w_t^k, \mathcal{W}_t) and (w_t^r, \mathcal{B}_t) are Wiener processes for $k \in W$, $r \in B$. Let $u \in \mathfrak{D}$ satisfy equation (4.1) (in the sense of distributions), and let a, f, σ, g be \mathcal{W}_t -adapted. Finally, assume that there exists an $n \in (-\infty, \infty)$ such that $f \in \mathbb{H}_2^n(T)$, $g \in \mathbb{H}_2^n(T, l_2)$ for any $T \in (0, \infty)$ and $u(0, \cdot)$ is W_0 -measurable and*

$$E\|u(0, \cdot)\|_{n,2}^2 < \infty.$$

Then in \mathfrak{D} there exists a unique solution \tilde{u} of the equation

$$(4.21) \quad d\tilde{u} = (a^{ij}\tilde{u}_{x^i x^j} + f) dt + \sum_{k \in W} (\sigma^{ik}\tilde{u}_{x^i} + g^k) dw_t^k, \quad t > 0.$$

In addition, for any $\phi \in C_0^\infty$ and $t \geq 0$,

$$(4.22) \quad (\tilde{u}(t, \cdot), \phi) = E\{(u(t, \cdot), \phi) | \mathcal{W}_t\} \quad (\text{a.s.}).$$

Proof. (i) As always, we can take $f \equiv 0, g \equiv 0$, and $u_0 \equiv 0$, and, by Lemma 4.7, it suffices to consider only the case $\sigma \equiv 0$. In this case, for any given $\phi \in C_0^\infty$ we have

$$(u(t, \cdot), \phi) = \int_0^t (u(s, \cdot), L(s)\phi) ds, \quad t \geq 0,$$

almost surely. Putting here $\phi(\cdot - x)$ instead of ϕ and observing that both sides are continuous and bounded in (t, x) on $[0, T] \times \mathbb{R}^d$ for any $T < \infty$ (cf. (4.15)), we get that the function $F(t, x) := (u(t, \cdot), \phi(\cdot - x))$ is bounded in (t, x) on $[0, T] \times \mathbb{R}^d$ for any $T < \infty$, infinitely differentiable in x , and satisfies the equation

$$F(t, x) = \int_0^t L(s)F(s, x) ds \quad \forall t, x$$

(a.s.). From the theory of parabolic equations, it follows that $F(t, x) = 0$ for all t, x (a.s.). This means that $(u(t, \cdot), \phi) = 0$ for all t (a.s.). Now take ϕ with unit integral. Then for any x and n with probability 1, $(u(t, \cdot), n^d\phi(n(\cdot - x))) = 0$ for all t . Since this function is continuous in x , we have $(u(t, \cdot), n^d\phi(n(\cdot - x))) = 0$ for all t and x with probability 1. Finally, $(u(t, \cdot), n^d\phi(n(\cdot - x))) \rightarrow u(t, x)$ as $n \rightarrow \infty$ for all (ω, t, x) in the sense of distributions, which implies that, with probability 1, we have $u(t, \cdot) = 0$ for all t as stated.

(ii) First, notice that, according to [30], equation (4.1) has a unique solution v in the space $\mathbb{H}_2^{n+1}(T)$ for any T . The definition of solutions in $\mathbb{H}_2^{n+1}(T)$ from [30] is slightly different, but v is continuous (a.s.) as an H_2^n -valued process and

$$(4.23) \quad E \sup_{t \leq T} \|v(t, \cdot)\|_{n,2}^2 < \infty \quad \forall T < \infty,$$

so that v is a \mathfrak{D} -solution of (4.1). It follows from (i) that our function u coincides with v and therefore belongs to $\mathbb{H}_2^{n+1}(T)$ for any T , and (4.23) holds for u . Furthermore, with probability 1 for all t at once,

$$u(t) = u(0) + \int_0^t [a^{ij}(s)u_{x^i x^j}(s) + f(s)] ds + \int_0^t [\sigma^{ik}(s)u_{x^i}(s) + g(s)] dw_s^k,$$

where all integrals are taken in the sense of the Hilbert space H_2^{n-1} (see [30]). By Theorem 1.4.7 of [30], or rather by its Hilbert-space counterpart, there exists an H_2^{n+1} -valued, \mathcal{W}_t -predictable function $\bar{u}(t)$ such that, for almost any t , we have

$$\bar{u}(t) = E\{u(t)|\mathcal{W}_t\}, \quad \bar{u}_x(t) = E\{u_x(t)|\mathcal{W}_t\}, \quad \bar{u}_{xx}(t) = E\{u_{xx}(t)|\mathcal{W}_t\} \quad (\text{a.s.})$$

(conditional expectations of Hilbert-space valued random elements) and

(4.24)

$$\bar{u}(t) = u(0) + \int_0^t [a^{ij}(s)\bar{u}_{x^i x^j}(s) + f(s)] ds + \sum_{k \in W} \int_0^t [\sigma^{ik}(s)\bar{u}_{x^i}(s) + g^k(s)] dw_s^k$$

for almost all t and ω . The right-hand side here is a continuous H_2^{n-1} -valued process which we denote by \tilde{u} and we show that \tilde{u} is the function we need.

By definition and by the equality $\bar{u} = \tilde{u}$ (a.e.), \tilde{u} satisfies (4.24) for all t with probability 1 and also is a continuous process in H_2^{n-1} . This implies that $\tilde{u} \in \mathfrak{D}$ and \tilde{u} is a solution of (4.21). To prove (4.22) for any t , it remains only to observe that, again by Theorem 1.4.7 of [30], the conditional expectation $E\{u(t)|\mathcal{W}_t\}$ for any t is equal to the right-hand side of (4.24) almost surely. The lemma is proved.

4.3. General Equation (4.1) with Coefficients Independent of x .

THEOREM 4.10. *Take $n \in \mathbb{R}$ and let $f \in \mathbb{H}_p^{n-1}$, $g \in \mathbb{H}_p^n(l_2)$. Then*

- (i) *equation (4.1) with zero initial condition has a unique solution $u \in \mathcal{H}_p^{n+1}$;*
- (ii) *for this solution, we have*

$$(4.25) \quad \|u_{xx}\|_{\mathbb{H}_p^{n-1}} \leq N(\|f\|_{\mathbb{H}_p^{n-1}} + \|g\|_{\mathbb{H}_p^n(l_2)}), \quad \|u\|_{\mathcal{H}_p^{n+1}} \leq N\|(f, g)\|_{\mathcal{F}_p^{n-1}},$$

where $N = N(d, p, \delta, K)$;

- (iii) *we have $u \in C_{loc}([0, \infty), H_p^n)$ almost surely and for any $\lambda, T > 0$,*

$$(4.26) \quad E \sup_{t \leq T} (e^{-p\lambda t} \|u(t, \cdot)\|_{n,p}^p) \leq N(\|e^{-\lambda t} f\|_{\mathbb{H}_p^{n-1}(T)}^p + \|e^{-\lambda t} g\|_{\mathbb{H}_p^n(T, l_2)}^p),$$

where $N = N(d, p, \delta, K, \lambda)$.

Proof. Since one can apply the operator $(I - \Delta)^{n/2}$ to both sides of (4.1), it suffices to prove the theorem only for $n = 0$. Furthermore, as we have already noticed in (4.17), any function $u \in \mathcal{H}_p^1$ also belongs to \mathfrak{D} . This and Lemma 4.9 prove the uniqueness in (i). Also, the fact that our norms are translation invariant, combined with Lemma 4.7, shows that, to prove the existence in (i) and all other assertions of the theorem, we only need to consider the case $\sigma \equiv 0$. As in the proof of Theorem 4.2, we can assume that f and g are as in (4.7). In this case, as we know from [25] or [30], equation (4.1) has a unique \mathfrak{D} -solution u that belongs to $C_b([0, T] \times \mathbb{R}^d)$ and $C([0, T], L_2)$ almost surely for any $T < \infty$. It follows that $u \in C([0, T], L_p)$ almost surely for any $T < \infty$. Estimate (4.26) also follows from [25] or [30] as in the proof of Theorem 4.2. It remains only to prove that $u \in \mathcal{H}_p^1$ and that (4.25) holds. Since u is a \mathfrak{D} -solution, to prove that $u \in \mathcal{H}_p^1$, it suffices to prove that $u \in \mathbb{H}_p^1(T)$ for any $T < \infty$.

Since the matrix a is uniformly non-degenerate, by making a nonrandom time change, we can reduce the general case to the case $a \geq I$. In this case, define the matrix-valued function $\bar{\sigma}(t) = \bar{\sigma}^*(t) \geq 0$ as a solution of the equation $\bar{\sigma}^2(t) + 2I = 2a(t)$. Furthermore, without loss of generality, we assume that on our probability space we are also given a d -dimensional Wiener process B_t independent of \mathcal{F}_t .

Now, consider the equation

$$(4.27) \quad dv(t, x) = [\Delta v(t, x) + f(t, x - \int_0^t \bar{\sigma}(s) dB_s)] dt + g^k(t, x - \int_0^t \bar{\sigma}(s) dB_s) dw_t^k$$

with zero initial condition. Replace the predictable σ -field \mathcal{P} with predictable σ -field generated by $\mathcal{F}_t \vee \sigma(B_s; s \leq t)$. Then the spaces \mathcal{H}_p^n become larger. By Theorem 4.2 there is a solution v of (4.27) possessing properties (i) through (iii) listed in Theorem 4.2 (with new \mathcal{H}_p^1). Use again that, after changing, if necessary, v on a set of probability zero, the function v becomes a \mathfrak{D} -solution of (4.27). Then, by Lemma 4.7, the function $z(t, x) := v(t, x + \int_0^t \bar{\sigma}(s) dB_s)$ is a \mathfrak{D} -solution of

$$dz(t, x) = (a^{ij}(t) z_{x^i x^j}(t, x) + f(t, x)) dt + g^k(t, x) dw_t^k + z_{x^i}(t, x) \bar{\sigma}^{ij}(t) dB_t^j,$$

and by Lemma 4.9 there is a solution $\tilde{u} \in \mathfrak{D}$ of

$$d\tilde{u}(t, x) = (a^{ij}(t) \tilde{u}_{x^i x^j}(t, x) + f(t, x)) dt + g^k(t, x) dw_t^k,$$

which is (4.1) in our case. In addition, for any $\phi \in C_0^\infty$ and $t \geq 0$,

$$(\tilde{u}(t, \cdot), \phi) = E\{(z(t, \cdot), \phi) | \mathcal{F}_t\} = E\{(v(t, \cdot + \int_0^t \bar{\sigma}(s) dB_s), \phi) | \mathcal{F}_t\} \quad (\text{a.s.}).$$

In particular, it follows from this equality that \tilde{u} is a \mathfrak{D} -solution with respect to the initial predictable σ -field \mathcal{P} , and from uniqueness we get $\tilde{u} = u$. Therefore,

$$(4.28) \quad (u(t, \cdot), \phi) = E\{(v(t, \cdot + \int_0^t \bar{\sigma}(s) dB_s), \phi) | \mathcal{F}_t\} \quad (\text{a.s.}).$$

It follows that

$$(4.29) \quad |(u(t, \cdot), \phi)|^p \leq E\{\|v(t, \cdot)\|_{1,p}^p | \mathcal{F}_t\} \|\phi\|_{-1,q}^p$$

(a.s.) for any $\phi \in C_0^\infty$ and $t \geq 0$, where $q = p/(p-1)$.

Next, take a countable family $\Phi \subset C_0^\infty$ dense in C_0^∞ . Observe that, given a distribution ψ , we have $\psi \in H_p^1$ if and only if for any $\phi \in \Phi$ we have $|(\psi, \phi)| \leq N\|\phi\|_{-1,q}$ with a constant N independent of ϕ . Indeed, in this case there exists a bounded linear functional ℓ on H_q^{-1} such that $\ell(\phi) = (\psi, \phi)$ for any $\phi \in \Phi$. Since $\ell(\phi) = ((1-\Delta)^{-1/2}h, \phi)$ with an $h \in L_p$ and Φ is dense in C_0^∞ , we have $\psi = (1-\Delta)^{-1/2}h \in H_p^1$ indeed. In particular, this fact implies that the set $\{(\omega, t) : w(\omega, t, \cdot) \in H_p^1\}$ is measurable (even predictable) for any $w \in \mathfrak{D}$, say for $w = u$.

We also know that $v \in \mathcal{H}_p^1$, which implies that $E\|v(t, \cdot)\|_{1,p}^p < \infty$ for almost all t . Fix such a t . Then there exists a set Ω' of probability 1 such that $E\{\|v(t, \cdot)\|_{1,p}^p | \mathcal{F}_t\} < \infty$ on Ω' and (4.29) holds for all $\omega \in \Omega'$ and $\phi \in \Phi$. Hence, $u(t, \cdot) \in H_p^1$ for the chosen t and all $\omega \in \Omega'$. In particular, $u(t, \cdot) \in H_p^1$ for almost all (ω, t) , and from (4.29) it follows that

$$\|u(t, \cdot)\|_{1,p}^p \leq E\{\|v(t, \cdot)\|_{1,p}^p | \mathcal{F}_t\} \quad (\text{a.s.}), \quad \|u\|_{\mathbb{H}_p^1(T)} \leq \|v\|_{\mathbb{H}_p^1(T)} < \infty.$$

Thus $u \in \mathbb{H}_p^1(T)$ for any $T < \infty$ indeed and $u \in \mathcal{H}_p^1$.

Similarly, from the equality

$$(u_{xx}(t, \cdot), \phi) = E\{(v_{xx}(t, \cdot + \int_0^t \bar{\sigma}(s) dB_s), \phi) | \mathcal{F}_t\} \quad (\text{a.s.}),$$

one gets

$$\|u_{xx}(t, \cdot)\|_{-1,p}^p \leq E\{\|v_{xx}(t, \cdot)\|_{-1,p}^p | \mathcal{F}_t\} \quad (\text{a.s.}).$$

This and the properties of v immediately yield (4.25). The theorem is proved.

REMARK 4.11. By using the self-similarity of equation (4.1), it is possible to obtain further estimates from estimates like (4.26). For instance, remembering that $H_p^1 = W_p^1$, one sees that, for $n = 1$ and $\lambda = 1/p$, estimate (4.26) implies that

$$\begin{aligned} E \sup_{t \leq T} \{ \|u_x(t, \cdot)\|_p^p + \|u(t, \cdot)\|_p^p \} \\ \leq N(d, p, \delta, K) e^T (\|f\|_{L_p(T)}^p + \|g_x\|_{L_p(T, l_2)}^p + \|g\|_{L_p(T, l_2)}^p). \end{aligned}$$

Let us take a constant $c > 0$ and consider $(c^2 f, cg)(c^2 t, cx)$, $c^{-1} w_{c^2 t}$, and $u(c^2 t, cx)$ instead of f, g, w , and u . Then, from the last estimate, we get

$$\begin{aligned} E \sup_{t \leq T} \{ c^{p-d} \|u_x(c^2 t, \cdot)\|_p^p + c^{-d} \|u(c^2 t, \cdot)\|_p^p \} \\ \leq N e^T (c^{2p-(d+2)} \|f\|_{L_p(c^2 T)}^p + c^{2p-(d+2)} \|g_x\|_{L_p(c^2 T, l_2)}^p + c^{p-(d+2)} \|g\|_{L_p(c^2 T, l_2)}^p), \end{aligned}$$

the constant N being the same as above. It follows that, for $c, T \geq 0$,

$$\begin{aligned} E \sup_{t \leq T} \{ \|u_x(t, \cdot)\|_p^p + c^{-p} \|u(t, \cdot)\|_p^p \} \\ \leq N e^{T/c^2} c^{p-2} (\|f\|_{L_p(T)}^p + \|g_x\|_{L_p(T, l_2)}^p + c^{-p} \|g\|_{L_p(T, l_2)}^p). \end{aligned}$$

Upon setting $c^2 = T$ and considering $(1 - \Delta)^{(n-1)/2} u$ instead of u , we conclude that, under the conditions of Theorem 4.2, for any $T > 0$,

$$\begin{aligned} E \{ \sup_{t \leq T} \|u_x(t, \cdot)\|_{n-1,p}^p + T^{-p/2} \sup_{t \leq T} \|u(s, \cdot)\|_{n-1,p}^p \} \leq \\ N(d, p, \delta, K) T^{(p-2)/2} (\|f\|_{H_p^{n-1}(T)}^p + \|g_x\|_{H_p^{n-1}(T, l_2)}^p + T^{-p/2} \|g\|_{H_p^{n-1}(T, l_2)}^p). \end{aligned}$$

We will later prove a much deeper estimate than (4.26).

5. Equations with Variable Coefficients

Here we are going to state main results concerning L_p -theory of SPDEs in the whole space, the most important of which, Theorem 5.1, will be proved in Sec. 6. Take a stopping time $\tau \leq T$ with T being a finite constant. Fix $n \in (-\infty, \infty)$ and fix a number $\gamma \in [0, 1)$ such that $\gamma = 0$ if $n = 0, \pm 1, \pm 2, \dots$; otherwise $\gamma > 0$ and is such that $|n| + \gamma$ is not an integer. Define

$$\begin{aligned} B^{|n|+\gamma} &= \begin{cases} B(\mathbb{R}^d) & \text{if } n = 0, \\ C^{|n|-1,1}(\mathbb{R}^d) & \text{if } n = \pm 1, \pm 2, \dots, \\ C^{|n|+\gamma}(\mathbb{R}^d) & \text{otherwise;} \end{cases} \\ B^{|n|+\gamma}(l_2) &= \begin{cases} B(\mathbb{R}^d, l_2) & \text{if } n = 0, \\ C^{|n|-1,1}(\mathbb{R}^d, l_2) & \text{if } n = \pm 1, \pm 2, \dots, \\ C^{|n|+\gamma}(\mathbb{R}^d, l_2) & \text{otherwise,} \end{cases} \end{aligned}$$

where $B(\mathbb{R}^d)$ is the Banach space of bounded functions on \mathbb{R}^d , $C^{|n|-1,1}(\mathbb{R}^d)$ is the Banach space of $|n|-1$ times continuously differentiable functions whose derivatives of $(|n|-1)$ st order satisfy the Lipschitz condition on \mathbb{R}^d , $C^{|n|+\gamma}(\mathbb{R}^d)$ is the usual Hölder space, and l_2 means that instead of real-valued functions we consider l_2 -valued ones.

Consider the following nonlinear equation on $[0, \tau]$:

$$(5.1) \quad \begin{aligned} du(t, x) = & [a^{ij}(t, x)u_{x^i x^j}(t, x) + f(u, t, x)] dt \\ & + [\sigma^{ik}(t, x)u_{x^i}(t, x) + g^k(u, t, x)] dw_t^k, \end{aligned}$$

where a^{ij} and f are real-valued, and σ^i and g are l_2 -valued functions defined for $\omega \in \Omega$, $t \geq 0$, $x \in \mathbb{R}^d$, $u \in H_p^{n+2}$, $i, j = 1, \dots, d$. We consider this equation in the sense of Definition 3.5 (where we take $n+2$ instead of n).

We make the following assumptions, where, as in Section 4, we define

$$(5.2) \quad \alpha^{ij}(t, x) = \frac{1}{2}(\sigma^i(t, x), \sigma^j(t, x))_{l_2}.$$

ASSUMPTION 5.1 (coercivity). For any $\omega \in \Omega$, $t \geq 0$, $x, \lambda \in \mathbb{R}^d$, we have

$$K|\lambda|^2 \geq [a^{ij}(t, x) - \alpha^{ij}(t, x)]\lambda^i \lambda^j \geq \delta|\lambda|^2,$$

where K, δ are fixed strictly positive constants.

ASSUMPTION 5.2 (uniform continuity of a and σ). For any $\varepsilon > 0$, i, j , there exists a $\kappa_\varepsilon > 0$ such that

$$(5.3) \quad |a^{ij}(t, x) - a^{ij}(t, y)| + |\sigma^i(t, x) - \sigma^i(t, y)|_{l_2}^2 \leq \varepsilon$$

whenever $|x - y| \leq \kappa_\varepsilon$, $t \geq 0$, $\omega \in \Omega$.

This assumption is actually used only if $n = 0$, and even then we need a stronger condition on σ . For other values of n we impose stronger conditions on a and σ .

ASSUMPTION 5.3. For any i, j, k , the functions $a^{ij}(t, x)$ and $\sigma^{ik}(t, x)$ are real-valued $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$ -measurable functions, and for any $\omega \in \Omega$ and $t \geq 0$, we have

$$a^{ij}(t, \cdot) \in B^{|n|+\gamma}, \quad \sigma^i(t, \cdot) \in B^{|n+1|+\gamma}(l_2).$$

ASSUMPTION 5.4. For any $u \in H_p^{n+2}$, the functions $f(u, t, x)$ and $g(u, t, x)$ are predictable as functions taking values in H_p^n and $H_p^{n+1}(\mathbb{R}^d, l_2)$, respectively.

ASSUMPTION 5.5. For any $t \geq 0, \omega, i, j$,

$$\|a^{ij}(t, \cdot)\|_{B^{|n|+\gamma}} + \|\sigma^i(t, \cdot)\|_{B^{|n+1|+\gamma}(l_2)} \leq K, \quad (f(0, \cdot, \cdot), g(0, \cdot, \cdot)) \in \mathcal{F}_p^n(\tau).$$

ASSUMPTION 5.6. The functions f, g are continuous in u . Moreover, for any $\varepsilon > 0$, there exists a constant K_ε such that, for any $u, v \in H_p^{n+2}$, t, ω , we have

$$(5.4) \quad \begin{aligned} \|f(u, t, \cdot) - f(v, t, \cdot)\|_{n,p} + \|g(u, t, \cdot) - g(v, t, \cdot)\|_{n+1,p} \\ \leq \varepsilon \|u - v\|_{n+2,p} + K_\varepsilon \|u - v\|_{n,p}. \end{aligned}$$

THEOREM 5.1. Let Assumptions 5.1–5.6 be satisfied and let

$$u_0 \in L_p(\Omega, \mathcal{F}_0, H_p^{n+2-2/p}).$$

Then the Cauchy problem for equation (5.1) on $[0, \tau]$ with the initial condition $u(0, \cdot) = u_0$ has a unique solution $u \in \mathcal{H}_p^{n+2}(\tau)$. For this solution, we have

$$\|u\|_{\mathcal{H}_p^{n+2}(\tau)} \leq N \{ \|f(0, \cdot, \cdot)\|_{\mathcal{H}_p^n(\tau)} + \|g(0, \cdot, \cdot)\|_{\mathcal{H}_p^{n+1}(\tau, l_2)} + (E \|u_0\|_{n+2-2/p,p}^p)^{1/p} \},$$

where the constant N depends only on $d, n, \gamma, p, \delta, K, T$, and the functions κ_ε and K_ε .

To discuss the theorem, we need the following lemma.

LEMMA 5.2. Let $\zeta \in C_0^\infty$ be a nonnegative function such that $\int \zeta(x) dx = 1$ and define $\zeta_k(x) = k^d \zeta(kx)$, $k = 1, 2, 3, \dots$. We assert that, for any $u \in H_p^n$, we have the following:

(i) $\|au\|_{n,p} \leq N\bar{a}\|u\|_{n,p}$, where $\bar{a} = \|a\|_{B^{|\alpha|+\gamma}}$ and the constant N depends only on d, p, n , and γ ;

(ii) $\|u * \zeta_k\|_{n,p} \leq \|u\|_{n,p}$, $\|u - u * \zeta_k\|_{n,p} \rightarrow 0$.

The same assertions hold true for Banach-space valued a with natural definition of \bar{a} .

Proof. If $n \neq 0, \pm 1, \pm 2, \dots$ (and $\gamma > 0$), then one gets (i) by Corollary 2.8.2 (ii) of [33]. If n is a nonnegative integer, then (i) follows from the Leibnitz rule (and the fact that $H_p^n = W_p^n$). For negative integers (and generally for negative n) (i) follows easily by duality, that is, by using the fact that if $u = (1 - \Delta)^{-n/2}f$, then

$$(au, \phi) = (f, (1 - \Delta)^{-n/2}(a\phi)) \leq \|f\|_p \|a\phi\|_{-n,q} \quad q = p/(p-1).$$

As for (ii), the first inequality follows from Minkowski's inequality and the second one is derived as usual owing to denseness of C_0^∞ in H_p^n . The lemma is proved.

REMARK 5.3. As we have said above, by solution to the Cauchy problem for equation (5.1) on $[0, \tau]$ with the given initial condition u_0 , we understand a function $u \in \mathcal{H}_p^{n+2}(\tau)$ such that for any test function $\phi \in C_0^\infty$, one has almost surely

$$\begin{aligned} (u(t, \cdot), \phi) &= (u_0, \phi) + \int_0^t (a^{ij}(s, \cdot)u_{x^i x^j}(s, \cdot) + f(u, s, \cdot), \phi) ds \\ &\quad + \int_0^t (\sigma^{ik}(s, \cdot)u_{x^i}(s, \cdot) + g^k(u, s, \cdot), \phi) dw_s^k, \quad \forall t \in [0, \tau]. \end{aligned}$$

It is important that, under our assumptions, the equation makes sense for $u \in \mathcal{H}_p^{n+2}(\tau)$, since by Lemma 5.2 we have $a^{ij}u_{x^i x^j} \in H_p^n$, $\sigma^i u_{x^i} \in H_p^{n+1}(\mathbb{R}^d, l_2)$ whenever $u \in \mathcal{H}_p^{n+2}$.

REMARK 5.4. Two main ideas in the proof of this theorem are quite standard. The first one, reduction to equations with constant coefficients, will be seen very clearly. The second one, which is somewhat hidden, consists of introducing the new unknown function $v = (1 - \Delta)^{n/2}u$, which reduces the case of general n to the case $n = 0$. The function v satisfies

$$\begin{aligned} dv &= \left\{ (1 - \Delta)^{n/2}(a^{ij}(1 - \Delta)^{-n/2}v_{x^i x^j}) + (1 - \Delta)^{n/2}f \right\} dt \\ &\quad + \left\{ (1 - \Delta)^{n/2}(\sigma^{ik}(1 - \Delta)^{-n/2}v_{x^i}) + (1 - \Delta)^{n/2}g^k \right\} dw_t^k. \end{aligned}$$

This is a pseudo-differential equation, and we note that more general pseudo-differential equations can be considered too. Also, this equations shows a need to have smoothness assumptions on a, σ in x if we are interested in $n \neq 0$ both positive or negative.

REMARK 5.5. By Theorem 14.2 of [14], for any $u \in H_p^{n+2}$ and $m \in [n, n+2]$, we have

$$\|u\|_{m,p} \leq N\|u\|_{n+2,p}^\theta \|u\|_{n,p}^{1-\theta} \leq N\theta\varepsilon\|u\|_{n+2,p} + N(1-\theta)\varepsilon^{-\theta/(1-\theta)}\|u\|_{n,p},$$

where $\theta = (m-n)/2$ and N depends only on d, n, m , and p . This shows that the right-hand side in (5.4) can be replaced by $\|u - v\|_{n+\varepsilon+1,p}$ once $|\varepsilon| < 1$. As an

example, one can take $f = f_0(x) \sup_x |u_x|$ if $(n+1)p > d$ and $f_0 \in H_p^n$. Indeed, by Sobolev's embedding theorems, $H_p^{n+1+\varepsilon} \subset C^1$ if $(n+\varepsilon)p > d$. Therefore,

$$\|f(u, t, \cdot) - f(v, t, \cdot)\|_{n,p} \leq \|f_0\|_{n,p} \sup_x |(u - v)_x| \leq N \|u - v\|_{n+1+\varepsilon,p}.$$

REMARK 5.6. A typical application of Theorem 5.1 occurs when $f(u, t, x) = b^i(t, x)u_{x^i} + c(t, x)u + f(t, x)$ and $g(u, t, x) = \nu(t, x)u + g(t, x)$, so that (5.1) becomes

$$(5.5) \quad du = (a^{ij}u_{x^i x^j} + b^i u_{x^i} + cu + f) dt + (\sigma^{ik}u_{x^i} + \nu^k u + g^k) dw_t^k.$$

To describe the appropriate assumptions, we take $\varepsilon \in (0, 1)$ and denote

$$\begin{aligned} n_b &= n + \gamma & \text{if } n \geq 0, & n_b = 0 & \text{if } n \in (-1, 0], \\ n_\nu &= n + 1 + \gamma & \text{if } n \geq -1, & n_\nu = 0 & \text{if } n \in (-2, -1], \\ n_c &= n + \gamma & \text{if } n \geq 0, & n_c = 0 & \text{if } n \in (-2, 0], \\ n_b &= -n - 1 + \varepsilon & \text{if } n \leq -1, \\ n_\nu &= -n - 2 + \varepsilon & \text{if } n \leq -2, \\ n_c &= -n - 2 + \varepsilon & \text{if } n \leq -2. \end{aligned}$$

Assume that b, c , and ν are appropriately measurable and

$$\begin{aligned} b^i(t, \cdot) &\in B^{n_b}, \quad c(t, \cdot) \in B^{n_c}, \quad \nu(t, \cdot) \in B^{n_\nu}(\mathbb{R}^d, l_2), \\ f(t, \cdot) &\in H_p^n, \quad g(t, \cdot) \in H_p^{n+1}(\mathbb{R}^d, l_2). \end{aligned}$$

$$\|b^i(t, \cdot)\|_{B^{n_b}} + \|c(t, \cdot)\|_{B^{n_c}} + \|\nu(t, \cdot)\|_{B^{n_\nu}(\mathbb{R}^d, l_2)} \leq K, \quad (f(\cdot, \cdot), g(\cdot, \cdot)) \in \mathcal{F}_p^n(\tau).$$

It turns out then that the assumptions of Theorem 5.1 about $f(u, t, x)$ and $g(u, t, x)$ are satisfied. To show this, it suffices to apply Remark 5.5 and to notice that, for instance, if $n \geq -1$, then $\|\nu u\|_{n+1,p} \leq N \|u\|_{n+1,p}$ by Lemma 5.2; if $n \in (-2, -1]$, then obviously $\|\nu u\|_{n+1,p} \leq \|\nu u\|_p \leq N \|u\|_p = N \|u\|_{n+1+(-n-1),p}$, and $-n-1 \in [0, 1]$; if $n \leq -2$, then Lemma 5.2 yields $\|\nu u\|_{n+1,p} \leq \|\nu u\|_{n+2-\varepsilon_1,p} \leq N \|u\|_{n+2-\varepsilon_1,p}$, where $\varepsilon_1 \in (0, \varepsilon)$. The terms $\|b^i u_{x^i}\|_{n,p}, \|cu\|_{n,p}$ are considered similarly.

Actually, the above conditions on b, c , and ν can be considerably relaxed if one makes use of deeper theorems about multipliers from [33].

Conditions (5.6) and (5.7) of the following theorem are discussed in Remarks 5.9 and 5.8.

THEOREM 5.7. Assume that for $m = 1, 2, 3, \dots$, we are given $a_m^{ij}, \sigma_m^i, f_m, g_m$, and u_{0m} having the same sense as in Theorem 5.1 and verifying the same assumptions as a^{ij}, σ^i, f, g , and u_0 with the same constants $\delta, K, \kappa_\varepsilon$, and K_ε . Let $\zeta(x)$ be a real function of class C_0^∞ such that $\zeta(x) = 1$ if $|x| \leq 1$ and $\zeta(x) = 0$ if $|x| \geq 2$. Define $\zeta_r(x) = \zeta(x/r)$ and assume that, for any $r = 1, 2, 3, \dots$, $i, j = 1, \dots, d$, $t \geq 0$, and $\omega \in \Omega$,

$$(5.6) \quad \|\zeta_r \{a^{ij}(t, \cdot) - a_m^{ij}(t, \cdot)\}\|_{n,p} + \|\zeta_r \{\sigma^i(t, \cdot) - \sigma_m^i(t, \cdot)\}\|_{n+1,p} \rightarrow 0$$

as $m \rightarrow \infty$. Finally, let $E \|u_{0m} - u_0\|_{n+2-2/p,p}^p \rightarrow 0$ and

$$(5.7) \quad \|f(u, \cdot, \cdot) - f_m(u, \cdot, \cdot)\|_{\mathbb{H}_p^n(\tau)} \rightarrow 0, \quad \|g(u, \cdot, \cdot) - g_m(u, \cdot, \cdot)\|_{\mathbb{H}_p^{n+1}(\tau, l_2)} \rightarrow 0$$

whenever $u \in \mathcal{H}_p^{n+2}(\tau)$. Take the function u from Theorem 5.1 and for any m define $u_m \in \mathcal{H}_p^{n+2}(\tau)$ as the (unique) solution of the Cauchy problem for the equation

$$(5.8) \quad \begin{aligned} du_m(t, x) &= [a_m^{ij}(t, x)u_{mx^i x^j}(t, x) + f_m(u_m, t, x)] dt \\ &\quad + [\sigma_m^{ik}(t, x)u_{mx^i}(t, x) + g_m^k(u_m, t, x)] dw_t^k \end{aligned}$$

on $[0, \tau]$ with initial condition $u_m(0, \cdot) = u_{0m}$. Then $\|u - u_m\|_{\mathcal{H}_p^{n+2}(\tau)} \rightarrow 0$ as $m \rightarrow \infty$.

Proof. For $v_m = u - u_m$ we have

$$dv_m(t) = [a_m^{ij} v_{mx^i x^j} + F_m(v_m)] dt + [\sigma_m^{ik} v_{mx^i} + G_m^k(v_m)] dw_t^k,$$

where

$$\begin{aligned} F_m(v) &= (a^{ij} - a_m^{ij}) u_{x^i x^j} + f(u) - f_m(u - v), \\ G_m^k(v) &= (\sigma^{ik} - \sigma_m^{ik}) u_{x^i} + g^k(u) - g_m^k(u - v). \end{aligned}$$

Hence, from our assumptions and by Theorem 5.1, we obtain

$$\|u - u_m\|_{\mathcal{H}_p^{n+2}(\tau)} \leq NI_m,$$

where N is independent of m and

$$\begin{aligned} I_m &= \|(a^{ij} - a_m^{ij}) u_{x^i x^j}\|_{\mathbb{H}_p^n(\tau)} + \|f(u) - f_m(u)\|_{\mathbb{H}_p^n(\tau)} \\ &+ \|(\sigma^i - \sigma_m^i) u_{x^i}\|_{\mathbb{H}_p^{n+1}(\tau, l_2)} + \|g(u) - g_m(u)\|_{\mathbb{H}_p^{n+1}(\tau, l_2)} + (E\|u_0 - u_{0m}\|_{n+2-2/p, p}^p)^{1/p}. \end{aligned}$$

Next, by our assumptions about convergence of f_m, g_m, u_{0m} ,

$$(5.9) \quad \limsup_{m \rightarrow \infty} I_m \leq \limsup_{m \rightarrow \infty} \{ \|(a^{ij} - a_m^{ij}) u_{x^i x^j}\|_{\mathbb{H}_p^n(\tau)} + \|(\sigma^i - \sigma_m^i) u_{x^i}\|_{\mathbb{H}_p^{n+1}(\tau, l_2)} \}.$$

Here, by Lemma 5.2, for any $v \in C_0^\infty$ and r so large that $v\zeta_r = v$, we have

$$\begin{aligned} (5.10) \quad \|(a^{ij} - a_m^{ij}) u_{x^i x^j}\|_{n, p} &\leq N \|(u - v)_{x^i x^j}\|_{n, p} + \|(a^{ij} - a_m^{ij}) v_{x^i x^j}\|_{n, p}, \\ \|(a^{ij} - a_m^{ij}) v_{x^i x^j}\|_{n, p} &= \|\zeta_r(a^{ij} - a_m^{ij}) v_{x^i x^j}\|_{n, p} \leq N \|\zeta_r(a^{ij} - a_m^{ij})\|_{n, p} \|v\|_{B^{|\alpha|+2+\gamma}}, \end{aligned}$$

where the constants N do not depend on m and r . Thus,

$$\limsup_{m \rightarrow \infty} \|(a^{ij} - a_m^{ij}) u_{x^i x^j}\|_{n, p} \leq N \|(u - v)_{x^i x^j}\|_{n, p},$$

and, from the arbitrariness of v , we conclude that the left-hand side is zero for those ω and t for which $u \in H_p^{n+2}$. If we again apply Lemma 5.2, then we see that the p th power of the left-hand side of (5.10) is bounded by an integrable function. This and the dominated convergence theorem imply that

$$\lim_{m \rightarrow \infty} \|(a^{ij} - a_m^{ij}) u_{x^i x^j}\|_{\mathbb{H}_p^n(\tau)} = 0.$$

Similar arguments take care of the remaining term in (5.9). The theorem is proved.

REMARK 5.8. Condition (5.7) is satisfied for any $u \in \mathcal{H}_p^{n+2}(\tau)$ if and only if it is satisfied for $u(t, x) \equiv \phi(x)$ with any $\phi \in C_0^\infty$.

Indeed, the “only if” part is obvious. In the proof of “if” part notice that, under the “if” assumption, (5.7) is automatically satisfied for u of type

$$v = \sum_{i=1}^j I_{(\tau_{i-1}, \tau_i]}(t) v_i(x),$$

where τ_i are bounded stopping times and $v_i \in C_0^\infty$. By Theorem 3.10, one can approximate any $u \in \mathcal{H}_p^{n+2}(\tau) \subset \mathbb{H}_p^{n+2}(\tau)$ with functions like v . It remains only to notice that, by the assumptions of the theorem, for instance,

$$\begin{aligned} \|f(u, \cdot, \cdot) - f_m(u, \cdot, \cdot)\|_{\mathbb{H}_p^n(\tau)} &\leq \|f(u, \cdot, \cdot) - f(v, \cdot, \cdot)\|_{\mathbb{H}_p^n(\tau)} + \|f(v, \cdot, \cdot) - f_m(v, \cdot, \cdot)\|_{\mathbb{H}_p^n(\tau)} \\ + \|f_m(v, \cdot, \cdot) - f_m(u, \cdot, \cdot)\|_{\mathbb{H}_p^n(\tau)} &\leq N \|u - v\|_{\mathcal{H}_p^{n+2}(\tau)} + \|f(v, \cdot, \cdot) - f_m(v, \cdot, \cdot)\|_{\mathbb{H}_p^n(\tau)}. \end{aligned}$$

REMARK 5.9. While checking conditions (5.6) and (5.7), it is useful to bear in mind that, if one defines $\sigma_m^{ik} = \sigma^{ik}$ and $g_m^k = g^k$ for $k \leq m$ and $\sigma_m^{ik} = g_m^k = 0$ for $k > m$, then

$$\|\zeta_r\{\sigma^i(t, \cdot) - \sigma_m^i(t, \cdot)\}\|_{n+1,p} \rightarrow 0, \quad \|g(u, \cdot, \cdot) - g_m(u, \cdot, \cdot)\|_{\mathbb{H}_p^{n+1}(\tau, l_2)} \rightarrow 0.$$

Indeed, for instance,

$$\|g(u, \cdot, \cdot) - g_m(u, \cdot, \cdot)\|_{\mathbb{H}_p^{n+1}(\tau, l_2)}^p = E \int_0^\tau \left\| \left(\sum_{k>m} |(1-\Delta)^{(n+1)/2} g^k(u, t, \cdot)|^2 \right)^{1/2} \right\|_p^p dt,$$

which goes to zero by the dominated convergence theorem.

This fact allows one to approximate solutions of (5.1) by solutions of

$$(5.11) \quad \begin{aligned} du_m(t, x) = & [a^{ij}(t, x) u_{mx^i x^j}(t, x) + f(u_m, t, x)] dt \\ & + \sum_{k \leq m} [\sigma^{ik}(t, x) u_{mx^i}(t, x) + g^k(u_m, t, x)] dw_t^k. \end{aligned}$$

Before stating the following corollary, take the functions ζ_k from Lemma 5.2 and, for a function $h = h(u, t, x)$ defined for $t > 0$, $x \in \mathbb{R}^d$, and u in a function space, let

$$(5.12) \quad \hat{h}_m(u, t, x) = \int_{\mathbb{R}^d} h(u(\cdot + y), t, x - y) \zeta_m(y) dy.$$

To get a better idea about this definition, notice that if, for instance, $h(u, t, x) = c(t, x)u(t, x)$, where $c(t, x)$ is a given function, then $h(u(\cdot + y), t, x - y) = c(t, x - y)u(t, x)$, so that $\hat{h}_m(u, t, x) = c_m(t, x)u(t, x)$ with $c_m(t, x) = c(t, x) * \zeta_m(x)$.

COROLLARY 5.10. Under the assumptions of Theorem 5.1, define

$$(a_m, \sigma_m)(t, x) = (a, \sigma)(t, x) * \zeta_m(x),$$

$$(f_m, g_m) = (\hat{f}_m, \hat{g}_m) \quad \text{or} \quad (f_m, g_m)(u, t, x) = (f, g)(u, t, x) * \zeta_m(x).$$

Then the assumptions of Theorem 5.7 are satisfied, and if we define $u_m \in \mathcal{H}_p^{n+2}(\tau)$ as a solution of the Cauchy problem (5.8) with initial condition $u_m(0, \cdot) = u_0 * \zeta_m$, then $\|u - u_m\|_{\mathcal{H}_p^{n+2}(\tau)} \rightarrow 0$.

To prove this, notice that

$$(\sigma_m^i, \sigma_m^j)_{l_2} \lambda^i \lambda^j = |(\sigma^i \lambda^i) * \zeta_m|_{l_2}^2 \leq \zeta_m * [(\sigma^i, \sigma^j)_{l_2} \lambda^i \lambda^j],$$

which guarantees that the approximating equations satisfy the same coercivity condition. Also, for any $u \in H_p^{n+2}$ and $\phi \in C_0^\infty$, the function $(f(u(\cdot + z), t, \cdot - y), \phi) = (f(u(\cdot + z), t, \cdot), \phi(\cdot + y))$ is continuous in z owing to Assumption 5.6 and infinitely differentiable in y as it has to be for any distribution. Therefore, the definition of \hat{f}_m according to (5.12) makes sense as an integral of distribution with respect to the parameter y , namely

$$(\hat{f}_m(u, t, \cdot), \phi) := \int_{\mathbb{R}^d} \zeta_m(y) (f(u(\cdot + y), t, \cdot - y), \phi) dy.$$

Furthermore, $\hat{f}_m = \hat{f}_{0m} + \hat{f}_{1m}$, where $f_0 = f(0, t, x)$ and $f_1 = f(u, t, x) - f(0, t, x)$. By Lemma 5.2, we have $\hat{f}_{0m}(t, \cdot) \in H_p^n$. By Assumption 5.6,

$$|(\hat{f}_{1m}(u, t, \cdot), \phi)| \leq N \int_{\mathbb{R}^d} \zeta_m(y) \|\phi\|_{-n,q} \|u(\cdot + y)\|_{n+2,p} dy = N \|\phi\|_{-n,q} \|u\|_{n+2,p},$$

where $q = p/(p-1)$. Hence, \hat{f}_m is a H_p^n -valued function. Also, $f(u, t, x) * \zeta_m(x)$ is well defined because of Assumption 5.4. Similarly, g_m is well defined. Minkowski's inequality implies that all Assumptions 5.1–5.6 are satisfied for any m with the same $K, \delta, \kappa_\epsilon, K_\epsilon$. Other assumptions of Theorem 5.7 are satisfied due to continuity in the mean of summable functions and Assumptions 5.1–5.6. For instance, for any $\phi \in C_0^\infty$,

$$\|f(\phi, \cdot, \cdot) - \hat{f}_m(\phi, \cdot, \cdot)\|_{\mathbb{H}_p^n(\tau)} \leq I_{1m} + I_{2m},$$

where

$$\begin{aligned} I_{1m} &:= \|f(\phi, \cdot, \cdot) - f_m(\phi, \cdot, \cdot)\|_{\mathbb{H}_p^n(\tau)}, \quad f_m(\phi, t, x) = \int_{\mathbb{R}^d} f(\phi, t, x-y) \zeta_m(y) dy, \\ I_{2m} &:= \|f_m(\phi, \cdot, \cdot) - \hat{f}_m(\phi, \cdot, \cdot)\|_{\mathbb{H}_p^n(\tau)} \\ &\leq \int_{\mathbb{R}^d} \|f(\phi, \cdot, \cdot-y) - f(\phi(\cdot+y), \cdot, \cdot-y)\|_{\mathbb{H}_p^n(\tau)} \zeta_m(y) dy \\ &= \int_{\mathbb{R}^d} \|f(\phi, \cdot, \cdot) - f(\phi(\cdot+y), \cdot, \cdot)\|_{\mathbb{H}_p^n(\tau)} \zeta_m(y) dy \leq N \int_{\mathbb{R}^d} \|\phi - \phi(\cdot+y)\|_{n+2,p} \zeta_m(y) dy. \end{aligned}$$

Clearly (cf. Lemma 5.2) $I_{1m} + I_{2m} \rightarrow 0$ as $m \rightarrow \infty$, and the application of Remark 5.8 ends the argument.

COROLLARY 5.11. Let the assumptions of Theorem 5.1 be satisfied and also let them be satisfied for a $p = q$, where $q \geq 2$. Then the solution u from Theorem 5.1 belongs also to $\mathcal{H}_q^{n+2}(\tau)$.

To prove this, without loss of generality, assume $p < q$ and let v be the solution of the same initial value problem but belonging to $\mathcal{H}_q^{n+2}(\tau)$ (such a unique v exists by Theorem 5.1). We have only to show that $v = u$. In light of Corollary 5.10, we can approximate both v and u by solutions of equations with smooth coefficients and with $(f_m, g_m)(u, t, x) = (f, g)(u, t, x) * \zeta_m(x)$. We only need to show that the approximating solutions coincide. Observe that, for any $r \geq n$, for instance,

$$\begin{aligned} \|f_m(u, \cdot, \cdot) - f_m(v, \cdot, \cdot)\|_{r,p} &\leq N \|f(u, \cdot, \cdot) - f(v, \cdot, \cdot)\|_{n,p} \leq N\varepsilon \|u - v\|_{n+2,p} \\ &+ NK_\epsilon \|u - v\|_{n,p} \leq N\varepsilon \|u - v\|_{r+2,p} + NK_\epsilon \|u - v\|_{r,p}, \end{aligned}$$

where N 's depend only on m, r, n, p . This shows that, for any fixed m , Assumptions 5.1–5.6 are satisfied with any large n both for p and q .

Therefore, we can suppose from the very beginning that Assumptions 5.1–5.6 are satisfied for any large n for p and q . In this case, by Theorem 5.1, $u \in \mathcal{H}_p^r(\tau)$ for any r , and hence, also invoking Theorem 3.7,

$$E \int_0^\tau \|u_{xx}(t, \cdot)\|_{r,p}^p dt < \infty, \quad E \sup_{t \leq \tau} \|u(t, \cdot)\|_{C^r}^p < \infty$$

for any $r = 1, 2, 3, \dots$. Since $\|u_{xx}(t, \cdot)\|_{r,i}^i \leq N \|u_{xx}(t, \cdot)\|_{r,p}^p \|u(t, \cdot)\|_{C^{r+2}}^{i-p}$ for $i \geq p$, it follows that

$$\int_0^\tau \|u_{xx}(t, \cdot)\|_{r,i}^i dt < \infty \quad (\text{a.s.})$$

for any $i \geq p$. Take here $r = 0$ and $i = q$ and define

$$\tau_k = \tau \wedge \inf\{t : \int_0^t \|u_{xx}(s, \cdot)\|_q^q ds \geq k\}.$$

Then obviously $u \in \mathcal{H}_q^2(\tau_k)$. Since v lies in the same class, by uniqueness $u(t, \cdot) = v(t, \cdot)$ for $t \leq \tau_k$ (a.s.). It remains only to observe that $\tau_k \uparrow \tau$ when $k \rightarrow \infty$.

Next general result concerns the maximum principle.

THEOREM 5.12 (maximum principle). *Let the assumptions of Theorem 5.1 be satisfied and let u be the function existence of which is asserted in this theorem. Assume that*

$f(v, t, x) = b^i(t, x)v_{x^i}(t, x) + v(t, x)c(t, x) + f(t, x)$, $g^k(v, t, x) = v(t, x)\nu^k(t, x)$, where $b^i(t, x), c(t, x), \nu^k(t, x)$ are certain bounded functions on $(0, \tau] \times \mathbb{R}^d$ and $f(t, x) = f(0, t, x) \geq 0$. Also assume that for any ω we have $u_0 \geq 0$. Then $u(t, \cdot) \geq 0$ for all $t \in [0, \tau]$ almost surely.

Proof. Because of Remark 5.9, we can only concentrate on equations with finite number of Wiener processes like (5.11), which in our case is the following

$$(5.13) \quad du(t, x) = [a^{ij}(t, x)u_{x^i x^j}(t, x) + b^i(t, x)u_{x^i}(t, x) + u(t, x)c(t, x) + f(t, x)] dt + \sum_{k \leq m} [\sigma^{ik}(t, x)u_{x^i}(t, x) + u(t, x)\nu^k(t, x)] dw_t^k.$$

Corollary 5.10 and the explanation after (5.12) allow us to assume that $u_0(x)$ and the coefficients and f in (5.13) are infinitely differentiable in x . After this, by multiplying u_0 and f by a cut-off function of x and by using Remark 5.6 and Theorem 5.7, we convince ourselves that we can assume that u_0 and $f(t, x)$ have supports with respect to x in a fixed ball. In this case the assumptions of Theorem 5.1 are satisfied for $p = 2$ again by Remark 5.6 (and Hölder's inequality), which, by Corollary 5.11, yields $u \in \mathcal{H}_2^n(\tau)$ for any n . Now our assertion follows from the maximum principle proved in [25] (see Theorem 4.2 there). The only point to mention is that in [25] we considered (5.13) on $[0, T]$, but we always can continue our data in an appropriate way after τ , which is assumed to be less than T . The theorem is proved.

By the way the proof of the maximum principle in [25] is based on representation formulas like (4.28). A heuristic derivation of many other formulas of that kind can be found in [15].

6. Proof of Theorem 5.1

The proof we present here is quite typical for proofs of solvability of equations with variable coefficients on the basis of solvability of equations with constant ones. The same type of arguments is commonly used in the theory of partial differential equations for proving the solvability in Sobolev or Hölder spaces. First we need some auxiliary constructions. Fix a $T \in (0, \infty)$.

DEFINITION 6.1. Assume that, for $\omega \in \Omega$ and $t \geq 0$, we are given operators

$$L(t, \cdot) : H_p^{n+2} \rightarrow H_p^n, \quad \Lambda(t, \cdot) : H_p^{n+2} \rightarrow H_p^{n+1}(\mathbb{R}^d, l_2).$$

Assume that

- (i) for any ω and t , the operators $L(t, u)$ and $\Lambda(t, u)$ are continuous (with respect to u);
- (ii) for any $u \in H_p^{n+2}$, the processes $L(t, u)$ and $\Lambda(t, u)$ are predictable;
- (iii) for any $\omega \in \Omega$, $t \geq 0$, and $u \in H_p^{n+2}$, we have

$$\|L(t, u)\|_{n,p} + \|\Lambda(t, u)\|_{n+1,p} \leq N_{L,\Lambda}(1 + \|u\|_{n+2,p}),$$

where $N_{L,\Lambda}$ is a constant.

Then for a function $u \in \mathcal{H}_p^{n+2}(T)$, we write

$$(L, \Lambda)u = -(f, g)$$

if $(f, g) \in \mathcal{F}_p^n(T)$, and, in the sense of Definition 3.5, for $t \in [0, T]$, we have that $\mathbb{D}u(t) = L(t, u(t)) + f(t)$ and $\mathbb{S}u(t) = \Lambda(t, u(t)) + g(t)$, or put otherwise

$$u(t) = u(0) + \int_0^t (L(s, u(s)) + f(s)) ds + \int_0^t (\Lambda^k(s, u(s)) + g^k(s)) dw_s^k \quad (\text{a.s.}).$$

REMARK 6.2. By virtue of our conditions on L and Λ , for any $u \in \mathcal{H}_p^{n+2}(T)$, we have $(L(u), \Lambda(u)) \in \mathcal{F}_p^n(T)$. Also, $(L, \Lambda)u = (L(u) - \mathbb{D}u, \Lambda(u) - \mathbb{S}u)$. In particular, the operator (L, Λ) is well defined on $\mathcal{H}_p^{n+2}(T)$, and, as follows easily from Definition 6.1 (iii),

$$\|(L, \Lambda)u\|_{\mathcal{F}_p^n(T)} \leq (1 + 2N_{L, \Lambda})\|u\|_{\mathcal{H}_p^{n+2}(T)} + 2N_{L, \Lambda}T^{1/p}.$$

In terms of Definition 6.1, Theorem 4.10 has the following version.

THEOREM 6.3. *Let a and σ satisfy the assumptions from the beginning of Sec. 4. Define*

$$Lu = a^{ij}u_{x^i x^j}, \quad \Lambda u = \sigma^i u_{x^i}.$$

Then the operator (L, Λ) is a one-to-one operator from $\mathcal{H}_{p,0}^{n+2}(T)$ onto $\mathcal{F}_p^n(T)$ and the norm of its inverse is less than a constant depending only on d, p, δ , and K (thus independent of T).

Next, we prove a perturbation result. It needs a proof because we do not allow ε to depend on T .

THEOREM 6.4. *Take the operators L and Λ from Theorem 6.3, and let some operators L_1 and Λ_1 satisfy the requirements from Definition 6.1. We assert that there exists a constant $\varepsilon \in (0, 1)$ depending only on d, p, δ , and K such that if, for a constant K_1 and any $u, v \in H_p^{n+2}$, $t \geq 0$, $\omega \in \Omega$, we have*

$$(6.1) \quad \begin{aligned} & \|L_1(t, u) - L_1(t, v)\|_{n,p} + \|\Lambda_1(t, u) - \Lambda_1(t, v)\|_{n+1,p} \\ & \leq \varepsilon \|u_{xx} - v_{xx}\|_{n,p} + K_1 \|u - v\|_{n+1,p}, \end{aligned}$$

then, for any $(f, g) \in \mathcal{F}_p^n(T)$, there exists a unique solution $u \in \mathcal{H}_{p,0}^{n+2}(T)$ of the equation

$$(6.2) \quad (L + L_1, \Lambda + \Lambda_1)u = -(f, g).$$

Furthermore, for this solution u , we have

$$(6.3) \quad \|u\|_{\mathcal{H}_p^{n+2}(T)} \leq N\|(L_1(\cdot, 0) + f, \Lambda_1(\cdot, 0) + g)\|_{\mathcal{F}_p^n(T)},$$

where N depends only on d, p, δ, K, K_1 , and T and N is independent of T if $K_1 = 0$.

Proof. First notice that, by interpolation theorems, $\|u\|_{n+1,p} \leq \varepsilon \|u_{xx}\|_{n,p} + N(\varepsilon, d, p)\|u\|_{n,p}$. Therefore, without loss of generality we assume that instead of (6.1) we have

$$\begin{aligned} & \|L_1(t, u) - L_1(t, v)\|_{n,p} + \|\Lambda_1(t, u) - \Lambda_1(t, v)\|_{n+1,p} \\ & \leq \varepsilon \|u_{xx} - v_{xx}\|_{n,p} + K_1 \|u - v\|_{n+1,p}. \end{aligned}$$

Now fix $(f, g) \in \mathcal{F}_p^n(T)$. Take $u \in \mathcal{H}_{p,0}^{n+2}(T)$, observe that $(L_1(u), \Lambda_1(u)) \in \mathcal{F}_p^n(T)$, and, by using Theorem 6.3, define $v \in \mathcal{H}_{p,0}^{n+2}(T)$ as the unique solution of the equation $(L, \Lambda)v = -(f + L_1(u), g + \Lambda_1(u))$. By denoting $v = Ru$, we define

an operator $R : \mathcal{H}_{p,0}^{n+2}(T) \rightarrow \mathcal{H}_{p,0}^{n+2}(T)$. Equation (6.2) is equivalent to the equation $u = Ru$. Therefore, to prove the existence and uniqueness of solutions to (6.2), it suffices to show that, for an integer $m > 0$, the operator R^m is a contraction in $\mathcal{H}_{p,0}^{n+2}(T)$.

By Theorem 6.3, for $t \leq T$,

$$\begin{aligned} \|Ru - Rv\|_{\mathcal{H}_p^{n+2}(t)}^p &\leq N \| (L_1(u) - L_1(v), \Lambda_1(u) - \Lambda_1(v)) \|_{\mathcal{F}_p^n(t)}^p \\ &\leq N_0 \varepsilon^p \|u - v\|_{\mathcal{H}_p^{n+2}(t)}^p + N_0 K_1^p \int_0^t E \|u(s) - v(s)\|_{n,p}^p ds, \end{aligned}$$

with a constant N_0 depending only on d, p, δ , and K . This gives the desired result if $K_1 = 0$. Also, in this case estimate (6.3) follows obviously with N independent of T .

In the general case, by Theorem 3.7,

$$E \|u(s) - v(s)\|_{n,p}^p \leq N_1 \|u - v\|_{\mathcal{H}_p^{n+2}(s)}^p,$$

where $s \leq T$ and N_1 depends only on d, p , and T . It follows that, for $t \leq T$ and $\theta := N_0 \varepsilon^p$, we have

$$\|Ru - Rv\|_{\mathcal{H}_p^{n+2}(t)}^p \leq \theta \|u - v\|_{\mathcal{H}_p^{n+2}(t)}^p + N_2 \int_0^t \|u - v\|_{\mathcal{H}_p^{n+2}(s)}^p ds,$$

where N_2 depends only on d, p, δ, K, K_1 , and T . Hence, by induction,

$$\begin{aligned} &\|R^m u - R^m v\|_{\mathcal{H}_p^{n+2}(t)}^p \leq \theta^m \|u - v\|_{\mathcal{H}_p^{n+2}(t)}^p \\ &+ \sum_{k=1}^m \binom{m}{k} \theta^{m-k} N_2^k \int_0^t \frac{(t-s)^{k-1}}{(k-1)!} \|u - v\|_{\mathcal{H}_p^{n+2}(s)}^p ds, \quad \|R^m u - R^m v\|_{\mathcal{H}_p^{n+2}(T)}^p \\ &\leq \sum_{k=0}^m \binom{m}{k} \theta^{m-k} \frac{1}{k!} (TN_2)^k \|u - v\|_{\mathcal{H}_p^{n+2}(T)}^p \leq 2^m \theta^m \max_k \frac{1}{k!} (TN_2/\theta)^k \|u - v\|_{\mathcal{H}_p^{n+2}(T)}^p. \end{aligned}$$

This allows us to find ε depending only on d, p, δ , and K and m depending on the same things plus K_1 and T , so that R^m is a contraction in $\mathcal{H}_{p,0}^{n+2}(T)$ with coefficient $1/2$. Of course, this yields all our assertions. The theorem is proved.

We finish our preparations by showing how Lemma 5.2 will be used.

REMARK 6.5. To some extent, in what follows, the most important consequence of assertion (i) in Lemma 5.2 is that if $\bar{a} = \|a\|_{B^{|n|+\gamma}} < \infty$, then there exists a new norm $\|\cdot\|_{[n,p]}$ in H_p^n such that

$$\|au\|_{[n,p]} \leq 2N \|a\|_B \|u\|_{[n,p]} \quad (\|a\|_B = \sup_{x \in \mathbb{R}^d} |a(x)|),$$

where N is the same constant as in Lemma 5.2. To show this, it suffices to observe that, for $a_m(x) = a(x/m)$, $u_m(x) = u(x/m)$, and $m \geq 1$, we have

$$\begin{aligned} \|(m^2 - \Delta)^{n/2}(au)\|_p &= m^{n-d/p} \|(1 - \Delta)^{n/2}(a_m u_m)\|_p \leq \\ &N \bar{a}_m m^{n-d/p} \|(1 - \Delta)^{n/2} u_m\|_p = N \bar{a}_m \|(m^2 - \Delta)^{n/2} u\|_p \leq \\ &N \left(\|a\|_B + \frac{1}{m^{(|n|+\gamma)\wedge 1}} \bar{a} \right) \|(m^2 - \Delta)^{n/2} u\|_p. \end{aligned}$$

Alternatively, it would be sufficient for our needs to know that

$$\|au\|_{[n,p]} \leq N (\|a\|_B \|u\|_{[n,p]} + \|a\|_{B^{|n|+\gamma}} \|u\|_{n-1,p}).$$

The author could not find the last inequality in the literature, though some very interesting information related to the subject can be found in [13].

Now we perform the main step in proving Theorem 5.1.

LEMMA 6.6. *Let Assumptions 5.1, 5.3, 5.4, and 5.5 be satisfied. Then there exists an $\varepsilon = \varepsilon(d, p, n, \gamma, \delta, K) > 0$ such that if $\tau = T$ and*

- (i) *inequality (5.3) holds with this ε for all x, y, t , and ω and*
- (ii) *f and g are independent of u ,*

then there exists a unique solution $u \in \mathcal{H}_{p,0}^{n+2}(T)$ of equation (5.1). Furthermore, for this solution u , we have

$$\|u\|_{\mathcal{H}_p^{n+2}(T)} \leq N \|(f, g)\|_{\mathcal{F}_p^n(T)},$$

where N depends only on d, p, δ, K , and T .

Proof. Define $a(t) = a(t, 0)$ and $\sigma(t) = \sigma(t, 0)$, take operators L and Λ from Theorem 6.3 corresponding to these $a(t)$ and $\sigma(t)$, and let

$$L_1(t, u)(x) = [a^{ij}(t, x) - a^{ij}(t)]u_{x^i x^j}(x), \quad \Lambda_1(t, u)(x) = [\sigma^i(t, x) - \sigma^i(t)]u_{x^i}(x).$$

In view of Theorem 6.4, to prove existence and uniqueness, we have only to check that if ε in (5.3) is sufficiently small, then the operators L_1 and Λ_1 satisfy condition (6.1) with as small ε as we like and with K_1 under control. Observe that by Lemma 5.2,

$$\|L_1(t, u)\|_{n,p} \leq N \|a(t, \cdot) - a(t, 0)\|_{B^{|n|+\gamma}} \|u_{xx}\|_{n,p},$$

$$\|\Lambda_1(t, u)\|_{n+1,p} \leq N \|\sigma(t, \cdot) - \sigma(t, 0)\|_{B^{|n+1|+\gamma}(I_2)} \|u_x\|_{n+1,p}.$$

Since $\|u_x\|_{n+1,p} \leq N(\|u_{xx}\|_{n,p} + \|u\|_{n+1,p})$, our lemma holds true if

$$\|a(t, \cdot) - a(t, 0)\|_{B^{|n|+\gamma}} + \|\sigma(t, \cdot) - \sigma(t, 0)\|_{B^{|n+1|+\gamma}(I_2)} \leq \varepsilon_0 = \varepsilon_0(d, p, n, \gamma, \delta, K) \quad \forall t.$$

Next, observe that, for $a_m(t, x) = a(t/m^2, x/m)$ and $m \geq 1$, we have

$$\|a_m(t, \cdot) - a_m(t, 0)\|_{B^{|n|+\gamma}} \leq \varepsilon + m^{-[(|n|+\gamma)\wedge 1]} K$$

(K comes from Assumption 5.5), and for $|n| + \gamma = 0$ we can even drop the second term on the right. An analogous inequality holds for σ . It follows that, for ε sufficiently small and m sufficiently large, the statements of the lemma are true if we replace a, σ, w_t, f, g , and T in equation (5.1) by $a_m, \sigma_m, m w_{t/m^2}, m^{-2} f(t/m^2, x/m), m^{-1} g(t/m^2, x/m)$, and $m^2 T$, respectively. After this, it remains only to fix an appropriate m and make an obvious change of the unknown function in the above-mentioned modification of equation (5.1) (and use $1 - \Delta \sim m^2 - \Delta$ so that the norms $\|\cdot\|_{n,p}$ of $u(t/m^2, x/m)$ and $u(t, x)$ are comparable). The lemma is proved.

Finally, we need the following result from [17], which, in a sense, is essentially covered by Theorem 2.4.7 from [34].

LEMMA 6.7. *Let $\delta > 0$ and let $\zeta_k \in C^\infty$, $k = 1, 2, 3, \dots$. Assume that for any multi-index α and $x \in \mathbb{R}^d$,*

$$\sup_{x \in \mathbb{R}^d} \sum_k |D^\alpha \zeta_k(x)| \leq M(\alpha),$$

where $M(\alpha)$ are some constants. Then there exists a constant $N = N(d, n, M)$ such that, for any $f \in H_p^n$,

$$\sum_k \|\zeta_k f\|_{H_p^n}^p \leq N \|f\|_{H_p^n}^p.$$

If in addition

$$\sum_k |\zeta_k(x)|^p \geq \delta,$$

then for any $f \in H_p^n$,

$$\|f\|_{n,p}^p \leq N(d, n, M, \delta) \sum_k \|\zeta_k f\|_{n,p}^p.$$

REMARK 6.8. We will also use a natural extension of this lemma to the case of Banach-space valued f .

Proof of Theorem 5.1. By Theorem 2.1, for any nonrandom $z \in H_p^{n+2-2/p}$ there exists a unique solution $u \in \mathcal{H}_p^{n+2}(\tau)$ of the equation $du = \Delta u dt$ with initial condition z . This theorem also provides estimate (2.7) of the norm of u . From this theorem and the estimate, it follows that there exists a unique solution $\bar{u} \in \mathcal{H}_p^{n+2}(\tau)$ of the equation $du = \Delta u dt$ with initial condition u_0 and

$$\|\bar{u}\|_{\mathcal{H}_p^{n+2}(\tau)}^p \leq N E \|u_0\|_{n+2-2/p,p}^p.$$

After replacing the unknown function u with $v + \bar{u}$, it becomes obvious that, to solve (5.1) in $\mathcal{H}_p^{n+2}(\tau)$ with the initial condition u_0 , it suffices to solve the equation

$$\begin{aligned} du(t, x) &= [a^{ij}(t, x)u_{x^i x^j}(t, x) + \bar{f}(u, t, x)] dt \\ &\quad + [\sigma^{ik}(t, x)u_{x^i}(t, x) + \bar{g}^k(u, t, x)] dw_t^k \end{aligned}$$

in the space $\mathcal{H}_{p,0}^{n+2}(\tau)$, where (remember $d\bar{u} = \Delta \bar{u} dt$)

$$\bar{f}(u, t, x) = \{f(u + \bar{u}, t, x) + a^{ij}(t, x)\bar{u}_{x^i x^j}(t, x) - \Delta \bar{u}\} I_A(t),$$

$$\bar{g}^k(u, t, x) = \{\sigma^{ik}(t, x)\bar{u}_{x^i}(t, x) + g^k(u + \bar{u}, t, x)\} I_A(t),$$

and A is the set of ω, t for which $\bar{u}(t, \cdot) \in H_p^{n+2}$.

From Lemma 5.2, it follows that \bar{f} and \bar{g} satisfy the same conditions as f and g . This allows us to assume that $u_0 = 0$. Furthermore, one can always extend f, g after time τ by equaling them to zero. That is why without loss of generality we assume that $u_0 = 0$ and $\tau = T$.

Now, define

$$Lu = a^{ij}(t, x)u_{x^i x^j}(t, x), \quad \Lambda u = \sigma^i(t, x)u_{x^i}(t, x),$$

and let $\{\zeta_k : k = 1, 2, 3, \dots\}$ be a standard partition of unity such that, for any k , the support of ζ_k lies in a ball B_k of radius $(1/4)\kappa_{\varepsilon/2}$, where κ_{ε} is taken from Assumption 5.2 and ε from Lemma 6.6. Also for any k , we take a function $\eta_k \in C_0^\infty$ such that $\eta_k = 1$ on B_k , $\eta_k = 0$ outside doubled B_k , and $0 \leq \eta_k \leq 1$. Denote by x_k the center of B_k , define $L_k(t, x) = \eta_k(x)L(t, x) + (1 - \eta_k(x))L(t, x_k)$, and similarly define Λ_k .

Observe that, for any k , the operators L_k, Λ_k satisfy condition (i) of Lemma 6.6. Therefore, if we denote $(f_k, g_k) := (L_k, \Lambda_k)(u\zeta_k)$, then by this lemma

$$\|u\zeta_k\|_{\mathcal{H}_p^{n+2}(T)} \leq N \|(f_k, g_k)\|_{\mathcal{F}_p^n(T)} = N \|(L_k, \Lambda_k)(u\zeta_k)\|_{\mathcal{F}_p^n(T)}.$$

Furthermore, by using that $\eta_k = 1$ everywhere where $\zeta_k u \neq 0$ we easily check that

$$(L_k, \Lambda_k)(u\zeta_k) = (L, \Lambda)(u\zeta_k) = \zeta_k(L, \Lambda)u + (uL\zeta_k + 2\zeta_{kx} \cdot a u_x, u\Lambda\zeta_k),$$

so that

$$\|u\zeta_k\|_{\mathcal{H}_p^{n+2}(T)} \leq N \|\zeta_k(L, \Lambda)u\|_{\mathcal{F}_p^n(T)} + N \|(uL\zeta_k + 2\zeta_{kx} \cdot a u_x, u\Lambda\zeta_k)\|_{\mathcal{F}_p^n(T)}.$$

We sum up the p th powers of the extreme terms and apply Lemma 6.7 and Lemma 5.2 in the estimates like the following one:

$$\sum_k \|ua^{ij}\zeta_{kx^i x^j}\|_{n,p} \leq N\|ua^{ij}\|_{n,p} \leq N\|u\|_{n,p}.$$

Then we conclude that, for any $u \in \mathcal{H}_{p,0}^{n+2}(T)$,

$$(6.4) \quad \|u\|_{\mathcal{H}_p^{n+2}(T)} \leq N\|(L, \Lambda)u\|_{\mathcal{F}_p^n(T)} + N\|u\|_{\mathbb{H}_p^{n+1}(T)}.$$

Next, we show that the last term on the right can be dropped. Indeed, $\|u\|_{n+1,p} \leq \varepsilon\|u_{xx}\|_{n,p} + N(\varepsilon, d, p)\|u\|_{n,p}$. Therefore, (6.4) can be modified by replacing the last term with $N\|u\|_{\mathbb{H}_p^n(T)}$. This can be done with any $t \leq T$ in place of T . After this, by Theorem 3.7, the inequality

$$E\|u(t, \cdot)\|_{n,p}^p \leq N\|(L, \Lambda)u\|_{\mathcal{F}_p^n(T)}^p + N \int_0^t E\|u(s, \cdot)\|_{n,p}^p ds$$

holds for any $t \leq T$. By Gronwall's inequality, this yields that

$$\|u\|_{\mathbb{H}_p^n(T)}^p \leq N\|(L, \Lambda)u\|_{\mathcal{F}_p^n(T)}^p$$

which, along with the modified (6.4), proves that

$$(6.5) \quad \|u\|_{\mathcal{H}_p^{n+2}(T)} \leq N\|(L, \Lambda)u\|_{\mathcal{F}_p^n(T)}$$

for any $u \in \mathcal{H}_{p,0}^{n+2}(T)$.

This is an a priori estimate. Now we use the standard method of continuity. For $\lambda \in [0, 1]$ we consider the equation

$$(6.6) \quad du = (L_\lambda u + f) dt + (\Lambda_\lambda^k u + g^k) dw_t^k$$

with zero initial condition, where

$$L_\lambda u = \lambda \Delta + (1 - \lambda)L, \quad \Lambda_\lambda u = (1 - \lambda)\Lambda$$

and (f, g) is an arbitrary element in $\mathcal{F}_p^n(T)$.

Observe that the a priori estimate (6.5) holds with the same constant N for all $L_\lambda, \Lambda_\lambda$ in place of L, Λ . Next, take a $\lambda_0 \in [0, 1]$ and assume that for $\lambda = \lambda_0$ equation (6.6) with zero initial data has a unique solution $u \in \mathcal{H}_{p,0}^{n+2}(T)$ for any $(f, g) \in \mathcal{F}_p^n(T)$. By the way, this assumption is satisfied for $\lambda_0 = 1$ by Theorem 6.3. Then we have an operator

$$\mathcal{R}_{\lambda_0} : \mathcal{F}_p^n(T) \rightarrow \mathcal{H}_{p,0}^{n+2}(T)$$

such that $\mathcal{R}_{\lambda_0}(f, g) = u$. From (6.5) we get that

$$(6.7) \quad \|\mathcal{R}_{\lambda_0}(f, g)\|_{\mathcal{H}_p^{n+2}(T)} \leq N\|(f, g)\|_{\mathcal{F}_p^n(T)}.$$

For other $\lambda \in [0, 1]$ we rewrite (6.6) as

$$du = (L_{\lambda_0}u + \{(\lambda - \lambda_0)(\Delta - L)u + f\}) dt + (\Lambda_{\lambda_0}^k u + \{(\lambda_0 - \lambda)\Lambda^k u + g^k\}) dw_t^k$$

and we solve the last equation by iterations. Define $u_0 = 0$ and

$$u_{j+1} = \mathcal{R}_{\lambda_0}((\lambda - \lambda_0)(\Delta - L)u_j + f, (\lambda_0 - \lambda)\Lambda u_j + g).$$

Then by (6.7)

$$\begin{aligned} \|u_{j+1} - u_j\|_{\mathcal{H}_p^{n+2}(T)} &\leq N|\lambda - \lambda_0| \|((\Delta - L)(u_j - u_{j-1}), \Lambda(u_j - u_{j-1}))\|_{\mathcal{F}_p^n(T)} \\ &\leq N_1 |\lambda - \lambda_0| \|u_j - u_{j-1}\|_{\mathcal{H}_p^{n+2}(T)}, \end{aligned}$$

where N_1 is independent of j , λ , and λ_0 . If $N_1|\lambda - \lambda_0| \leq 1/2$, then u_j is a Cauchy sequence in $\mathcal{H}_p^{n+2}(T)$, which converges by Theorem 3.7. Its limit satisfies

$$u = \mathcal{R}_{\lambda_0}((\lambda - \lambda_0)(\Delta - L)u + f, (\lambda_0 - \lambda)\Lambda u + g),$$

which is equivalent to (6.6).

In this way we show that if (6.6) is solvable for a λ_0 , then it is solvable for λ satisfying $N_1|\lambda - \lambda_0| \leq 1/2$. In finite number of steps starting with $\lambda = 1$, we get to $\lambda = 0$. This proves the theorem if f, g are independent of u .

To consider general f and g , it remains only to repeat the proof of Theorem 6.4 taking $f(u, t, x)$ and $g(u, t, x)$ instead of $f + L_1(u)$ and $g + \Lambda_1(u)$ there. The theorem is proved.

7. Embedding Theorems for $\mathcal{H}_p^n(\tau)$

The above theory of solvability looks satisfactory only until one tries to apply it to concrete problems when one is interested in getting not only solvability but also some qualitative properties of solutions, like continuity, decay at infinity, compactness of support, and so on. To answer such questions, one has to understand what qualitative properties the solutions have. Since solutions are just arbitrary functions from $\mathcal{H}_p^n(\tau)$, we are actually interested in properties of functions from this space. Let us fix a $T \in [0, \infty)$ and a stopping time $\tau \leq T$.

The first two assertions of the following theorem are straightforward corollaries of two Sobolev's theorems. One says that $H_p^n \subset C^\alpha$ if $\alpha := n - d/p > 0$, where $C^\alpha = C^\alpha(\mathbb{R}^d)$ is the Zygmund space (which differs from the usual Hölder space $C^\alpha = C^\alpha(\mathbb{R}^d)$ only if α is an integer, see [33]). The second one says that $H_p^n \subset H_q^m$ if $m < n$ and $n - d/p = m - d/q$.

THEOREM 7.1. (i) If $\alpha := n - d/p > 0$ and $u \in \mathcal{H}_p^n(\tau)$, then $u \in L_p([0, \tau], C^\alpha)$, where C^α is the Zygmund space. In addition,

$$E \int_0^\tau \|u(t, \cdot)\|_{C^\alpha}^p dt \leq N(d, n, p) \|u\|_{\mathcal{H}_p^n(\tau)}^p.$$

(ii) If $m < n$, $n - d/p = m - d/q$, and $u \in \mathcal{H}_p^n(\tau)$, then

$$E \int_0^\tau \|u(t, \cdot)\|_{m,q}^p dt \leq N(d, n, m, p) \|u\|_{\mathcal{H}_p^n(\tau)}^p.$$

(iii) For any function $u \in \mathcal{H}_2^n(\tau)$, we have $u \in C([0, \tau], H_2^{n-1})$ (a.s.) and

$$E \sup_{t \leq \tau} \|u(t, \cdot)\|_{n-1,2}^2 \leq N(d, n, T) \|u\|_{\mathcal{H}_2^n(\tau)}^2.$$

Proof. As we have said before the theorem, we only need to prove the third assertion. By Remark 3.4, we may assume that $n = 1$. Denote $u_0 = u(0)$, $v = T_t u_0$. Observe that by Theorem 2.1 we have $v \in \mathcal{H}_2^1(\tau)$. By Minkowski's inequality, $\|\zeta * u\|_p \leq \|u\|_p \|\zeta\|_1$, so that

$$\|T_t u_0\|_2 \leq \|u_0\|_2, \quad E \sup_t \|v(t, \cdot)\|_2^2 \leq E \|u_0\|_2^2.$$

In addition, almost obviously, $T_t u_0$ is a continuous (analytic) L_2 -valued function in t for $t > 0$. Also, we have $(T_t u_0)'_t = T_t \Delta u_0$, which implies that $\|T_t u_0 - u_0\|_2 \leq t \|\Delta u_0\|_2 \rightarrow 0$ if $u_0 \in H_2^2$. Adding that the set H_2^2 is dense in L_2 , we conclude that T_t is a continuous semigroup in L_2 . This means that $v \in C([0, \tau], H_2^{n-1})$ and shows

that we need only to consider $u - v$. In other words, in the rest of the proof we may and will assume that $u_0 = 0$.

In this case, denote $f = (\mathbb{D}u - \Delta u)I_{t \leq \tau}$, and $g = (\mathbb{S}u)I_{t \leq \tau}$. Solve equation (4.2) on $[0, \infty)$ with zero initial data. By uniqueness, the solution coincides with u on $[0, \tau]$. By Theorem 4.2, assertion (iii) holds. The theorem is proved.

Further results about some basic properties of the spaces $\mathcal{H}_p^n(\tau)$ are collected in the following theorem. As we have seen in Theorem 4.2, assertion (iii) of Theorem 7.1 is true not only for $p = 2$ but also for any $p \geq 2$. However, for $p > 2$ a much stronger statement (i) of Theorem 7.2 holds.

THEOREM 7.2. (i) If $p > 2$, $1/2 > \beta > \alpha > 1/p$, then for any function $u \in \mathcal{H}_p^n(\tau)$, we have $u \in C^{\alpha-1/p}([0, \tau], H_p^{n-2\beta})$ (a.s.) and for any stopping time $\eta \leq \tau$,

$$(7.1) \quad E\|u(t \wedge \eta, \cdot) - u(s \wedge \eta, \cdot)\|_{n-2\beta, p}^p \leq N(d, \beta, p, T)|t-s|^{\beta p-1}\|u\|_{\mathcal{H}_p^n(\tau)}^p \quad \forall t, s \leq T;$$

$$(7.2) \quad E\|u(t, \cdot)\|_{C^{\alpha-1/p}([0, \tau], H_p^{n-2\beta})}^p \leq N(d, \beta, \alpha, p, T)\|u\|_{\mathcal{H}_p^n(\tau)}^p.$$

(ii) If $q \geq p > 2$ and $\theta \in (0, 1)$, then for

$$m < n + \frac{d}{q} - \frac{d+2(1-\theta)}{p}, \quad u \in \mathcal{H}_p^n(\tau),$$

we have $u \in L_{p/\theta}((0, \tau), H_q^m)$ (a.s.) and

$$E\left(\int_0^\tau \|u(t, \cdot)\|_{m,q}^{p/\theta} dt\right)^\theta \leq N(d, p, q, n, m, \theta, T)\|u\|_{\mathcal{H}_p^n(\tau)}^p.$$

In particular (take $\theta = p/q$),

$$E\left(\int_0^\tau \|u(t, \cdot)\|_{m,q}^q dt\right)^{p/q} \leq N(d, p, q, n, m, T)\|u\|_{\mathcal{H}_p^n(\tau)}^p$$

if

$$q > p > 2, \quad m < n - (d+2)\left(\frac{1}{p} - \frac{1}{q}\right).$$

To prove the theorem, we need two lemmas. Remember the notation T_t introduced in (2.3).

LEMMA 7.3. For any $h \in L_p$, $\theta \in [0, 1]$, and $t > 0$, we have

$$(7.3) \quad \|e^{-t}T_t h\|_p \leq N \frac{1}{t^\theta} \|h\|_{-2\theta, p}, \quad \|(T_t - 1)h\|_p \leq N t^\theta \|h\|_{2\theta, p},$$

where $N = N(d, p, \theta)$.

Proof. These inequalities follow from Theorem 14.11 of [14]. For the sake of completeness, we prove them. The derivative $(T_t h)'_t$ can be represented as t^{-1} times a convolution of h with a function having finite L_1 -norm, the norm being independent of t . By Minkowski's inequality, $\|\zeta * h\|_p \leq \|h\|_p \|\zeta\|_1$ so that

$$(7.4) \quad \|(T_t h)'_t\|_p \leq N(d, p) \frac{1}{t} \|h\|_p, \quad \|(e^{-t}T_t h)'_t\|_p \leq N(d, p) \frac{1}{t} \|h\|_p.$$

Hence for $\theta \in (0, 1)$, (see (2.2))

$$\|(1 - \Delta)^\theta (e^{-t}T_t h)\|_p = c(\theta) \left\| \int_0^\infty \frac{e^{-(s+t)}T_s u - e^{-t}T_t h}{s^\theta} \frac{ds}{s} \right\|_p$$

$$\leq N \int_0^t \frac{1}{s^\theta} \|e^{-(s+t)} T_{s+t} u - e^{-t} T_t h\|_p \frac{ds}{s} + N \int_t^\infty \frac{ds}{s^{1+\theta}} \|h\|_p.$$

Here by (7.4)

$$\begin{aligned} \|e^{-(s+t)} T_{s+t} u - e^{-t} T_t h\|_p &\leq N \frac{s}{t} \|h\|_p, \\ \int_0^t \frac{1}{s^\theta} \|e^{-(s+t)} T_{s+t} u - e^{-t} T_t h\|_p \frac{ds}{s} &\leq \|h\|_p \int_0^t \frac{1}{ts^\theta} ds = N \|h\|_p \frac{1}{t^\theta}. \end{aligned}$$

This proves that

$$(7.5) \quad \|e^{-t} T_t (1 - \Delta)^\theta h\|_p \leq N \|h\|_p t^{-\theta},$$

which gives the first inequality in (7.3) after replacing $(1 - \Delta)^\theta h$ with h . If $\theta = 0$, then one can take $N = 1$ in the first inequality in (7.3), which follows from Minkowski's inequality. If $\theta = 1$, then we use (7.4) and the fact that $(e^{-t} T_t h)'_t = e^{-t} T_t (\Delta - 1)h$.

To prove the second inequality in (7.3) for $t \in [0, 1]$ and $\theta > 0$, it suffices to notice that $(T_t h)'_t = \Delta T_t h$ and $\Delta = [\Delta(1 - \Delta)^{-1}](1 - \Delta)^{1-\theta}(1 - \Delta)^\theta$ and use (7.5) in the following estimates

$$\begin{aligned} \|(T_t - 1)h\|_p &\leq \int_0^t \|[\Delta(1 - \Delta)^{-1}](1 - \Delta)^{1-\theta} T_s [(1 - \Delta)^\theta h]\|_p ds \\ &\leq N \int_0^t \|(1 - \Delta)^{1-\theta} T_s [(1 - \Delta)^\theta h]\|_p ds \leq N \int_0^t s^{\theta-1} ds \|h\|_{2\theta,p} = N t^\theta \|h\|_{2\theta,p}. \end{aligned}$$

For $t \geq 1$ or $\theta = 0$, the second inequality in (7.3) is trivial since $\|h\|_p \leq \|h\|_{2\theta}$ and $t^0 = 1$. The lemma is proved.

LEMMA 7.4. *Let $\alpha p > 1$ and $p \geq 1$. Then, for any continuous L_p -valued function $h(t)$, and $s \leq t$, we have*

$$\begin{aligned} (7.6) \quad \|h(t) - h(s)\|_p^p &\leq N(\alpha, p)(t-s)^{\alpha p-1} \int_s^t \int_s^t I_{r_2 > r_1} \frac{\|h(r_2) - h(r_1)\|_p^p}{|r_2 - r_1|^{1+\alpha p}} dr_1 dr_2 \\ &= N(\alpha, p)(t-s)^{\alpha p-1} \int_0^{t-s} \frac{d\gamma}{\gamma^{1+\alpha p}} \int_s^{t-\gamma} \|h(r+\gamma) - h(r)\|_p^p dr \quad (\frac{0}{0} := 0). \end{aligned}$$

This is one of embedding theorems for Slobodetskii's spaces (see, for instance, [33]). It is to be said that embedding theorems were applied for studying continuity properties of random processes for quite long time; see, for instance, [1], [9], [11], [19], [35]. The following consequence of (7.6),

$$(7.7) \quad E \sup_{0 \leq s < t \leq T} \frac{\|h(t) - h(s)\|_p^p}{(t-s)^{\alpha p-1}} \leq N(\alpha, p) \int_0^T \int_0^T I_{r_2 > r_1} \frac{E\|h(r_2) - h(r_1)\|_p^p}{|r_2 - r_1|^{1+\alpha p}} dr_1 dr_2,$$

can be used wherever one uses Kolmogorov's continuity criterion (see, [1]). By the way, embedding theorems are (for the most part) known for multi-dimensional case as well, and one can use them for studying random fields (cf. [35]).

In [19], equation (7.7) with $T = \pi$ was used to prove that Fourier's series

$$\frac{1}{\sqrt{\pi}} t \eta_0 + \sqrt{\frac{2}{\pi}} \sum_{n=1}^{\infty} \eta_n \frac{1}{n} \sin nt,$$

where η_n are $N(0, 1)$ i.i.d. random variables, converges uniformly on $[0, \pi]$. This is the way the Wiener process is introduced in [19]. Estimate (7.7) also gives a

very good modulus of continuity for the Wiener process. Of course, an advantage of normal centered variables is that $(E|\xi|^p)^{1/p} = N(p)(E|\xi|^2)^{1/2}$. Instead of this, in our applications here we use Burkholder–Davis–Gundy inequalities. By the way, a very short proof of (7.6) may be found in [19]. Finally, notice that the space L_p in (7.6) and (7.7) can be replaced with any Banach space.

Proof of Theorem 7.2. Take $u \in \mathcal{H}_p^n(\tau)$, and define $f = (\mathbb{D}u - \Delta u)I_{t \leq \tau}$ and $g = (\mathbb{S}u)I_{t \leq \tau}$. Notice that the function u on $[0, \tau]$ satisfies the equation

$$(7.8) \quad dv = (\Delta v + f) dt + g^k dw_t^k.$$

By Theorem 5.1, equation (7.8) on $[0, T]$ with initial condition $v(0) = u(0)$ has a unique solution $v \in \mathcal{H}_p^n(T)$. The difference $u - v$ satisfies the heat equation on $[0, \tau]$ with zero initial condition. It follows that $u(t, \cdot) = v(t, \cdot)$ on $[0, \tau]$, and by Theorem 5.1,

$$\|v\|_{\mathcal{H}_p^n(T)} \leq N(\|f\|_{\mathbb{H}_p^{n-2}(\tau)} + \|g\|_{\mathbb{H}_p^{n-1}(\tau, t_2)} + (E\|u(0)\|_{n-2/p, p}^p)^{1/p}) \leq N\|u\|_{\mathcal{H}_p^n(\tau)}.$$

The inequality $\|v\|_{\mathcal{H}_p^n(T)} \leq N\|u\|_{\mathcal{H}_p^n(\tau)}$ and the equality $u(t, \cdot) = v(t, \cdot)$ on $[0, \tau]$ show that we only need to prove the theorem for v, T in place of u, τ . In particular, we may and will assume that $\tau = T$. Also, observe that

$$\|f\|_{\mathbb{H}_p^{n-2}(T)} + \|g\|_{\mathbb{H}_p^{n-1}(T, t_2)} + (E\|u(0)\|_{n-2/p, p}^p)^{1/p} \leq N\|u\|_{\mathcal{H}_p^n(T)}.$$

This allows us to concentrate only on solutions of equations (7.8) on $[0, T]$ and to prove our assertions with $\|(f, g)\|_{\mathcal{F}_p^{n-2}(T)} + (E\|u(0)\|_{n-2/p, p}^p)^{1/p}$ in place of $\|u\|_{\mathcal{H}_p^n(T)}$. Therefore, below we take $\tau = T$ and take the function $u \in \mathcal{H}_p^n(T)$ as a solution of (7.8). As in the proof of Theorem 4.2, we may and will assume that f and g are as in (4.7) and $u(0) \in C_0^\infty$.

After these preparations, we are going to prove assertion (i). By Remark 3.4, without loss of generality we assume that $n = 2\beta$. Denote

$$u_1(t) = T_t u_0 + \int_0^t T_{t-s} f(s) ds.$$

It is easy to see that

$$u_1(r + \gamma) - u_1(r) = (T_\gamma - 1)u_1(r) + \int_0^\gamma T_{\gamma-\rho} f(r + \rho) d\rho.$$

Therefore, $E\|u_1(r + \gamma) - u_1(r)\|_p^p \leq N(A_1(r, \gamma) + B_1(r, \gamma))$, where

$$A_1(r, \gamma) = E\|(T_\gamma - 1)u_1(r)\|_p^p, \quad B_1(r, \gamma) = E\|\int_0^\gamma T_{\gamma-\rho} f(r + \rho) d\rho\|_p^p,$$

and, by Lemma 7.4,

$$E\|u_1(t) - u_1(s)\|_p^p \leq N(t-s)^{\alpha p-1}(I_1(t, s) + J_1(t, s)),$$

with

$$I_1(t, s) = \int_0^{t-s} \frac{d\gamma}{\gamma^{1+\alpha p}} \int_s^{t-\gamma} A_1(r, \gamma) dr, \quad J_1(t, s) = \int_0^{t-s} \frac{d\gamma}{\gamma^{1+\alpha p}} \int_s^{t-\gamma} B_1(r, \gamma) dr.$$

By using Hölder's inequality and Lemma 7.3 and observing that $(\beta - 1)q > -1$ for $q = p/(p - 1)$, we get (remember $n = 2\beta$)

$$B_1(r, \gamma) = E \int_{\mathbb{R}^d} \left| \int_0^\gamma \rho^{\beta-1} \rho^{1-\beta} T_\rho f(r + \gamma - \rho) d\rho \right|^p dx$$

$$\begin{aligned} &\leq \left(\int_0^\gamma \rho^{(\beta-1)q} d\rho \right)^{p/q} E \int_0^\gamma \rho^{(1-\beta)p} \int_{\mathbb{R}^d} |T_\rho f(r + \gamma - \rho)|^p dx d\rho \\ &\leq N \gamma^{\beta p - 1} E \int_0^\gamma \|f(r + \gamma - \rho)\|_{n-2,p}^p d\rho = N \gamma^{\beta p - 1} E \int_0^\gamma \|f(r + \rho)\|_{n-2,p}^p d\rho. \end{aligned}$$

This and the inequality $\alpha < \beta$ implies that

$$\begin{aligned} J_1(t, s) &\leq N \int_0^{t-s} \frac{d\gamma}{\gamma^{2+(\alpha-\beta)p}} \int_0^\gamma d\rho E \int_s^{t-\gamma} \|f(r + \rho)\|_{n-2,p}^p dr \\ &\leq N \int_0^{t-s} \frac{d\gamma}{\gamma^{2+(\alpha-\beta)p}} \int_0^\gamma d\rho E \int_0^t \|f(r)\|_{n-2,p}^p dr \\ &= N(t-s)^{(\beta-\alpha)p} E \int_0^t \|f(r)\|_{n-2,p}^p dr. \end{aligned}$$

To estimate I_1 , we use Lemma 7.3 and Theorem 2.1 to obtain

$$\begin{aligned} A_1(r, \gamma) &\leq N \gamma^{\beta p} E \|u_1(r)\|_{n,p}^p \leq N \gamma^{\beta p} \{E \|u(0)\|_{n-2/p,p}^p + E \int_0^r \|f(s)\|_{n-2,p}^p ds\} \\ I_1(t, s) &\leq \int_0^{t-s} \frac{d\gamma}{\gamma^{1+\alpha p}} \int_0^t A_1(r, \gamma) dr \\ &\leq N(t-s)^{(\beta-\alpha)p} \{E \|u(0)\|_{n-2/p,p}^p + E \int_0^t \|f(r)\|_{n-2,p}^p dr\}. \end{aligned}$$

For $u_2 := u - u_1$ we have

$$\begin{aligned} u_2(r + \gamma) - u_2(r) &= (T_\gamma - 1)u_2(r) + \int_r^{r+\gamma} T_{r+\gamma-\rho} g^k(\rho) dw_\rho^k, \\ E \|u_2(r + \gamma) - u_2(r)\|_p^p &\leq N(A_2(r, \gamma) + B_2(r, \gamma)), \end{aligned}$$

where

$$A_2(r, \gamma) = E \|(T_\gamma - 1)u_2(r)\|_p^p, \quad B_2(r, \gamma) = E \left\| \int_r^{r+\gamma} T_{r+\gamma-\rho} g^k(\rho) dw_\rho^k \right\|_p^p.$$

By Lemma 7.4,

$$E \|u_2(t) - u_2(s)\|_p^p \leq N(t-s)^{\alpha p - 1} (I_2(t, s) + J_2(t, s)),$$

with

$$I_2(t, s) = \int_0^{t-s} \frac{d\gamma}{\gamma^{1+\alpha p}} \int_s^{t-y} A_2(r, \gamma) dr, \quad J_2(t, s) = \int_0^{t-s} \frac{d\gamma}{\gamma^{1+\alpha p}} \int_s^{t-\gamma} B_2(r, \gamma) dr.$$

Here we use Lemma 7.3, the Burkholder–Davis–Gundy inequalities, and the inequality $(2\beta - 1)q > -1$, where $q = p/(p - 2)$. Then, similarly to the above calculations, we obtain

$$\begin{aligned} B_2(r, \gamma) &\leq N E \int_{\mathbb{R}^d} \left[\int_0^\gamma \rho^{2\beta-1} \rho^{1-2\beta} \|T_\rho g(r + \gamma - \rho)\|_{l_2}^2 d\rho \right]^{p/2} dx \\ &\leq N \gamma^{\beta p - 1} E \int_0^\gamma \rho^{(1-2\beta)p/2} \|T_\rho g(r + \gamma - \rho)\|_p^p d\rho \\ &\leq N \gamma^{\beta p - 1} E \int_0^\gamma \|(1 - \Delta)^{\beta-1/2} g(r + \gamma - \rho)\|_p^p d\rho, \end{aligned}$$

$$\begin{aligned}
J_2(t, s) &\leq NE \int_0^{t-s} \frac{d\gamma}{\gamma^{2+(\alpha-\beta)p}} \int_s^{t-\gamma} dr \int_0^\gamma \|g(r+\gamma-\rho)\|_{n-1,p}^p d\rho \\
&\leq NE \int_0^{t-s} \frac{d\gamma}{\gamma^{2+(\alpha-\beta)p}} \int_0^\gamma d\rho \int_0^t \|g(r)\|_{n-1,p}^p dr \\
&= N(t-s)^{(\beta-\alpha)p} E \int_0^t \|g(r)\|_{n-1,p}^p dr.
\end{aligned}$$

Finally, again by Lemma 7.3 and Theorem 4.2, we conclude

$$\begin{aligned}
I_2(t, s) &\leq NE \int_0^{t-s} \frac{d\gamma}{\gamma^{1+(\alpha-\beta)p}} \int_s^{t-\gamma} \|u_2(r)\|_{2\beta,p}^p dr \\
&\leq N(t-s)^{(\beta-\alpha)p} E \int_0^t \|u_2(r)\|_{n,p}^p dr \leq N(t-s)^{(\beta-\alpha)p} E \int_0^t \|g(r)\|_{n-1,p}^p dr.
\end{aligned}$$

Collecting all these estimates, we get (7.1) at least for $\eta = T (= \tau)$. In the general case, observe that if $u(t \wedge \eta) - u(s \wedge \eta)$ is not zero, then $s \wedge \eta = s$ and $s \leq s \wedge \eta \leq t \wedge \eta \leq t$. After this it suffices to notice that, instead of points t and s on the left in (7.6), we can obviously take any two points between them.

The proof of (7.2) goes exactly the same way, the only difference being that this time we use (7.7).

To prove assertion (ii) notice that from an interpolation theorem (Theorem 2.4.2 in [33]) we have

$$\|u\|_{m(\theta)-d/p+d/q,q} \leq N(d, p, q, m(0), m(1), \theta) \|u\|_{m(0),p}^{1-\theta} \|u\|_{m(1),p}^\theta$$

whenever $1 < p < q < \infty$, $\theta \in (0, 1)$, $m(\theta) := (1-\theta)m(0) + \theta m(1) \neq m(0)$, $m(i) \leq n$, and $u \in H_p^n$. Theorem 14.2 from [14] shows that the case $p = q$ actually need not be excluded. Note also that, under the conditions in (ii), there is a β such that $1/2 > \beta > 1/p$ and $m \leq n - 2\beta(1-\theta) - d/p + d/q = m(\theta) - d/p + d/q$, where $m(\theta) := (1-\theta)(n - 2\beta) + \theta n$. Therefore,

$$\begin{aligned}
E \left(\int_0^T \|u(t, \cdot)\|_{m,q}^{p/\theta} dt \right)^\theta &\leq E \left(\int_0^T \|u(t, \cdot)\|_{m(\theta)-d/p+d/q,q}^{p/\theta} dt \right)^\theta \\
&\leq NE \left(\int_0^T \|u(t, \cdot)\|_{n-2\beta,p}^{(1-\theta)p/\theta} \|u(t, \cdot)\|_{n,p}^p dt \right)^\theta \\
&\leq NE \sup_{t \leq T} \|u(t, \cdot)\|_{n-2\beta,p}^{(1-\theta)p} \left(\int_0^T \|u(t, \cdot)\|_{n,p}^p dt \right)^\theta.
\end{aligned}$$

To prove (ii), it remains only to apply Hölder's inequality, (3.4), and (7.2) with an α such that $\beta > \alpha > 1/p$. The theorem is proved.

8. Applications

8.1. Filtering Equation. Take $p \geq 2$ and an integer $d_1 > d$. Let (Ω, \mathcal{F}, P) be a complete probability space and let w_t be a d_1 -dimensional Wiener process on this space. Consider a d_1 -dimensional two-component process $z_t = (x_t, y_t)$ with x_t being d -dimensional and y_t $d_1 - d$ -dimensional. We assume that z_t is defined as a solution of the system

$$\begin{aligned}
(8.1) \quad dx_t &= b(t, z_t) dt + \theta(t, z_t) dw_t, \\
dy_t &= B(t, z_t) dt + \Theta(t, y_t) dw_t
\end{aligned}$$

with some initial data. The coefficients of (8.1) are assumed to be vectors or matrices of appropriate dimensions satisfying the following assumptions.

ASSUMPTION 8.1. The functions b, θ, B, Θ are Borel measurable and bounded functions of their arguments. Each of them satisfies the Lipschitz condition with respect to z with constant K . Moreover, $\theta(t, x, y)$ is continuously differentiable with respect to x (not z) and its first derivatives with respect to x^i 's are continuous in y and satisfy the Lipschitz condition, with constant K , with respect to x (not z).

ASSUMPTION 8.2. The symmetric matrix $\Theta\Theta^*$ is invertible and $\Psi := (\Theta\Theta^*)^{-1/2}$ is a bounded function (of (t, y)).

ASSUMPTION 8.3. For any $\xi \in \mathbb{R}^d$ and $z = (x, y) \in \mathbb{R}^{d_1}$ and $t > 0$,

$$|Q(t, y)\theta^*(t, z)\xi|^2 \geq \delta|\xi|^2,$$

where Q is the orthogonal projector on $\ker \Theta$. In other words

$$(\theta(1 - \Theta^*\Psi^2\Theta)\theta^*\xi, \xi) \geq \delta|\xi|^2.$$

One can easily check that this assumption is satisfied if the diffusion matrix of system (8.1) is block upper-triangular:

$$\begin{pmatrix} \theta \\ \Theta \end{pmatrix} = \begin{pmatrix} \theta_1 & \theta_2 \\ 0 & \Theta_2 \end{pmatrix}$$

and the $d \times d$ matrix θ_1 is non-degenerate so that $|\theta_1^*\xi|^2 \geq \delta|\xi|^2$.

We also make the following assumption regarding the initial data for (8.1).

ASSUMPTION 8.4. The random vectors x_0, y_0 are independent of the process w_t . The conditional distribution of x_0 given y_0 has a density, which we denote by $\pi_0(x) = \pi_0(\omega, x)$. We have $\pi_0 \in L_p(\Omega, H_p^{2-2/p})$.

Let

$$a(t, x) = (1/2)\theta\theta^*(t, x, y_t), \quad \sigma(t, x) = \theta\Theta^*\Psi(t, x, y_t),$$

$$\beta(t, x) = \Psi B(t, x, y_t), \quad \beta_t = \beta(t, x_t),$$

and let β^k be the k th coordinate of the vector β where $k = 1, \dots, d_1 - d$. Also, for a function u define

$$Lu = \sum_{i,j \leq d} (a^{ij}u)_{x^i x^j} - \sum_{i \leq d} (b^i u)_{x^i}, \quad \Lambda^k u = \beta^k u - \sum_{i \leq d} (\sigma^{ik}u)_{x^i}.$$

Finally, let $\mathcal{F}_t^y = \sigma\{y_s : s \leq t\}$.

THEOREM 8.1. Under the above assumptions there exists $\pi \in \cap_{T>0} \mathcal{H}_p^2(T)$ such that for any $t \geq 0$, the conditional distribution of x_t given \mathcal{F}_t^y has a density coinciding with $\pi(t, \cdot)$ almost surely. In addition, π satisfies the equation

$$(8.2) \quad d\pi = L\pi dt + \sum_{k \leq d_1-d} [\Lambda^k \pi - \bar{\beta}_t^k \pi] \{ \Psi^k \Theta dw_t + [\beta_t^k - \bar{\beta}_t^k] dt \},$$

with initial data $\pi(0, \cdot) = \pi_0$, where the missing arguments are t, x, y_t , vector Ψ^k is the k th row of Ψ , and $\bar{\beta}_t := (\pi(t, \cdot), \beta(t, \cdot, y_t))$.

To prove the theorem, we need a lemma, but first we make some comments. Notice that a little bit strange way of writing equation (8.2), with dt in two terms, is related to the fact that

$$\bar{w}_t := \int_0^t \{\Psi\Theta dw_s + [\beta_s - \bar{\beta}_s] ds\}$$

is a $d_1 - d$ -dimensional \mathcal{F}_t^y -adapted Wiener process (the so-called *innovation process*). Here we do not need this fact.

We will only use that in (8.2) we have $\Psi^k \Theta dw_t = d\tilde{w}_t^k$, where \tilde{w}_t is a $d_1 - d$ -dimensional \mathcal{F}_t -adapted Wiener process. This is easily proved by using Levy's theorem, since it follows from the definition of Ψ that $\langle \Psi\Theta dw_t \rangle = Idt$, where I is the $d_1 - d$ unit matrix. In turn, this shows that in notation (5.2) applied to (8.2) we have

$$\alpha^{ij} = \frac{1}{2} \sum_{k \leq d_1 - d} \sigma^{ik} \sigma^{jk},$$

which, by Assumption 8.3, guarantees that Assumption 5.1 is satisfied for equation (8.3) below with a K and $\delta/2$ instead of δ . Assumption 8.1 provides certain smoothness of the coefficients of (8.3), which leads us to the conclusion that the function $\bar{\pi}$ considered in the following lemma exists by Theorem 5.1 and Remark 5.6. By the way, the requirement in Assumption 8.1 that first derivatives of $\theta(t, x, y)$ with respect to x^i 's be continuous in y , which is not needed at this point, will be needed later in some passages to the limit. It is worth noting that this requirement follows automatically from other requirements in Assumption 8.1. Indeed, one knows that if a function $f(x, y)$ is Lipschitz continuous in (x, y) and has bounded generalized second-order derivatives in x , then f_x is Hölder continuous in y with exponent $1/2$.

LEMMA 8.2. *Let $\bar{\pi}$ be a unique solution in $\cap_T \mathcal{H}_p^1(T)$ of the equation*

$$(8.3) \quad d\bar{\pi} = L\bar{\pi} dt + \sum_{k \leq d_1 - d} \Lambda^k \bar{\pi} \{\Psi^k \Theta dw_t + \beta_t^k dt\}$$

with initial data $\bar{\pi}(0, \cdot) = \pi_0$, where we use the same notation as in (8.2). Then $\bar{\pi}(t, \cdot)$ is \mathcal{F}_t^y -measurable.

Proof. Observe that now we write dt in two terms on the right for the reason that $\Psi^k \Theta dw_t + \beta_t^k dt = \Psi^k dy_t$, which makes the assertion quite natural since everything in (8.3) depends only on y_t . Further (cf. [24] or [30]), there is an \mathcal{F}_t^y -adapted measurable process $\tilde{\beta}_t = (\tilde{\beta}_t^1, \dots, \tilde{\beta}_t^{d_1-d})$ such that the process

$$(8.4) \quad \tilde{w}_t := \int_0^t \{\Psi\Theta dw_s + [\beta_s - \tilde{\beta}_s] ds\}$$

is a $d_1 - d$ -dimensional \mathcal{F}_t^y -adapted Wiener process (actually $\tilde{\beta}_t$ is a modification of $E\{\beta_t | \mathcal{F}_t^y\}$). By Itô's formula the function $\tilde{\pi}$ defined by

$$\tilde{\rho}_t \tilde{\pi}(t, x) := \bar{\pi}(t, x), \quad \tilde{\rho}_t := \exp\left\{\int_0^t \tilde{\beta}_s d\tilde{w}_s + \frac{1}{2} \int_0^t |\tilde{\beta}_s|^2 ds\right\}$$

satisfies

$$(8.5) \quad d\tilde{\pi} = L\tilde{\pi} dt + \sum_{k \leq d_1 - d} [\Lambda^k \tilde{\pi} - \tilde{\beta}_t^k \tilde{\pi}] d\tilde{w}_t^k$$

with initial data $\bar{\pi}(0, \cdot) = \pi_0$. One can consider this equation on the probability space (Ω, \mathcal{F}, P) with filtration \mathcal{F}_t^y and Wiener process \tilde{w}_t . Let us denote by $\tilde{\mathcal{H}}_p^n(\tau)$ the $\mathcal{H}_p^n(\tau)$ spaces constructed on the basis of \mathcal{F}_t^y and \tilde{w}_t . Define

$$\tau_n := \inf\{t \geq 0 : \tilde{\rho}_t \leq \frac{1}{n}\} \wedge n.$$

Observe that τ_n are \mathcal{F}_t^y -stopping times. By the theory from previous sections, equation (8.5) with initial data u_0 has a unique solution in $\tilde{\mathcal{H}}_p^1(\tau_n)$ (even in $\tilde{\mathcal{H}}_p^2(\tau_n)$). On the other hand, we can replace \tilde{w}_t in (8.5) with its value from (8.4) and get an equation in terms of the original $\mathcal{H}_p^n(\tau_n)$. This equation has a unique solution in the original $\mathcal{H}_p^1(\tau_n)$. Since $\tilde{\mathcal{H}}_p^1(\tau_n) \subset \mathcal{H}_p^1(\tau_n)$, this implies that any $\mathcal{H}_p^1(\tau_n)$ -solution belongs to $\tilde{\mathcal{H}}_p^1(\tau_n)$, and since $\tilde{\pi}$ is obviously a $\mathcal{H}_p^1(\tau_n)$ -solution, we get that $\tilde{\pi}(t \wedge \tau_n, \cdot)$ is \mathcal{F}_t^y -adapted, which, for $n \rightarrow \infty$, implies that $\tilde{\pi}(t, \cdot)$ is \mathcal{F}_t^y -adapted. Finally, $\tilde{\rho}_t$ is \mathcal{F}_t^y -adapted, and therefore $\tilde{\pi}(t, \cdot)$ is \mathcal{F}_t^y -adapted. The lemma is proved.

Proof of Theorem 8.1. We imitate some proofs from [24] or [30] omitting some simple arguments which the reader can find there. First, we know (cf. Theorems 2.1 and 2.2 of [24]) that there is an \mathcal{F}_t^y -adapted measure-valued process π_t such that, for any $\phi \in C_0^\infty$ and $t \geq 0$, we have $E\{\phi(x_t) | \mathcal{F}_t^y\} = (\pi_t, \phi)$ (a.s.). We also know that π_t satisfies the following equation for any $\phi \in C_0^\infty$

$$(8.6) \quad d(\pi(t, \cdot), \phi) = (\pi(t, \cdot), L^* \phi) dt + (\pi(t, \cdot), \Lambda^{k*} \phi - \bar{\beta}_t^k \phi) \{ \Psi^k \Theta dw_t + [\beta_t^k - \bar{\beta}_t^k] dt \},$$

where $\bar{\beta}_t = (\pi(t, \cdot), \beta(t, \cdot, y_t))$ and L^* , Λ^* are the operators formally adjoint to L , Λ .

Now, let us assume for a moment that, in addition to our hypotheses, each derivative of b, B, θ with respect to x exists and is bounded. Notice that, by Sobolev's embedding theorem, any measure on \mathbb{R}^d belongs to H_p^{-n} with $n > (p-1)d/p$. This implies that π_t is a $\cap_{T>0} \mathcal{H}_p^{-n}(T)$ -solution of (8.6). If we look at $\bar{\beta}$ as a given process, then by Theorem 5.1 and Remark 5.6 there is a *unique* solution of (8.6) in the class $\cap_{T>0} \mathcal{H}_p^{-n}(T)$ with the initial condition $\pi(0, \cdot) = \pi_0$. On the other hand, since the coefficients are infinitely differentiable (but π_0 belongs only to $L_p(\Omega, H_p^{2-2/p})$), by the same Theorem 5.1 and Remark 5.6 equation (8.6) has a solution in $\cap_{T>0} \mathcal{H}_p^2(T)$. Owing to the uniqueness, this solution is π . Thus $\pi \in \cap_{T>0} \mathcal{H}_p^2(T)$, which allows us to rewrite (8.6) as (8.2) after integrating by parts, and finishes the proof in the particular case of very smooth coefficients.

Passing to the general case notice that we only need to prove that π belongs to $\cap_{T>0} \mathcal{H}_p^2(T)$, since equation (8.2) is derived from (8.6) as above. Further, take a nonnegative function $\zeta \in C_0^\infty$ with unit integral and for $\varepsilon > 0$ define $\zeta_\varepsilon(x) = \varepsilon^{-d} \zeta(x/\varepsilon)$, $h^{(\varepsilon)}(t, x, y) = h(t, x, y) * \zeta_\varepsilon(x)$, where the convolution is taken with respect to x . Consider the process $z_t^\varepsilon = (x_t^\varepsilon, y_t^\varepsilon)$ defined as a solution of

$$(8.7) \quad \begin{aligned} dx_t^\varepsilon &= b^{(\varepsilon)}(t, z_t^\varepsilon) dt + \theta^{(\varepsilon)}(t, z_t^\varepsilon) dw_t, \\ dy_t^\varepsilon &= B^{(\varepsilon)}(t, z_t^\varepsilon) dt + \Theta(t, y_t^\varepsilon) dw_t \end{aligned}$$

with initial condition $z_0^\varepsilon = z_0$. If we denote $\mathcal{F}_t^\varepsilon = \sigma\{y_s^\varepsilon : s \leq t\}$ and $\pi^\varepsilon(t, x)$ the conditional distribution of x_t^ε given $\mathcal{F}_t^\varepsilon$, then, by the above, $\pi^\varepsilon \in \cap_{T>0} \mathcal{H}_p^2(T)$ and

π^ε satisfies the following version of (8.2)

$$(8.8) \quad d\pi^\varepsilon = L_\varepsilon \pi^\varepsilon dt + \sum_{k \leq d_1 - d} [\Lambda_{\varepsilon k} \pi^\varepsilon - \bar{\beta}_{\varepsilon t}^k \pi^\varepsilon] \{ \Psi_\varepsilon^k \Theta_\varepsilon dw_t + [\beta_{\varepsilon t}^k - \bar{\beta}_{\varepsilon t}^k] dt \},$$

with initial data $\pi^\varepsilon(0, \cdot) = \pi_0$, where $L_\varepsilon, \Lambda_\varepsilon, \Psi_\varepsilon, \Theta_\varepsilon, \beta_{\varepsilon t}$ are defined in the same way as $L, \Lambda, \Psi, \Theta, \beta_t$ with y_t^ε in place of y_t , and $\beta_{\varepsilon t} := (\pi^\varepsilon(t, \cdot), \beta_\varepsilon(t, \cdot, y_t^\varepsilon))$.

By Theorem 5.1, we have $\|\pi^\varepsilon\|_{\mathcal{H}_p^2(T)} \leq N$ for any T , where N is independent of ε . By Theorem 3.11 there is a $\pi^0 \in \cap_{T>0} \mathcal{H}_p^2(T)$ and a sequence $\varepsilon_n \rightarrow 0$ such that, for any $\xi \in L_q(\Omega)$ with $q = p/(p-1)$ and $\phi \in C_0^\infty$ and $t \geq 0$,

$$(8.9) \quad \lim_{n \rightarrow \infty} E(\pi^{\varepsilon_n}(t, \cdot), \phi) \xi = E(\pi^0(t, \cdot), \phi) \xi.$$

This equality and the fact that $(\pi^{\varepsilon_n}(t, \cdot), \phi) = E\{\phi(x_t^{\varepsilon_n}) | \mathcal{F}_t^{\varepsilon_n}\}$ (a.s.) can be used instead of equality (3.14) of [24]. As there, first for ξ of type $g(y_{t_1}, \dots, y_{t_m})$ and then for arbitrary \mathcal{F}_t^y -measurable $\xi \in L_q(\Omega)$, we get that

$$E\xi(\pi^0(t, \cdot), \phi) = E\xi E\{\phi(x_t) | \mathcal{F}_t^y\}.$$

If we knew that $\pi^0(t, \cdot)$ is \mathcal{F}_t^y -measurable, then this equality would imply that

$$(8.10) \quad (\pi^0(t, \cdot), \phi) = E\{\phi(x_t) | \mathcal{F}_t^y\} \quad (\text{a.s.})$$

which is nothing but assertion that the conditional distribution of x_t given \mathcal{F}_t^y has a density coinciding with $\pi^0(t, \cdot)$. Since $\pi^0 \in \cap_{T>0} \mathcal{H}_p^2(T)$, equality (8.10) would imply that π^0 satisfies (8.2) in the same way as at the beginning of the proof. Therefore, in the remaining part of the proof we only need to prove the \mathcal{F}_t^y -measurability of $\pi^0(t, \cdot)$.

By Itô's formula and (8.8), the function

$$(8.11) \quad \bar{\pi}^\varepsilon(t, x) := \pi^\varepsilon(t, x) \rho_t^\varepsilon,$$

$$\rho_t^\varepsilon := \exp \left\{ \int_0^t (\Theta_\varepsilon^* \Psi_\varepsilon \bar{\beta}_{\varepsilon s}, dw_s) + \int_0^t (\bar{\beta}_{\varepsilon s}, \beta_{\varepsilon s}) ds - \frac{1}{2} \int_0^t |\bar{\beta}_{\varepsilon s}|^2 ds \right\}$$

satisfies the linear equation

$$d\bar{\pi}^\varepsilon = L_\varepsilon \bar{\pi}^\varepsilon dt + \sum_{k \leq d_1 - d} \Lambda_\varepsilon^k \bar{\pi}^\varepsilon \{ \Psi_\varepsilon^k \Theta_\varepsilon dw_t + [\beta_{\varepsilon t}^k - \bar{\beta}_{\varepsilon t}^k] dt \},$$

with initial data $\bar{\pi}^\varepsilon(0, \cdot) = \pi_0^\varepsilon$. By using that $\sup_{t \leq T} |y_t^\varepsilon - y_t| \rightarrow 0$ in probability and by applying Theorem 5.7 with $n = -1$ (here we also use Remark 5.8 and the continuity of θ_x with respect to y), one easily proves that $\bar{\pi}^\varepsilon \rightarrow \bar{\pi}$ in $\mathcal{H}_p^1(T)$ for any T , where $\bar{\pi}$ satisfies (8.3). By Lemma 8.2 we have that $\bar{\pi}(t, \cdot)$ is \mathcal{F}_t^y -adapted.

Next, by using the fact that $\pi^\varepsilon(t, x)$ is a probability density, one gets from (8.11) that (cf. [24])

$$\rho_t^\varepsilon \rightarrow \int_{\mathbb{R}^d} \bar{\pi}(t, x) dx =: \bar{\rho}_t$$

in probability and the last term is again \mathcal{F}_t^y -adapted. The process $\bar{\rho}_t$ satisfies the equation obtained by integrating (8.3), which means that

$$(8.12) \quad d\bar{\rho}_t = \bar{\rho}_t \hat{\beta}_t dy_t,$$

where

$$\hat{\beta}_t = \int_{\mathbb{R}^d} \beta(t, x) \bar{\pi}(t, x) dx \left(\int_{\mathbb{R}^d} \bar{\pi}(t, x) dx \right)^{-1} \quad (0 \cdot 0^{-1} := 0).$$

It follows from (8.12) and the boundedness of β and $\hat{\beta}_t$ that $\bar{\rho}_t \neq 0$ with probability one.

Thus, for any $\phi \in C_0^\infty$,

$$(\pi^\varepsilon(t, \cdot) \rho_t^\varepsilon, \phi) \rightarrow (\bar{\pi}(t, \cdot), \phi),$$

$$(\pi^\varepsilon(t, \cdot), \phi) = \frac{(\pi^\varepsilon(t, \cdot) \rho_t^\varepsilon, \phi)}{\rho_t^\varepsilon} \rightarrow (\bar{\pi}(t, \cdot), \phi) (\int_{\mathbb{R}^d} \bar{\pi}(t, x) dx)^{-1}$$

in probability. This implies that

$$(\pi^0(t, \cdot), \phi) = (\bar{\pi}(t, \cdot), \phi) (\int_{\mathbb{R}^d} \bar{\pi}(t, x) dx)^{-1} \quad (\text{a.s.}),$$

where the last expression is \mathcal{F}_t^y -adapted. The theorem is proved.

REMARK 8.3. This theorem, along with embedding theorems from Sec. 7, extends and generalizes the corresponding results from [24] and [30] in the way we were talking about in the Introduction.

8.2. On the Notion of Stochastic Integral. To consider applications to equations with infinitely dimensional Wiener processes, we want to discuss the notion of stochastic integral and show that “basically” there is nothing more general than series of usual one-dimensional stochastic integrals. This will show that equations like (2.1), containing series of integrals with respect to Wiener processes, are of a quite general nature.

The first notion was introduced by Paley, Wiener, and Zygmund in [29], where the stochastic integral of a *nonrandom smooth* function $f(t)$ against a one dimensional Wiener process w_t is defined as

$$(8.13) \quad \int_0^1 f(s) dw_s = f(1)w_1 - \int_0^1 w_s f'(s) ds.$$

Then it is verified that the $L_2(\Omega)$ -norm of the stochastic integral is equal to the $L_2(0, 1)$ -norm of f , which allows one to extend the stochastic integral from smooth functions to all $f \in L_2(0, 1)$. One obtains the same integral if, instead of (8.13), one starts with

$$(8.14) \quad \int_0^1 f(s) dw_s = \sum_{i=1}^n a_i (w_{s_i} - w_{s_{i-1}})$$

for functions f such that $f(t) = a_i$ on $(s_{i-1}, s_i]$ and $0 = s_0 \leq s_1 \leq \dots \leq s_n = 1$.

The definition based on (8.14) has an advantage that it can be easily generalized to define an integral of a *nonrandom* function against a random orthogonal measure on a σ -finite measure space (X, \mathcal{X}, m) . More precisely, assume that we are given a π -system Π of subsets of X such that $\sigma(\Pi) = \mathcal{X}$, and a random (complex-valued) variable $\mu(\gamma)$ defined for each $\gamma \in \Pi$ (and perhaps not for all $\gamma \in \mathcal{X}$). Assume that $\mu(\gamma) \in L_2(\Omega)$ and $E\mu(\gamma_1)\bar{\mu}(\gamma_2) = m(\gamma_1 \cap \gamma_2)$ for each $\gamma, \gamma_1, \gamma_2 \in \Pi$. Then for functions

$$f(x) = \sum_{i=1}^n a_i I_{\gamma_i}(x),$$

where a_i s are constant and $\gamma_i \in \Pi$, one defines the stochastic integral of f against μ as

$$(8.15) \quad \int_X f(x) \mu(dx) = \sum_{i=1}^n a_i \mu(\gamma_i),$$

and, again by isometry, one extends the stochastic integral to all $f \in L_2(X, m)$. Such integrals are used in the theory of stationary processes. Surprisingly enough, as we will see, one can also say that this is the most general stochastic integral in Itô's stochastic calculus.

Another advantage of (8.14) is that one can allow f to depend on ω , and if a_i are independent of the process $w_{t+s_{i-1}} - w_{s_{i-1}}$, $t \geq 0$, then (8.14) is again an isometry between a part of $L_2(\Omega \times (0, 1])$ and a part of $L_2(\Omega)$. Closing this isometry, K. Itô defines his famous integral.

It turns out that Itô's integral is a particular case of the integral based on (8.15). To be more precise, let w_t be (as usual) a Wiener process with respect to a filtration \mathcal{F}_t , \mathcal{P} be the predictable σ -field on $\Omega \times (0, 1]$, and Π be the set of all stochastic intervals $(0, \tau]$, where τ are stopping times ≤ 1 . Then one gets Itô's integral by taking $X = \Omega \times (0, 1]$, $\mathcal{X} = \mathcal{P}$, $\mu(0, \tau] = w_\tau$ (see more about this in [19]). In the same way one defines stochastic integrals with respect to any locally square integrable martingale.

K. Itô [12] was also the first to consider integration against measure-valued processes, which is a particular case of integration against martingale measures. Let $p(t, \Gamma)$, $t \geq 0$, be a square integrable process as a function of t with independent increments in time and a random orthogonal measures as a function of Γ for any t . Define $p((s, t], \Gamma) = p(t, \Gamma) - p(s, \Gamma)$. Itô's way of introducing the integral with respect to p is to replace the expression $a_i(w_{s_i} - w_{s_{i-1}})$ in (8.14) with

$$(8.16) \quad \int_X f_i(x) p((s_{i-1}, s_i], dx) = \int_X f_i(x) p(s_i, dx) - \int_X f_i(x) p(s_{i-1}, dx),$$

where f_i are assumed to be independent of the processes $p((s_{i-1}, t], \Gamma)$, $t \geq s_{i-1}$.

More generally, for any $\gamma \in \Pi$ let a process $p(t, \gamma)$ be given, which is a square integrable martingale with respect to a given filtration $\{\mathcal{F}_t\}$. Let $\langle p(\cdot, \gamma_1), p(\cdot, \gamma_2) \rangle_t = q(t, \gamma_1 \cap \gamma_2)$, where $q(t, \cdot)$ is a σ -finite measure on (X, \mathcal{X}) for any ω, t , and $q(t, \Gamma)$ increases in t for any $\Gamma \in \mathcal{X}$ and ω . Then there is a measure $q(dt, dx)$ such that

$$q(t, \gamma) = \int_0^1 \int_\gamma q(ds, dx).$$

By following Itô's method based on (8.16), for any $\mathcal{P} \times \mathcal{X}$ -measurable $f = f(\omega, t, x)$ such that

$$\int_0^1 \int_X f^2(s, x) q(ds, dx) < \infty,$$

one defines the stochastic integral

$$(8.17) \quad \int_0^1 \int_X f(s, x) p(ds, dx).$$

This integral is also a particular case of the integral of a *nonrandom* function against a random orthogonal measure. Indeed, define $\bar{X} = \Omega \times (0, 1] \times X$, $\bar{\mathcal{X}} = \mathcal{P} \times \mathcal{X}$, and let $m(d\omega dt dx) := P(d\omega)q(dt, dx)$. Also let

$$\bar{\Pi} = \{(0, \tau] \times \gamma : Eq((0, \tau] \times \gamma) < \infty\}, \quad \mu((0, \tau] \times \gamma) = p(\tau, \gamma).$$

Then on functions

$$(8.18) \quad f = \sum_{i \leq n} a_i I_{t \leq \tau_i} I_{\gamma_i}(x),$$

where $(0, \tau_i] \times \gamma_i \in \bar{\Pi}$ and a_i are some constants, integral (8.17) equals

$$\sum_{i \leq n} a_i p(\tau_i, \gamma_i) = \sum_{i \leq n} a_i \mu((0, \tau_i] \times \gamma_i),$$

which agrees with (8.15). Finally, by functions of type (8.18) one can approximate any $\mathcal{P} \times \mathcal{X}$ -measurable function for which (8.17) can be defined.

It is worth mentioning that there are also other notions of martingale measures with respect to which one can define stochastic integration (see [35], where the martingale measures discussed above are called orthogonal martingale measures).

Even though the notion of integral of nonrandom functions with respect to random orthogonal measures is very convenient for the purpose of introducing Itô's stochastic integrals (cf. [19]), one works almost always with stochastic integrals with variable limits, and a different notation is more appropriate.

In connection with this notice that it is shown in [10] how to reduce the stochastic integral with respect to a martingale measure to a series of usual stochastic integrals. This was further used in [10] to treat *stochastic equations* containing integrals against martingale measures using *the same notation* as in the case of equations containing just usual stochastic integrals.

To be more precise, it is assumed in [10] that \mathcal{X} is countably generated and

$$q(t, \Gamma) = \int_0^t q_s(\Gamma) dV_s,$$

where V_s is a predictable increasing process and $q_s(\Gamma)$ is a measure in Γ for any s and predictable in s for any $\Gamma \in \mathcal{X}$. Then it is shown that

$$(8.19) \quad \int_0^t \int_X f(s, x) p(ds, dx) = \sum_{k=1}^{\infty} \int_0^t f_k(s) dp_s^k,$$

where

$$p_t^k = \int_0^t \int_X \eta_k(s, x) p(ds, dx), \quad f_k(s) = \int_X \eta_k(s, x) f(s, x) q_s(dx),$$

and for any ω, s the system of functions $\{\eta_k(s, \cdot)\}$ forms an orthonormal basis in $L_2(X, q_s)$.

A particular case of the stochastic integral with respect to a martingale measure is the stochastic integral with respect to the two-dimensional Brownian sheet $W(t, x)$ defined for $t \geq 0, x \in \mathbb{R}$. In this case, one takes \mathcal{F}_t so that the random variables $W(t, x)$ are \mathcal{F}_t -measurable, and defines $p((0, \tau] \times (a, b]) = W(\tau, b) - W(\tau, a)$. This integral got very popular thanks to the article [35]. One can construct $W(s, x)$ by taking independent one-dimensional Wiener processes $w_t^k, k \geq 1$, and an orthonormal basis $\{\eta_k(x), k \geq 1\}$ in $L_2(\mathbb{R})$, and letting

$$W(t, x) = \sum_{k=1}^{\infty} w_t^k \int_0^x \eta_k(y) dy \quad t \geq 0, x \in \mathbb{R}.$$

Incidentally, observe that thus defined W is a Gaussian field and

$$EW(s, y)W(t, x) = s \wedge t \sum_{k=1}^{\infty} \int_0^x \eta_k(z) dz \int_0^y \eta_k(z) dz,$$

where

$$\sum_{k=1}^{\infty} \int_0^x \eta_k(z) dz \int_0^y \eta_k(z) dz = \begin{cases} \int_{\mathbb{R}} I_{(0,x)}(z)I_{(0,y)}(z) dz = x \wedge y & \text{if } x, y \geq 0, \\ \int_{\mathbb{R}} I_{(x,0)}(z)I_{(y,0)}(z) dz = |x| \wedge |y| & \text{if } x, y \leq 0, \\ 0 & \text{otherwise.} \end{cases}$$

In this particular case (8.19) becomes

$$(8.20) \quad \int_0^t \int_{\mathbb{R}} f(s, x) W(ds, dx) = \sum_{k=1}^{\infty} \int_0^t \left\{ \int_{\mathbb{R}} \eta_k(x) f(s, x) dx \right\} dw_s^k.$$

By the way, general one-dimensional equations driven by the cylindrical space-time white noise B_t were considered in [7], where the right-hand side of (8.20) is taken by definition as $\int_0^t \langle f(s, \cdot), dB_s \rangle$. There the series was introduced from the very beginning.

The last integral we want to discuss is the integral against a Hilbert-space valued Wiener processes (see, for instance, [30]). Let H be a Hilbert space and w_t be a H -valued Wiener process with covariance operator Q . This operator is known to be nuclear. If h_k are its unit eigenvectors with nonzero eigenvalues, then $w_t^k := (h^k, w_t)(Qh^k, h^k)^{-1/2}$ are independent standard Wiener processes and, for any H -valued process f_t for which the integral $\int_0^t f_t \cdot dw_t$ is defined, the integral can be written as

$$\sum_k \int_0^t f_s^k dw_s^k,$$

where $f_s^k = (f_s, h^k)(Qh^k, h^k)^{1/2}$ (see, for instance, [30]).

8.3. Equations Driven by Space-Time White Noise. In this subsection, we consider one-dimensional equations with space-time white noise. Thus $d = 1$. Very often (see, for instance [7]) one writes these equations in the form

$$(8.21) \quad du(t, x) = [a(t, x)u''(t, x) + b(t, x)u'(t, x) + f(t, x, u(t, x))] dt + h(t, x, u(t, x)) dB_t,$$

where B_t is a cylindrical Wiener process on L_2 . There are also very many articles (see, for instance [8]) where instead of dB_t one writes $(\partial^2 W / \partial t \partial x) dt$. As we explained in Subsec. 8.2 we may as well take $\sum_k \eta_k(x) dw_t^k$, where $\{\eta_k(x), k \geq 1\}$ is an orthonormal basis in L_2 and w_t^k are independent \mathcal{F}_t -adapted one-dimensional Wiener processes. Thus, instead of (8.21), we will be considering the equation

$$(8.22) \quad \begin{aligned} du(t, x) = & [a(t, x)u''(t, x) + b(t, x)u'(t, x) + f(t, x, u(t, x))] dt \\ & + g^k(t, x, u(t, x)) dw_t^k \end{aligned}$$

on a time interval $[0, \tau]$, where $g^k := h\eta_k$ and τ is a bounded stopping time.

ASSUMPTION 8.5. The functions $a(t, x) = a(\omega, t, x)$ and $b(t, x) = b(\omega, t, x)$ are real-valued functions defined on $(0, \tau] \times \mathbb{R}$.

(i) For any ω and $t \leq \tau(\omega)$, $a(\omega, t, \cdot) \in C^{1,1}$ ($= B^2$) and $b(\omega, t, \cdot) \in C^{0,1}$ ($= B^1$) and $\|a\|_{C^{1,1}} + \|b\|_{C^{0,1}} \leq K$. Also $K \geq a \geq \delta$.

(ii) For any $x \in \mathbb{R}$, the processes a and b are predictable.

To state the next assumption, take and fix $s \leq \infty$ and finite κ, p, r such that

$$(8.23) \quad \kappa \in (0, 1/2], \quad p \geq 2r \geq 2, \quad s \leq \infty, \quad \frac{1}{r} + \frac{1}{s} = 1, \quad r < \frac{1}{1 - 2\kappa}.$$

ASSUMPTION 8.6. The functions $f(t, x, u)$ and $h(t, x, u)$ are real-valued functions on $(0, \tau] \times \mathbb{R}^2$ such that

- (i) for any x and u , the processes $f(t, x, u)$ and $h(t, x, u)$ are predictable;
- (ii) for any ω, t, x, u , and v ,

$$(8.24) \quad |f(t, x, u) - f(t, x, v)| \leq K|u - v|, \quad |h(t, x, u) - h(t, x, v)| \leq \xi(t, x)|u - v|,$$

where ξ is certain function of ω, t, x satisfying $\|\xi(t, \cdot)\|_{2s} \leq K$.

Observe that one of possibilities is $r = 1$, and then $s = \infty$ and (8.24) just means that both f and h satisfy the Lipschitz condition in u with constant K .

To fit equation (8.22) in our general scheme, we need to find an appropriate n such that the assumptions of Theorem 5.1 are satisfied. Define

$$n = -\kappa - 3/2.$$

In the following lemma, we also set

$$R(x) = \chi|x|^{-(1-2\kappa)/2} \int_0^\infty t^{-(5-2\kappa)/4} e^{-tx^2-1/(4t)} dt,$$

so that, according to (2.3) and (2.4) for right choice of the constant χ , the function $R(x)$ is the kernel of the operator $(1 - \Delta)^{(n+1)/2}$. It is easy to prove that R is infinitely differentiable everywhere except the origin, decreases exponentially fast as $|x| \rightarrow \infty$, and behaves near the origin like $|x|^{-(1-2\kappa)/2}$ if $\kappa < (1/2)$ and like $-\log|x|$ if $\kappa = 1/2$. Finally, notice that generally speaking, the notation h, ξ, g is used in the lemma for functions different from the ones introduced above.

LEMMA 8.4. Take some functions $h \in L_p$, $\xi \in L_{2s}$, and set $g^k = \xi h \eta_k$. Then $g = \{g^k\}_{k \geq 1} \in H_p^{n+1}(l_2)$ and

$$(8.25) \quad \|g\|_{n+1,p} = \|\bar{h}\|_p \leq (N\|\xi\|_{2s}\|h\|_p) \wedge (\|\xi h\|_2\|R\|_p),$$

where $N = \|R\|_{2r} < \infty$ and

$$(8.26) \quad \bar{h}(x) := \left\{ \int_{\mathbb{R}} R^2(x-y) \xi^2(y) h^2(y) dy \right\}^{1/2}.$$

In addition, if $p(1 - 2\kappa) > 2$ and $\xi = 1$, then

$$(8.27) \quad \|g\|_{n+1,p} \leq N(\kappa) \|h\|_2^{2\kappa p/(p-2)} \|h\|_p^{1-2\kappa p/(p-2)}.$$

Proof. We know that

$$(1 - \Delta)^{(n+1)/2}(\xi h \eta_k)(x) = \int_{\mathbb{R}} R(x-y) \xi(y) h(y) \eta_k(y) dy.$$

It follows that by Parseval's theorem

$$\|(1 - \Delta)^{(n+1)/2} g(x)\|_{l_2}^2 = \sum_{k=1}^{\infty} \left(\int_{\mathbb{R}} R(x-y) \xi(y) h(y) \eta_k(y) dy \right)^2$$

$$= \int_{\mathbb{R}} R^2(x-y) \xi^2(y) h^2(y) dy = \bar{h}^2(x).$$

We thus get the equality in (8.25). To prove the inequality, notice that \bar{h}^2 is a convolution and by Minkowski's inequality the L_p -norm of a convolution is less than the L_1 -norm of one function times L_p -norm of another. This immediately gives $\|\bar{h}\|_p \leq \|\xi h\|_2 \|R\|_p$. Also, by Hölder's inequality,

$$\bar{h}^2(x) \leq \|\xi\|_{2s}^2 \left(\int_{\mathbb{R}} R^{2r}(x-y) h^{2r}(y) dy \right)^{1/r},$$

which leads to the second inequality in (8.25): $\|\bar{h}\|_p \leq N \|\xi\|_{2s} \|h\|_p$, again by Minkowski's inequality and by the assumption $p \geq 2r$. The finiteness of N follows from the assumption that $r < (1 - 2\kappa)^{-1}$.

To prove (8.27), we use that $R^2(y) \leq N|y|^{2\kappa-1}$ and we minimize with respect to $\varepsilon > 0$ after the following computations:

$$\begin{aligned} \left(\int_{\mathbb{R}} \left\{ \int_{\mathbb{R}} R^2(y) h^2(x-y) dy \right\}^{p/2} dx \right)^{2/p} &\leq \left(\int_{\mathbb{R}} \left\{ \int_{\mathbb{R}} I_{|y| \leq \varepsilon} R^2(y) h^2(x-y) dy \right\}^{p/2} dx \right)^{2/p} \\ + \left(\int_{\mathbb{R}} \left\{ \int_{\mathbb{R}} I_{|y| \geq \varepsilon} R^2(y) h^2(x-y) dy \right\}^{p/2} dx \right)^{2/p} &\leq \|I_{|y| \leq \varepsilon} R^2\|_1 \|h\|_p^2 + \|I_{|y| \geq \varepsilon} R^p\|_1^{2/p} \|h\|_2^2 \\ &\leq N\varepsilon^{2\kappa} \|h\|_p^2 + N\varepsilon^{-(1-2\kappa-2/p)} \|h\|_2^2. \end{aligned}$$

The lemma is proved.

THEOREM 8.5. *Let $\kappa \in (0, 1/2)$ and $u_0 \in L_p(\Omega, \mathcal{F}_0, H_p^{(1/2)-\kappa-2/p})$. Assume that*

$$(8.28) \quad I^p(\tau) := E \int_0^\tau \{ \|f(t, \cdot, 0)\|_{-3/2-\kappa, p}^p + \|\bar{h}(t, \cdot, 0)\|_p^p \} dt < \infty,$$

where

$$\bar{h}(t, x, 0) := \left\{ \int_{\mathbb{R}} R^2(x-y) h^2(t, y, 0) dy \right\}^{1/2}.$$

Then, in the space $\mathcal{H}_p^{(1/2)-\kappa}(\tau)$, equation (8.22) with the initial condition u_0 has a unique solution u . Moreover,

$$\|u\|_{\mathcal{H}_p^{1/2-\kappa}(\tau)} \leq N \{ I(\tau) + (E \|u_0\|_{1/2-\kappa-2/p, p}^p)^{1/p} \},$$

where the constant N depends only on κ, p, δ, K , and τ .

Proof. We will apply Theorem 5.1. Its assumptions concerning a and $\sigma (\equiv 0)$ are obviously satisfied. Next, if $u \in H_p^{n+2}$, then $bu' \in H_p^{n+1} \subset H_p^n$, $u \in L_p$, $f(u) - f(0) \in L_p \subset H_p^n$. Also, by (8.28) we have $f(t, \cdot, 0) \in H_p^n$ (a.e. on $(0, \tau]$). Furthermore, by Lemma 5.2,

$$\|bu'\|_{n,p} \leq \|bu'\|_{-1,p} \leq N\|u'\|_{-1,p} \leq N\|u\|_p = N\|u\|_{n+2-((1/2)-\kappa),p},$$

$$\|f(u) - f(v)\|_{n,p} \leq \|f(u) - f(v)\|_p \leq K\|u - v\|_p.$$

We emphasize that $\|\cdot\|_p = \|\cdot\|_{n+2-((1/2)-\kappa),p}$ and $n+2-((1/2)-\kappa) < n+2$ for $\kappa < 1/2$. Consequently (see Remark 5.5), the assumptions of Theorem 5.1 concerning $bu' + f(u)$ are satisfied. To check the remaining assumptions about $g(u)$, it suffices to notice that, by Lemma 8.4, we have

$$\begin{aligned} \|g(0)(t, \cdot)\|_{n+1,p} &= \|\bar{h}(t, \cdot, 0)\|_p, \quad \|g(u)(t, \cdot) - g(v)(t, \cdot)\|_{n+1,p} \\ &\leq N\|h(t, \cdot, u(t, \cdot)) - h(t, \cdot, v(t, \cdot))\|_p \leq N\|u(t, \cdot) - v(t, \cdot)\|_p. \end{aligned}$$

The theorem is proved.

REMARK 8.6. We have obtained Theorem 8.5 by checking that all Assumptions 5.1–5.6 are satisfied for $n = -3/2 - \kappa$. After this, of course, all other results from Sec. 5 are available. For example, by the approximation theorem (Theorem 5.7) and Remark 5.9, we have $\|u - u_m\|_{\mathcal{H}_p^{n+2}(\tau)} \rightarrow 0$, where u_m is a unique solution of the following version of (5.11)

$$du(t, x) = [a(t, x)u''(t, x) + b(t, x)u'(t, x) + f(t, x, u(t, x))] dt + \sum_{k=1}^m h(t, x, u(t, x))\eta_k dw_t^k$$

with initial condition u_0 .

REMARK 8.7. Additional information about Hölder continuity properties of the solution is readily obtained from the properties of elements of $\mathcal{H}_p^n(\tau)$ listed in Theorem 7.2.

For example,

$$(8.29) \quad u \in C^{\alpha-1/p}([0, \tau], H_p^{n+2-2\beta}) \quad (\text{a.s.})$$

provided $p > 2$ and $1/2 > \beta > \alpha > 1/p$. Here $H_p^{n+2-2\beta} \subset C^\gamma$ if

$$\gamma := n + 2 - 2\beta - 1/p = 1/2 - \kappa - 2\beta - 1/p > 0.$$

Hence, if the inequality $E\|u_0\|_{1/2-\kappa-2/p,p}^p < \infty$ and (8.28) are satisfied for any $\kappa \in (0, 1/2)$ and $p \geq 2$ (say $f(t, x, 0) = h(t, x, 0) = 0$, $\xi = K$, and u_0 is a deterministic smooth function with compact support), then, after taking p large enough and κ , α , and β small, we see that u satisfies the Hölder condition in x of order $1/2 - \varepsilon$ uniformly with respect to $t \in [0, \tau]$ (a.s.) for any $\varepsilon > 0$.

On the other hand, by taking p large enough, both α and β close to $1/4$, and κ small, we get that u satisfies the Hölder condition in t of order $1/4 - \varepsilon$ uniformly with respect to $x \in \mathbb{R}$ (a.s.) for any $\varepsilon > 0$. In terms of parabolic Hölder spaces, this means that

$$u \in C^{1/4-\varepsilon, 1/2-\varepsilon}_{t,x}([0, \tau] \times \mathbb{R}) \quad (\text{a.s.}).$$

This result agrees well with [35], where the same continuity is obtained for equations with constant coefficients. Also, notice that, using standard PDE methods, one can prove interior (with respect to t) estimates which would give similar continuity of u away from $t = 0$ under weaker assumptions on u_0 . One of results which can be obtained is discussed in the following remark.

REMARK 8.8. There are smoothing properties of equations in the sense that solutions may be much smoother than the initial data. For example, assume that $f(t, x, 0) = h(t, x, 0) = 0$, $\tau = T$, where T is a constant, $r = 1$, and $s = \infty$. Let $u_0 \in L_2(\Omega, \mathcal{F}_0, H_2^{-1+\varepsilon})$ with $\varepsilon \in (0, 1/2)$, so that u_0 may be a finite measure or just a delta-function. We claim then that, for any $\varepsilon \in (0, 1/4)$,

$$u \in C^{1/4-\varepsilon, 1/2-\varepsilon}_{t,x}([\varepsilon, T] \times \mathbb{R}) \quad (\text{a.s.}).$$

Indeed, by Theorem 8.5 with $p = 2$ and $\kappa = 1/2 - \varepsilon$, there is a unique solution $u \in \mathcal{H}_2^\varepsilon(T)$ of (8.22) and

$$\|u\|_{\mathcal{H}_2^\varepsilon(T)}^2 \leq N E\|u_0\|_{-1/2-\kappa,2}^2.$$

By (3.4) we have

$$\int_0^T E\|u(t, \cdot)\|_{\varepsilon, 2}^2 dt = \|u\|_{\mathbb{H}_2^{\varepsilon}(T)}^2 \leq NE\|u_0\|_{-1/2-\kappa, 2}^2 =: I.$$

By Chebyshev's inequality, for any $\gamma \in (0, T/2)$, Lebesgue measure of the set on $(\gamma, 2\gamma)$, where $E\|u(t, \cdot)\|_{\varepsilon, 2}^2 > 2I/\gamma$, is less than $\gamma/2$. This implies that for any $\gamma \in (0, T/2)$ there is a point $t_\gamma \in (\gamma, 2\gamma)$ such that

$$u(t_\gamma, \cdot) \in L_2(\Omega, \mathcal{F}_{t_\gamma}, H_2^\varepsilon), \quad E\|u(t_\gamma, \cdot)\|_{\varepsilon, 2}^2 \leq 2I/\gamma.$$

If we now consider (8.22) after time t_γ instead of 0 and if we define the spaces $\mathcal{H}_2^{1/2-\kappa}(t_\gamma, T)$ in an obvious way, we get by Theorem 8.5 that $u \in \mathcal{H}_2^{1/2-\kappa}(t_\gamma, T)$ for any $\kappa \in (0, 1/2)$. In addition,

$$(8.30) \quad \int_{t_\gamma}^T E\|u(t, \cdot)\|_{1/2-\kappa, 2}^2 dt \leq NI/\gamma, \quad E\|u(s_\gamma, \cdot)\|_{1/2-\kappa, 2}^2 \leq NI/\gamma^2,$$

where s_γ is a certain point in $(t_\gamma + \gamma, t_\gamma + 2\gamma)$ and $\gamma \leq T/4$.

Now we can go to larger powers. Take any $p > 4$ and define $A(p, \gamma, R) = \{\omega : \|u(s_\gamma, \cdot)\|_p \leq R\}$. By Sobolev's embedding theorem, $H_2^{1/2-\kappa} \subset H_p^m$ if $m < 1/2 - \kappa$ and $-\kappa = m - 1/p$. For $m = 0$ we have $\kappa = 1/p$. Since in (8.30) we can take $\kappa = 1/p$, we get

$$P\{A(p, \gamma, R)\} = P\{\|u(s_\gamma, \cdot)\|_p \leq R\} \geq P\{\|u(s_\gamma, \cdot)\|_{1/2-\kappa, 2} \leq NR\} \geq 1 - NI\gamma^{-2}R^{-2}.$$

Next, obviously, on the set $A(p, \gamma, R)$ and the time interval (s_γ, T) , the assumptions of Theorem 8.5 are satisfied with $\kappa = 1/2 - 2/p \in (0, 1/2)$. Therefore

$$\int_{s_\gamma}^T EI_{A(p, \gamma, R)}\|u(t, \cdot)\|_{2/p, p}^p dt \leq NEI_{A(p, \gamma, R)}\|u(s_\gamma, \cdot)\|_p^p \leq NR^p,$$

$$EI_{A(p, \gamma, R)}\|u(r_\gamma^1, \cdot)\|_{2/p, p}^p \leq NR^p/\gamma,$$

for an $r_\gamma^1 \in (s_\gamma + \gamma, s_\gamma + 2\gamma)$ and $\gamma < T/6$.

For $p > 8$, we can represent $2/p$ as $1/2 - \kappa - 2/p$ with $\kappa = 1/2 - 4/p \in (0, 1/2)$. For those p , by Theorem 8.5,

$$\int_{r_\gamma^1}^T EI_{A(p, \gamma, R)}\|u(t, \cdot)\|_{4/p, p}^p dt \leq NEI_{A(p, \gamma, R)}\|u(r_\gamma^1, \cdot)\|_{2/p, p}^p \leq NR^p/\gamma,$$

$$EI_{A(p, \gamma, R)}\|u(r_\gamma^2, \cdot)\|_{4/p, p}^p \leq NR^p/\gamma^2,$$

for some $r_\gamma^2 \in (r_\gamma^1 + \gamma, r_\gamma^1 + 2\gamma) \in (0, T)$. One can keep going this way and, for $p > 4n$, where $n = 1, 2, \dots$, find r_γ^n that are close to zero if γ is small, such that $I_{A(p, \gamma, R)}u \in \mathcal{H}_p^{2n/p}(r_\gamma^n, T)$. Here p can be taken arbitrary large, r_γ^n small, and $2n/p$ can be made as close to $1/2$ as we wish. Also, the probability of the set $A(p, \gamma, R)$ can be chosen close to 1. Therefore, we obtain our claim as in Remark 8.7.

REMARK 8.9. If $\varepsilon \geq 0$ and $\kappa + 2\varepsilon < 1/2$, then in (8.21) we could take a noise like $(1 - \Delta)^\varepsilon B_t$, which is even "whiter" than B_t . This would only lead to replacing η_k in (8.22) with $(1 - \Delta)^\varepsilon \eta_k$, and, under natural additional smoothness assumptions on h , would give the assertions of Theorem 8.5 with $\kappa + 2\varepsilon$ in place of κ .

8.4. Non-Explosion for a Nonlinear Equation. Take a, b, h satisfying Assumptions 8.5 and 8.6 from Subsec. 8.3 for $\tau = \infty$, $r = 1$, $s = \infty$, and $\xi \equiv 1$ and take a bounded real-valued $\mathcal{P} \times \mathcal{B}(\mathbb{R})$ -measurable function $c(t, x) = c(\omega, t, x)$. Assume that $h(t, x, u) = 0$ for $u \leq 0$. Fix a number $\lambda \in [0, 1/2)$ and let

$$g^k(t, x, u) = h(t, x, u)u_+^\lambda \eta_k(x),$$

where, as usual, $\{\eta_k\}$ form an orthonormal basis in L_2 . Here it is convenient to assume additionally that each η_k is bounded.

The results from Subsec. 8.3 can be easily applied to prove that if the initial condition u_0 is nonnegative and, say, is nonrandom and belongs to C_0^∞ , then the equation

$$(8.31) \quad \begin{aligned} du(t, x) = & [au''(t, x) + b(t, x)u'(t, x) + c(t, x)u(t, x)] dt \\ & + g^k(t, x, u(t, x)) dw_t^k \end{aligned}$$

has a solution defined for all t in the class of functions such that $\sup_{t \leq T, x} |u(t, x)|$ is finite (a.s.) for any $T < \infty$. These facts for equation (8.31) considered on a finite space interval with $a \equiv 1$, $b \equiv c \equiv 0$, and $h(u) = u_+$ and with zero boundary data were discovered in [26] with the help of a quite different approach. By using the maximum principle, one can show that our assertion implies the result of [26].

First, let us explain why (8.31) is solvable despite the high growth of g in u . It turns out that (8.31) possesses a kind of integral or conservation law. Observe that $u \geq 0$, which follows from the maximum principle if one notices that the solution u of (8.31) also satisfies the equation with $g^k(t, x, u(t, x))$ replaced by $\nu^k(t, x)u(t, x)$, where $\nu^k(t, x) = g^k(t, x, u(t, x))u^{-1}(t, x)$ and $|\nu^k| \leq N|u|^\lambda$. Moreover, if, for instance, $a = 1, b = c = 0$, then, by integrating (8.31) formally with respect to x , one obtains that $\|u(t, \cdot)\|_1$ is a local martingale. It is nonnegative, therefore its trajectories are bounded (a.s.).

This takes care of “almost $u^{1/2}$ ” in the diffusion term. Indeed, one can rewrite (8.31) with $\xi(t, x)u\eta_k(x)$ in place of $g^k(t, x, u)$, where

$$\xi(t, x) = h(t, x, u(t, x))u^{\lambda-1}(t, x)$$

and $|\xi(t, x)| \leq |u^\lambda(t, x)|$. By the above, the latter is summable to power $2s = 1/\lambda$, and $s > 1$ (which is required in Theorem 8.5) for $\lambda < 1/2$. Therefore, u satisfies a linear equation with coefficients under control, and we get that $u(t, x)$ is a bounded continuous function on $[0, T] \times \mathbb{R}$ for any $T < \infty$ as in Remark 8.7 (by letting $p \rightarrow \infty$). The rigorous treatment below follows this idea.

By Theorem 8.5, for any $m = 1, 2, 3, \dots$, the equation

$$(8.32) \quad du_m = (au_m'' + bu_m' + cu_m) dt + g^k(u_m \wedge m) dw_t^k$$

with the initial condition $u_m(0, \cdot) = u_0$ has a unique solution $u_m \in \mathcal{H}_p^{1/2-\kappa}(T)$ for any $\kappa \in (0, 1/2)$, $p \geq 2$, and $T < \infty$. By Remark 8.7, the function $u_m(t, x)$ is continuous in (t, x) (a.s.). To proceed with the argument, we need the following lemma to be proved later.

LEMMA 8.10. *For any $T < \infty$,*

$$(8.33) \quad \lim_{R \rightarrow \infty} \sup_m P\left\{ \sup_{t \leq T, x} |u_m(t, x)| \geq R \right\} = 0.$$

Define

$$\tau_m^R = \inf\{t \geq 0 : \sup_x |u_m(t, x)| \geq R\}.$$

Observe that if $m \geq R$ (and R is an integer), then the functions u_m and u_R satisfy the same equation on $[0, \tau_m^R]$ and therefore coincide by uniqueness. In particular, $\tau_m^m \geq \tau_m^R = \tau_R^R$ (a.s.). Therefore, there is no ambiguity in the definition

$$u(t, x) = u_m(t, x) \quad \text{on } [0, \tau_m^m].$$

Of course, u satisfies (8.31) on $(0, \lim \tau_m^m)$ and $|u(t, x)| \leq m$ for $t \leq \tau_m^m$. To finish the proof of our assertions about (8.31), it only remains to notice that $\lim \tau_m^m = \infty$ (a.s.), since, by (8.33),

$$P\{\tau_m^m \leq T\} = P\{\sup_{t \leq T, x} |u_m(t, x)| \geq m\} \leq \sup_n P\{\sup_{t \leq T, x} |u_n(t, x)| \geq m\} \rightarrow 0$$

as $m \rightarrow \infty$.

Proof of Lemma 8.10. Define

$$\xi_m(t, x) = h(t, x, u_m(t, x) \wedge m)(u_m(t, x) \wedge m)_+^\lambda u_m^{-1}(t, x) \quad (0 \cdot 0^{-1} := 0)$$

and notice that ξ_m is a bounded function. It follows from (8.32) that u_m is a solution of the equation

$$(8.34) \quad dv = (av'' + bv' + cv) dt + v \xi_m \eta_k dw_t^k.$$

Obviously, Assumptions 8.5 and 8.6 from Subsec. 8.3 are satisfied for (8.34). By Remark 8.6 and by virtue of our assumption about boundedness of η_k , Theorem 5.12 is valid for (8.34). Thus, $u_m \geq 0$ (a.s.).

Next, take $\zeta_k(x)$ from Theorem 5.7, multiply (8.32) by $\zeta_k e^{-Kt}$, where $K = \sup(|a''| + |b'| + |c|)$, integrate by parts (that is, use the definition of solutions), and take expectations. Then, for any constant T and stopping time $\tau \leq T$, we obtain

$$\begin{aligned} e^{-KT} E(\zeta_k, u_m(\tau, \cdot)) &\leq E(\zeta_k, u_m(\tau, \cdot)) e^{-K\tau} = (\zeta_k, u_0) \\ &+ E \int_0^\tau (a\zeta_k'' + (2a' - b)\zeta_k' + (a'' - b' + c - K)\zeta_k, u_m) e^{-Kt} dt \\ &\leq N + NE \int_0^\tau (|\zeta_k''| + |\zeta_k'|, u_m) dt \leq N + \frac{N}{k} k^{1-1/p} E \int_0^T \|u_m(t, \cdot)\|_p dt, \\ E(\zeta_k, u_m(\tau, \cdot)) &\leq N + Nk^{-1/p} (E \int_0^T \|u_m(t, \cdot)\|_p^p dt)^{1/p} \leq N + Mk^{-1/p}, \end{aligned}$$

where the last constant N is independent of m, k , and τ and M is independent of k . Since this inequality is true for any stopping time $\tau \leq T$, with the same N and M for any number $\gamma \in (0, 1)$ (see, for instance, Theorem III.6.8 of [19]),

$$E \sup_{t \leq T} \left(\int_{\mathbb{R}} \zeta_k(x) u_m(t, x) dx \right)^\gamma \leq 1 + \gamma \frac{N + Mk^{-1/p}}{1 - \gamma}, \quad E \sup_{t \leq T} \|u_m(t, \cdot)\|_1^\gamma \leq \frac{N + 1}{1 - \gamma},$$

where the latter relation is obtained from the former one by the monotone convergence theorem. It follows that

$$(8.35) \quad P\{\sup_{t \leq T} \|u_m(t, \cdot)\|_1 > S\} \leq \frac{N}{\sqrt{S}},$$

where N is independent of m and $S > 0$.

Now fix $m, S > 0$ and define

$$\tau_m(S) = \inf\{t \geq 0 : \|u_m(t, \cdot)\|_1 \geq S\}, \quad h_m(t, x, u) = \xi_m(t, x)u.$$

Observe that $\xi_m \leq u_m^\lambda$, which yields that for $t \leq \tau_m(S)$ we have $\|\xi_m(t, \cdot)\|_{2s} \leq S^{1/2s}$ with $s = 1/(2\lambda)$. Also, u_m satisfies (8.34). By Theorem 8.5, we obtain that, for any $T < \infty$, $\|u_m\|_{\mathcal{H}_p^{1/2-\kappa}(T \wedge \tau_m(S))}$ is bounded by a constant independent of m ,

whenever $p \geq 2r = 2/(1 - 2\lambda)$ and $1/2 > \kappa > \lambda$. The embedding theorems (cf. Remark 8.7) imply that $E \sup_{t \leq T \wedge \tau_m(S), x} |u_m(t, x)|^p$ is bounded independently of m for all large p (and hence for small p as well).

Thus

$$\begin{aligned} P\{\sup_{t \leq T, x} |u_m(t, x)| \geq R\} &\leq P\{\sup_{t \leq T \wedge \tau_m(S), x} |u_m(t, x)| \geq R\} + P\{\tau_m(S) < T\} \\ &\leq \frac{N}{R} + P\{\sup_{t \leq T} ||u_m(t, \cdot)||_1 > S\} \leq \frac{N}{R} + \frac{N}{\sqrt{S}} \end{aligned}$$

with the constants N independent of m, S, R . This leads to (8.33), and the lemma is proved.

REMARK 8.11. In [24] one can find a proof that $\sup_m E||u_m||_{C([0, T] \times \mathbb{R})}^\gamma < \infty$ for any $T < \infty$, where $0 < \gamma < 1 - 2\lambda$ and γ can be chosen arbitrary close to $1 - 2\lambda$.

9. Open Problems

- In the filtering problem, we only considered the case of equation (8.1) with noise in observations depending only on t and y_t , and thus independent of the signal x_t . Almost nothing is known about filtering problems in which Θ depends on the signal x_t . Very interesting jump processes may appear.
- The above theory applies to equations in the whole space. It is known that even for zero Dirichlet boundary data the L_2 -theory is much more complicated because one needs weighted Sobolev spaces (cf. [16]) or compatibility conditions (cf. [2], [6]). The L_2 -theory in [16] was developed by considering the stochastic term as a perturbation, which gave a perfect result due to obtaining some sharp estimates which turned out to be usable. Nothing of that kind can happen for general $p > 2$, that is why very often, say in [6], it is assumed that the stochastic term is sufficiently small and some compatibility conditions hold. A first encouraging step toward developing L_p -theory in weighted Sobolev spaces in bounded domains is made in [21] (see also [23]).

By the way, the fact that the derivatives of solutions blow up near the boundary shows that one needs to be careful constructing finite-difference approximations near the boundary.

- The above theory is an L_p -theory. One may ask whether a $C^{2+\alpha}$ -theory can be constructed. By a theory we mean not only results that, for f, g^k belonging to a space F , the solution belongs to some kind of stochastic $C^{2+\alpha}$ -spaces, but also that *every* element of this stochastic space can be obtained as a solution for certain f, g^k belonging to the same F .

We only have Theorems 5.1 and 7.2 and Sobolev's embedding $H_p^n \subset C^{n-d/p}$ if $n - d/p > 0$. Thus, if f, g^k are relatively smooth, the solution belongs to $C^{2+\alpha}$. This is a useful result; for instance, it essentially covers the main result of [3] obtained in the whole space for coefficients independent of x . The weakness of such results is that they do not allow natural perturbations and cannot be extended reasonably far. For example, if one knows such results only for equations with constant coefficients, one will not be able to extend them to equations with variable highest order coefficients.

However, a “right” $C^{2+\alpha}$ theory in the above sense is developed in [27], for equations with nonrandom coefficients and with no derivatives of the unknown function in the stochastic term.

- Fully nonlinear equations and uniqueness. In Assumption 5.6 we assume that nonlinear terms in equation (5.1) are strictly subordinated to the main terms. Nothing is known if, instead of linear main operators in (5.1), one has nonlinear ones. Even for the case of linear equations we do not know how strong this subordination has to be. For instance, take the following version of one-dimensional equation (8.22)

$$du(t, x) = u''(t, x) dt + h(x)u(t, x)\eta^k(x) dw_t^k,$$

where we only know that $h \in L_2$. Is it true that if $\sup_{t \leq T, x} |u(t, x)| < \infty$ (a.s.) and $u(0, \cdot) = 0$, then $u = 0$ on $[0, T]$? The point is that in Assumption 8.6 we have to take $s = 1$, and then $r = \infty$, $p = \infty$ and even $\kappa = 1/2$ is a little bit short of satisfying the last inequality in (8.23) which is needed to provide the right subordination. The author believes that the answer to this particular question is positive, but the general question still remains.

- Superdiffusions. One can derive SPDEs for superdiffusions in multi dimensional case (see [22]). The question arises whether one can get from these equations at least some part of huge and delicate information about superprocesses, available in the literature.

Bibliography

- [1] J.M. Bismut, M. Yor: An inequality for processes which satisfy Kolmogorov’s continuity criterion. Application to continuous martingales, *J. Funct. Anal.* **51** (1983), 166–173.
- [2] Z. Brzezniak: Stochastic partial differential equations in Banach spaces, II, (preprint).
- [3] P.-L. Chow, J.-L. Jiang: Stochastic partial differential equations in Hölder spaces, *Probab. Theory Related Fields*, **99** (1994), 1–27.
- [4] G. Da Prato: Some results on linear stochastic evolution equations in Hilbert spaces by the semi-group method, *Stochastic Anal. Appl.* **1** (1983), 57–88.
- [5] G. Da Prato and J. Zabczyk: *Stochastic equations in infinite dimensions*, Cambridge University Press, New York, NY, 1992.
- [6] F. Flandoli: Dirichlet boundary value problem for stochastic parabolic equations : compatibility relations and regularity of solutions. *Stochastics and Stochastics Reports* **29** (1990), 331–357.
- [7] T. Funaki: Random motion of strings and related evolution equations, *Nagoya Math. Journal* **89** (1983) 129–193.
- [8] I. Gyöngy, D. Nualart and M. Sanz-Solé: Approximation and support theorems in modulus spaces, *Probab. Theory and Related Fields* **101** (1995) 495–509.
- [9] A.M. Garsia, E. Rodemich, and H. Rumsey, Jr.: A real variable lemma and the continuity of paths of some Gaussian processes, *Indiana Univ. Math. J.* **20** (1970) 565–578.
- [10] I. Gyöngy and N.V. Krylov: On stochastic equations with respect to semi-martingales I, *Stochastics*, **4** (1980) 1–21.
- [11] I.A. Ibragimov: Properties of sample functions for stochastic processes and embedding theorems, *Theory Probab. and its Appl.* **18** (1973) 468–480 in Russian; English translation pp. 442–453.
- [12] K. Itô: On stochastic differential equations, *Mem. Amer. Math. Soc.* **4** (1951).
- [13] C.E. Kenig, G. Ponce and L. Vega: Well-posedness and scattering results for the generalized Korteweg–de Vries equation via the contraction principle, *Comm. Pure Appl. Math.* **46** (1993) 527–620.
- [14] M.A. Krasnoselskii, E.I. Pustylnik, P.E. Sobolevskii and P.P. Zabrejko: *Integral operators in spaces of summable functions*, Nauka, Moscow, 1966 in Russian; English translation: Noordhoff International Publishing, Leyden, 1976.

- [15] N.V. Krylov: On explicit formulas for solutions of evolutionary SPDE's (a kind of introduction to the theory), *Lect. Notes in Control and Inform. Sci.* **176** (1992) 153–164.
- [16] N.V. Krylov: A W_2^n -theory of the Dirichlet problem for SPDE in general smooth domains, *Probab. Theory Relat. Fields*, **98** (1994) 389–421.
- [17] N.V. Krylov: A generalization of the Littlewood-Paley inequality and some other results related to stochastic partial differential equations, *Ulam Quarterly*, **2** (1994) 16–26, <ftp://ftp.math.ufl.edu/pub/ulam/volume2/issue4/>
- [18] N.V. Krylov: A parabolic Littlewood-Paley inequality with applications to parabolic equations, *Topological Methods in Nonlinear Analysis, Journal of the Juliusz Schauder Center*, **4** (1994) 355–364.
- [19] N.V. Krylov: *Introduction to the theory of diffusion processes* Amer. Math. Soc., Providence, RI, 1995.
- [20] N.V. Krylov: On L_p -theory of stochastic partial differential equations, *SIAM J. Math. Anal.* **27** (1996) 313–340.
- [21] N.V. Krylov: One-dimensional SPDEs with constant coefficients in the positive half axis, Submitted to a volume dedicated to R. Liptser's 60th birthday.
- [22] N.V. Krylov: On SPDEs and superdiffusions, to appear in the Annals of Probability.
- [23] N.V. Krylov, S.V. Lototsky: A Sobolev space theory of SPDEs with constant coefficients on a half line, submitted to SIAM Journal on Mathematical Analysis.
- [24] N. V. Krylov, B. L. Rozovskii: On conditional distributions of diffusion processes, *Izvestija Akademii Nauk SSSR, serija matematicheskaja*, **42** (1978), 356–378; English translation: Math. USSR Izvestija, Vol. 12, No. 2 (1978), 336–356.
- [25] N. V. Krylov, B. L. Rozovskii: On the characteristics of degenerate second order parabolic Itô equations, *Trudy seminara imeni Petrovskogo*, **8** (1982) 153–168 in Russian; English translation: J. Soviet Math., Vol. 32 (1986), 336–348.
- [26] C. Mueller: Long time existence for the heat equation with a noise term, *Probab. Theory Related Fields*, **90** (1991) 505–517.
- [27] R. Mikulevicius, B.L. Rozovskii: *Introduction to SPDEs* (forthcoming book).
- [28] D. Nualart and E. Pardoux: White noise driven quasilinear SPDEs with reflection. *Probab. Theory Related Fields*, **93** (1992) 77–89.
- [29] R.E.A.C. Paley, N. Wiener and A. Zygmund, Note on random functions, *Math. Z.* **37** (1933) 647–668.
- [30] B. L. Rozovskii: *Stochastic evolution systems*, Kluwer, Dordrecht, 1990.
- [31] O. A. Ladyzhenskaia, V. A. Solonnikov and N. N. Ural'tceva: *Linear and quasi-linear equations of parabolic type*, Nauka, Moscow, 1967 in Russian; English translation: American Math. Soc., Providence, 1968.
- [32] S. Lapic: On the first initial-boundary problem for SPDEs on domains with limited smoothness at the boundary, submitted to Potential Anal.
- [33] H. Triebel: *Theory of function spaces*, Birkhäuser Verlag, Basel–Boston–Stuttgart, 1983.
- [34] H. Triebel: *Theory of function spaces II*, Birkhäuser Verlag, Basel–Boston–Berlin, 1992.
- [35] J. B. Walsh: An introduction to stochastic partial differential equations, *Lecture Notes in Math.* **1180** (1986) 266–439.

CHAPTER 6

Martingale Problems for Stochastic PDE's

R. MIKULEVICIUS and B. L. ROZOVSKII
 Center for Applied Mathematics Sciences
 University of Southern California
 Los Angeles, CA90089-1113

1. Introduction

In this Chapter we study nonlinear stochastic PDE's with non-smooth (in some cases singular) coefficients. The examples include stochastic Navier-Stokes equation, Langevin (stochastic quantization) equation in $P(\varphi)_2$ Euclidean quantum field theory, SPDE's for the super-Brownian motion and some related superprocesses, etc. We concentrate on existence, uniqueness, absolute continuity and singularity of distributions, and ergodicity problems for these equations.

To treat this fairly diverse set of equations in a unified way we view them as infinite-dimensional evolutional stochastic differential equations of the form

$$(1.1) \quad dX_t = A(t, x, X_t(x))dt + B(t, x, X_t(x))dW_t(x), \quad X_0(x) = u_0(x)$$

where W is a space-time white noise and A and B are partial differential operators. More specifically, we study martingale problems associated with equation (1.1).

In studying stochastic differential equations, ordinary or partial, one has to differentiate between (stochastically) strong and weak solutions. The relation (or the difference) between these two is similar to the one between a random variable and its law.

A strong solution is usually defined as a measurable function $X : (t, \{W_s, s \leq t\}) \mapsto X_t$ that satisfies equation (1.1) in a classical or generalized sense (see e.g. Ikeda, Watanabe [57]). Strong solutions exist for many classes of stochastic differential equations. These include: Ito equations with Lipschitz coefficients, Kushner's and Zakai's equations of nonlinear filtering (see e.g. [53]), quasi-linear SPDE's discussed in the previous Chapter, evolutional stochastic equations with monotone operators (Pardoux [47], Krylov, Rozovskii [51]), and many others. However, stochastic differential equations (ordinary or partial) with non-smooth coefficients often fail to allow a strong solution. Probably the simplest example of such situation was provided by Tanaka, who proved that the equation

$$(1.2) \quad dX_t = \sigma(X_t)dw_t, \quad X_0 = 0$$

where w_t is a one-dimensional Brownian motion and

$$\sigma(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases}$$

has no strong solutions. On the other hand, already in the Doob's book ([56], Ch. VI, Section 3) it was (almost explicitly) stated that if X_t is a one-dimensional random process and

$$M_t = X_t - X_0 - \int_0^t A(X_s)ds \text{ and } S_t = M_t^2 - \int_0^t B^2(X_s)ds$$

are martingales, then under certain conditions (possibly on another probability space)

$$X_t = X_0 + \int_0^t A(X_s)ds + \int_0^t B(X_s)dw_s,$$

where w_t is a Brownian motion. In the case of Tanaka's example, $M_t = X_t - X_0$ and $S_t = M_t^2 - t$ are martingales with respect to Wiener's measure \mathbf{P}_W on the space $D = (X : X \in C[0, \infty), X_0 = 0)$ furnished with the system of cylindrical σ -algebras. Thus, by Doob's result there exists a Brownian motion w'_t on (D, \mathbf{P}_W) so that $X_t = \int_0^t \sigma(X_s)dw'_s$ \mathbf{P}_W --a.s.

The problem of finding a measure on the space of trajectories so that M_t and S_t are martingales is usually referred to as martingale problem. The martingale problem approach is a generically stochastic way to relax the notion of a strong solution. Loosely speaking, in this approach one replaces the requirements on the integro-differential relations between the strong solution X and the Brownian motion W by the appropriate conditions on the probability law of the solution X , should the latter exists. If there is a measure that satisfies these conditions (a solution to the martingale problem in question), then one could construct a probability space, a white noise W' and a (not necessarily unique) stochastic process X'_t on this space that satisfy equation (1.1) with W' substituted for W . Such a process is usually referred to as a weak solution of (1.1). In general, one could not claim that X' is a measurable function $(t, \{W'_s, s \leq t\}) \mapsto X'_t$. However, X' has the same probability law as the strong solution would have if it exists, and in many cases (some would argue that always) the probability law is the only thing that matters.

Various classes of stochastic differential equations, where strong solutions do not exist or were their existence is very difficult to prove, can be handled using the martingale problem approach. For example, the martingale problem approach is very useful in situations with nonsmooth coefficients typical of many SPDE's arising in physics and other sciences. These include the majority of equations obtained as limits of branching particle systems (see Chapter 2), equations of stochastic hydrodynamics (Chapter 3), stochastic quantization equations in quantum field theory (Jona-Lasinio, Mitter [17]).

However, the idea of martingale problem approach can be traced to Doob, Stroock and Varadhan were the first who fully realized the potential of the martingale setting. The general concept of the martingale problem as well as the comprehensive development of related techniques is due to these authors, see [26].

Another approach to weak solutions of ordinary stochastic differential equations was introduced by Skorohod [50]. See also Ethier, Kurtz [14], Ikeda, Watanabe [57], Jacod, Memin [54], Zvonkin, Krylov [55] and references therein.

In applications to infinite-dimensional systems, and, in particular, to important classes of nonlinear SPDE's, the martingale problem approach was pioneered by Viot [46]. Further developments are due to Grigelionis, Mikulevicius [43], Kozlov [52], Kunita [3], Metivier [45] and others.

The outline of the rest of the Chapter is as follows. Section 1 is concerned with stochastic calculus in topological vector spaces. A theory of stochastic integration with respect to locally square integrable cylindrical martingales in topological vector spaces is developed in this section. Some particular cases of the general theory, including stochastic integration with respect to space-time white noise, Hilbert space-valued martingales, measure-valued martingales, stochastic flows etc. are also discussed in detail.

Section 2 deals with martingale problems in topological vector spaces and their applications for stochastic PDE's. This section begins with a discussion of absolute continuity and singularity of solutions to a general martingale problem. In particular, we derive the necessary and sufficient criteria for absolute continuity of solutions to martingale problems. One interesting application of this result is the characterization of all the measures absolutely continuous with respect to super-Brownian motion. We also apply the obtained criteria to establish uniqueness and existence of certain classes of SPDE's including quasi-linear SPDE's with measurable (but not necessarily continuous) "drift". The rest of the Section is concerned with extensions of Viot's compactness method and its applications in solving stochastic PDE's. Specifically, in this Section we investigate existence of weak solutions to stochastic Navier-Stokes type equations and some other quasi-linear SPDE's with polynomially growing coefficients.

In Section 3 we apply the methods developed in the previous chapters to study the $P(\varphi)_2$ stochastic quantization equation

$$(1.3) \quad \dot{X}_t(x) = -1/2(I - \Delta)X_t(x) - 2\lambda : X^3(x) : + \dot{W}_t(x), x \in U$$

where \dot{W}_t is a space-time white noise and $: X^n(x) :$ is the n th Wick product, and U is a bounded domain in \mathbf{R}^2 . . The equation of stochastic quantization was originally introduced and studied by physicists (see Parisi, Wu [38], Jona-Lasinio, Mitter [17]) with the intent to bring dynamics in $P(\varphi)_2$ Euclidean quantum field theory. The idea was to demonstrate that this equation plays the role of Langevin's equation, in that its solution is an ergodic Markov process whose unique invariant measure is the Euclidean $P(\varphi)_2$ measure. Due to the singular nature of the non-linear term, equation (1.3) is quite challenging. The predominant approach to investigate regularized versions of (1.3) (obtained by partial "coloration" of \dot{W}_t and regularization of the "drift") is based on absolute continuous change of measure arguments (see Jona-Lasinio, Mitter [17], Gatarek, Goldys [19], Hu, Kallianpur [21]). In this case a solution to the associate martingale problem is absolutely continuous with respect to the distribution of the linear version of the equation ($\lambda = 0$). In Section 3 we prove the long standing conjecture that in the non-regular case (equation (1.3)) these measures are singular. To consider both the regularized and the "white noise" cases in a unified way, we present a new approach based on the compactness method. We remark that many important issues regarding the

stochastic quantization equations remain unresolved, for example the uniqueness of weak solutions to equation (1.3) is an open problem.

2. Stochastic Integrals for Cylindrical Martingales in Topological Vector Spaces

In this section following to [29], [30], we develop a theory of stochastic integration with respect to locally square integrable cylindrical martingales in topological vector spaces.

2.1. Schwartz Pairs and Cylindrical Martingales. Let E be a topological vector space and E' its topological dual. Denote by $\langle \cdot, \cdot \rangle$ the canonical bilinear form (duality) on E, E' .

DEFINITION 2.1. The set (E, E') is called a **Schwartz pair** if:

- (i) E is locally convex and quasi-complete; the latter means that all bounded closed subsets of E are complete.
- (ii) E' is weakly separable.

As we will see later, Schwartz pairs provide a convenient setting for researching martingales and stochastic differential equations in topological vector spaces.

Denote by $\mathcal{L}^+(E)$ the space of symmetric non-negative definite forms Q from E' to E , that is, for $Q \in \mathcal{L}^+(E)$,

$$\langle y', Qy'' \rangle = \langle y'', Qy' \rangle \text{ and } \langle y', Qy' \rangle \geq 0 \quad \forall y', y'' \in E'.$$

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space with the right-continuous filtration of σ -algebras $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$. Let us denote by $\mathcal{P}(\mathbb{F})$ the \mathbb{F} -predictable σ -algebra and by $\mathcal{M}_{loc}(\mathbb{F}, \mathbf{P})$ the class of real valued local martingales on the stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$. Everywhere below, we assume that processes from $\mathcal{M}_{loc}(\mathbb{F}, \mathbf{P})$ are càdlàg. $\mathcal{M}_{loc}^2(\mathbb{F}, \mathbf{P})$ will stand for the class of locally square integrable (LSI) real valued martingales.

Suppose that we are given $M_t(y') \in \mathcal{M}_{loc}^2(\mathbb{F}, \mathbf{P})$ for each $y' \in E'$, an increasing $\mathcal{P}(\mathbb{F})$ -measurable process λ_t , and $\mathcal{P}(\mathbb{F})$ -measurable function

$$Q : [0, \infty) \times \Omega \rightarrow \mathcal{L}^+(E) \text{ (i.e. } \forall y', y'' \in E', \langle y', Qy'' \rangle \text{ is } \mathcal{P}(\mathbb{F}) - \text{measurable).}$$

DEFINITION 2.2. We say that $M_t = (M_t(y'))_{y' \in E'}$ is a locally square integrable (LSI) **cylindrical martingale** in E with the predictable quadratic variation $\int_0^t Q_s d\lambda_s$, if for all $y, y' \in E'$,

$$(2.1) \quad M_t(y')M_t(y'') - \int_0^t \langle y', Q_s y'' \rangle d\lambda_s \in \mathcal{M}_{loc}(\mathbb{F}, \mathbf{P}).$$

The function Q_s is called the **covariance operator function** of M .

Since everywhere below we consider only predictable quadratic variations, this term will be abbreviated to quadratic variation.

DEFINITION 2.3. A locally square integrable martingale $M_t(y')_{y' \in E'}$ is said to be **cylindrical Brownian motion** (CBM) in E with covariance operator $K \in \mathcal{L}^+(E)$ if its quadratic variation is tK . If $E = E'$ is a Hilbert space and K is the identity, then $M_t(y')_{y' \in E'}$ is usually referred to as the **standard cylindrical Brownian motion**

EXAMPLE 2.1. (Cylindrical Brownian Motion in the Space of Distributions). Let \mathcal{O} be an open connected set in \mathbf{R}^d . The space $C_0^\infty(\mathcal{O})$ with the inductive limit topology is denoted $\mathcal{D}(\mathcal{O})$. The dual space, written $\mathcal{D}'(\mathcal{O})$, is called the space of distributions (generalized functions) (see e.g. Reed, Simon [25]). we denote by \mathbf{I} the canonical embedding of $\mathcal{D}(\mathcal{O})$ into $\mathcal{D}'(\mathcal{O})$, i.e. the identity operator. It is readily checked that $(\mathcal{D}'(\mathcal{O}), \mathcal{D}(\mathcal{O}))$ is a Schwartz pair. The LSI cylindrical martingale $(W_t(f))_{f \in \mathcal{D}(\mathcal{O})}$ in $\mathcal{D}(\mathcal{O})$ with the quadratic variation $\int_0^t \mathbf{I} ds = t\mathbf{I}$ is the standard CBM in the space of distributions.

EXAMPLE 2.2. (Cylindrical Brownian Motions in Hilbert spaces). Suppose that H is a real separable Hilbert space and H' is identified with H . The norm and the inner product in H will be denoted by $|\cdot|$ and (\cdot, \cdot) , respectively. Let (E, E') be a Schwartz pair. Assume that $E' \subset H \subset E$ and the embeddings are continuous. Let $Q \in \mathcal{L}^+(H)$ be nuclear. It is a standard fact that there exists an orthonormal basis $(e_i)_{i \in \mathbb{N}}$ in H compiled of eigenvectors of Q . In addition we assume that all $e_i \in E'$.

Obviously, given H and a nuclear covariance operator Q on H , one can always construct a Schwartz pair (E, E') with the aforementioned properties, the simplest example is $E = E' = H$. A more important example that fits into the above setting is related to the countable Hilbert space generated by the Hilbert space $H = L^2(U)$, and the operator $Q = (I - \Delta)^{-\alpha}$, were U is a bounded smooth domain in \mathbf{R}^d , Δ is the Dirichlet Laplacian on G (see Appendix), and $\alpha > d/2$. In this case $E = \cap_s \mathcal{H}^s$, $E' = \cup_{s \geq 0} (\mathcal{H}^s)'$ where for $s \geq 0$, $\mathcal{H}^s = \{f \in H : |Q^{-s/2\alpha} f| < \infty\}$ and \mathcal{H}^{-s} is the completion of H in the norm $|f|_{-s} = |Q^{s/2\alpha} f|$. For more detail see Appendix.

Let $(W_t(y))_{y \in H}$ be a standard CBM in H with the identity covariance operator I . Then obviously $(W_t(e_i))_{i \in \mathbb{N}}$ is a sequence of standard independent Brownian motions. Write $W_t^N = \sum_{i=1}^N W_t(e_i)e_i$. Since $E|W_t^N|^2 = Nt$, we can not expect that W^N converges to an H -valued stochastic process. On the other hand it is well known (see e.g. Gihman, Skorohod [6]) that one can extend H in such a way that W^N converges strongly to a cylindrical Brownian motion in this space. This extension is not unique. Below we outline a simple and convenient extension of this type. For the sake of simplicity we will assume that Q is strictly positive.

Let H_- be the closure of H in the norm $|f|_{-1} = |Q^{1/2}f|$. Write $H_+ = \{f \in H : |f|_1 = |Q^{-1/2}f| < \infty\}$. Obviously, H_- is a separable Hilbert space under the norm $|\cdot|_{-1}$ and H_+ is its dual with respect to the scalar product in H , i.e. for $y \in H_-$, $y' \in H_+$, $\langle y, y' \rangle_{H_-, H_+} = (Q^{\frac{1}{2}}y, Q^{-\frac{1}{2}}y')$. It is readily checked that (H_-, H_+) is a Schwartz pair and $H_+ \subset H \subset H_-$ is a Gelfand triple.

Since $Qe_i = \lambda_i e_i$ and $\text{trace } Q = \sum_{i=1}^\infty \lambda_i < \infty$, we have

$$\begin{aligned} \mathbf{P} \left| \sup_{s \leq t} |W_t^{n+m} - W_t^m|_{-1} > \varepsilon \right| &= \\ \mathbf{P} \left| \sup_{s \leq t} \left| \sum_{i=m+1}^{m+n} W_t(e_i) \lambda_i^{\frac{1}{2}} e_i \right| > \varepsilon \right| &\\ \leq \varepsilon^{-2} \sum_{i=m}^{m+n} \lambda_i t &\rightarrow 0 \text{ as } n, m \rightarrow \infty \end{aligned}$$

It is readily checked now that the limit, $W_t = \sum_{i=1}^\infty W_t(e_i)e_i$ is an H_- -valued continuous square integrable martingale with quadratic characteristic Qt . Moreover,

$Q^{\frac{1}{2}}W_t = \sum_{i=1}^{\infty} W_t(e_i)Q^{\frac{1}{2}}e_i$ is an H -valued continuous square integrable martingale with quadratic characteristic Qt . For $y \in H$, set $Q^{\frac{1}{2}}W_t(y) = (Q^{\frac{1}{2}}W_t, y) = \sum_{i=1}^{\infty} W_t(e_i)(Q^{\frac{1}{2}}e_i, y)$. It is readily checked that $(Q^{\frac{1}{2}}W_t(y))_{y \in H}$ is a cylindrical Brownian motion in H with covariance operator Q . This implies that $(Q^{\frac{1}{2}}W_t(y))_{y \in E'}$ is a cylindrical Brownian motion in E with covariance operator Q . In fact, the inverse is also true. Specifically, we claim the following.

PROPOSITION 2.1. *If $(W_t^Q(y))_{y \in E'}$ is a CBM in E with covariance operator Q , then there is a standard CBM $(W_t^I(y))_{y \in H}$ in H with the identity covariance operator I such that for every $y \in E'$, $W_t^Q(y) = (Q^{\frac{1}{2}}W_t^I, y)$ for all $y \in E'$, and $((Q^{\frac{1}{2}}W_t^I, y))_{y \in H}$ is a CBM in H with covariance operator Q .*

Proof:

For all i, j ,

$$W_t^Q(e_i)W_t^Q(e_j) - t(Qe_i, e_j) = W_t^Q(e_i)W_t^Q(e_j) - \delta_{i,j}t\lambda_i \in \mathcal{M}_{loc}^c(\mathbb{F}, \mathbf{P}).$$

This yields that $(\lambda_i^{-\frac{1}{2}}W_t^Q(e_i))_{i \in \mathbb{N}}$ is a sequence of independent standard one-dimensional Brownian motions. Write $W_t^I = \sum_{i=1}^{\infty} \lambda_i^{-\frac{1}{2}}W_t^Q(e_i)e_i$. We already know that

$$Q^{\frac{1}{2}}W_t^I = \sum_{i=1}^{\infty} \lambda_i^{-\frac{1}{2}}W_t^Q(e_i)Q^{\frac{1}{2}}e_i = \sum_{i=1}^{\infty} W_t^Q(e_i)e_i$$

is an H -valued continuous square integrable martingale with quadratic characteristic Qt and $((Q^{\frac{1}{2}}W_t^I, y))_{y \in H}$ is a CBM in H with covariance operator Q . This yields, of course, that $((Q^{\frac{1}{2}}W_t^I, y))_{y \in E'}$ is a CBM in E with covariance operator Q . Thus, for all $y, y' \in E'$,

$$(Q^{\frac{1}{2}}W_t^I, y)(Q^{\frac{1}{2}}W_t^I, y') - (Qy, y')t \in \mathcal{M}_{loc}^c(\mathbb{F}, \mathbf{P}),$$

and we also have

$$W_t^Q(y)W_t^Q(y') - (Qy, y')t \in \mathcal{M}_{loc}^c(\mathbb{F}, \mathbf{P})$$

Thus for every $y \in E$, $W_t^Q(y) = (Q^{\frac{1}{2}}W_t^I, y)$ \mathbf{P} -a.s. for all t . ■

EXAMPLE 2.3. (Cylindrical Martingales in Hilbert Spaces). In fact, a more general result holds. Let $(M_t(y))_{y \in E}$ be a LSI martingale in H with quadratic variation $\int_0^t Q_s d\lambda_s$, and there exists a sequence of stopping times $T_n \uparrow \infty$ such that for every $t < \infty$ and n , $\sup_{s \leq t} \sum_{i=1}^{\infty} E|M_{t \wedge T_n}(e_i)|^2 < \infty$. (We employ the same notation as in the previous example). Write $M_t = \sum_{i=1}^{\infty} M_t(e_i)e_i$. Obviously, M_t is an H -valued stochastic process so that $(M_t, y) \in \mathcal{M}_{loc}^2(\mathbb{F}, \mathbf{P})$ for all $y \in H$, and for all n and t , $\sup_{s \leq t} E|M_{t \wedge T_n}|^2 < \infty$.

Let us denote by $\langle M \rangle_t$ the quadratic variation of $|M_t|$, that is, $\langle M \rangle_t$ is a $\mathcal{P}(\mathbb{F})$ -measurable increasing process so that $\langle M \rangle_0 = 0$ and $|M_t|^2 - \langle M \rangle_t \in \mathcal{M}_{loc}(\mathbb{F}, \mathbf{P})$. It is well known (see e.g. Meyer [2] Rozovskii [53]) that there exists a symmetric non-negative $\mathcal{PP}(F)$ -measurable nuclear operator Q_s^M on E such that $\text{trace } Q_s^M = 1$, and for all $y', y'' \in H$,

$$(M_t, y')(M_t, y'') - \int_0^t \langle y', Q_s^M y'' \rangle d\langle M \rangle_s \in \mathcal{M}_{loc}(\mathbb{F}, \mathbf{P}).$$

Thus $(M_t, y)_{y \in E}$ is a LSI martingale in H with quadratic variation $\int_0^t Q_s^M d\langle M \rangle_s$. It is readily checked that $\int_0^t Q_s d\lambda_s = \int_0^t Q_s^M d\langle M \rangle_s$, \mathbf{P} -a.e. We say, M_t is an H -valued representation of $(M_t(y))_{y \in H}$. For each y we have $(M_t, y) = M_t(y)$ \mathbf{P} -a.s.

EXAMPLE 2.4. (Measure-valued Cylindrical Martingales). Let $(U, \mathcal{B}(U))$ be a countable measurable space (i.e. $\mathcal{B}(U)$ is generated by a countable subset of sets $A \in \mathcal{B}(U)$). Denote by $B_0(U)$ be the set of all bounded $\mathcal{B}(U)$ -measurable functions f such that $\text{supp } f \subset U_n$ for some n . Let E be the set of all σ -additive measures on $(U, \mathcal{B}(U))$ so that $\mu|_{U_n}$ is bounded. Assume that E is furnished with the $\sigma(E, B_0(U))$ topology, the weakest topology in which all the functions

$$\mu \rightarrow \langle \mu, f \rangle = \int_U f d\mu \text{ for each } f \in B_0(U)$$

are continuous. It is a standard fact that in the above topology of E is locally convex and $B_0(U)$ is its dual (see e.g. [25] Theorem 4.20). Since $\mathcal{B}(U)$ is generated by a countable set, $B_0(U)$ is separable. By Theorem IV 9.4 [24], E is weakly complete.

Suppose now that there exists an increasing $\mathcal{P}(\mathbb{F})$ -measurable process λ_s , a $\mathcal{P}(\mathbb{F})$ -measurable family of non-negative σ -additive measures q_s on $(U, \mathcal{B}(U))$, and an increasing sequence of $\mathcal{B}(U)$ -measurable sets $(U_n)_{n \geq 1}$ such that $\cup_n U_n = U$, and

$$\int_0^t \int_{U_n} q_s(du) d\lambda_s < \infty \quad \mathbf{P}\text{-a.s. } \forall n, t.$$

Now set

$$\langle Q_s f, g \rangle = \int_U f(u)g(u) q_s(du), \quad f, g \in B_0(U).$$

Obviously the following statement holds.

LEMMA 2.1. *The system $M_t = (M_t(f))_{f \in B_0(U)}$ defined in such a way that for every $f \in B_0(U)$, $M_t(f) \in \mathcal{M}_{loc}^2(\mathbb{F}, \mathbf{P})$, and for all $f, g \in B_0(U)$,*

$$M_t(f)M_t(g) - \int_0^t \int_U f(u)g(u) q_s(du) d\lambda_s \in \mathcal{M}_{loc}(\mathbb{F}, \mathbf{P}),$$

is a LSI (measure-valued) cylindrical martingale in E with the quadratic variation $\int_0^t Q_s d\lambda_s$.

The covariance operator function of the martingale M_t can be identified with the measure $Q_t(du \otimes dt) = q_t(du)d\lambda_t$.

EXAMPLE 2.5. (Martingale Measures) We say that $A \in \tilde{\mathcal{B}}(U)$ if for some n , $A \in \mathcal{B}(U) \cap U_n$. For $A \in \tilde{\mathcal{B}}(U)$, set $M_t(A) = M_t(1_A)$. The process $(M_t(A))_{A \in \tilde{\mathcal{B}}(U)}$ will be called the martingale measure with the quadratic variation

$$\left(\langle M(A), M(A) \rangle_t = \int_A \int_0^t q_s(du) d\lambda_s \right)_{A \in \tilde{\mathcal{B}}(U)}.$$

A martingale measure $(M_t(A))_{A \in \tilde{\mathcal{B}}(U)}$ is orthogonal (cf. Walsh [28]) in that, for any two disjoint sets A and B in $\tilde{\mathcal{B}}(U)$, the quadratic co-variation $\langle M(A), M(B) \rangle_t = 0$.

Note that $M_t(A)$ is not necessarily an additive set function on $\tilde{\mathcal{B}}(U)$, only its quadratic variation $\langle M(A), M(A) \rangle_t$ is a measure (up to evanescence).

If the measure $q_s(du)$ is non-random, then the martingale measure with the quadratic variation $\int_A \int_0^t q_s(du) ds$ is white noise on $\mathbf{R}_+ \times E$ based on the measure $Q_t(du \otimes dt) = q_t(du)dt$, that is, (see Walsh [28]) a random set function W on $\mathbf{R}_+ \times E$ with these properties:

(i) For every $t \geq 0$ and $A \in \tilde{\mathcal{B}}(U)$, $W([0, t] \times A)$ is a zero mean Gaussian random variable with covariance $t \int_A q_s(du)$.

(ii) For any two disjoint sets A and B in $\tilde{\mathcal{B}}(U)$, $W([0, t] \times A)$ and $W([0, t] \times B)$ are independent and $W([0, t] \times A \cup [0, t] \times B) = W([0, t] \times A) + W([0, t] \times B)$ \mathbf{P} -a.s.

2.2. Hilbert Subspaces and Covariance Spaces in Schwartz Pairs. One important feature of Schwartz pairs is that there exists a one-to-one (linear) correspondence between the elements of $\mathcal{L}^+(E)$ and a family of Hilbert subspaces $(H_K)_{K \in \mathcal{L}^+(E)}$ continuously embedded into E (see [10]). In this Section we will use this fact to construct the family of Hilbert subspaces $(H_{Q_s})_{s \in \mathbf{R}_+}$ corresponding to the covariance operator function Q_s of a LSI cylindrical martingale M_t in (E, E') .

To begin with we recall some results from [10]. For $K \in \mathcal{L}^+(E)$, we define an inner product in KE' by $(Ky', Ky'')_K = \langle y', Ky'' \rangle \quad \forall y', y'' \in E'$. The following statements hold true (see [10]).

PROPOSITION 2.2. (See Proposition 10 in [10]). *There exists a completion H_K of KE' with respect to the inner product $(\cdot, \cdot)_K$ such that $H_K \subset E$ and the natural embedding is continuous.*

COROLLARY 2.1. (cf. Corollary to Proposition 7 in [10]). *Let T be a countable weakly dense subset of E' . Then KT is strongly dense in H_K , i.e., H_K is a separable Hilbert space.*

If $T = \{e'_1, e'_2, \dots\}$ is a countable weakly dense subset of E' , then we obtain an orthogonal basis (e^n) in H_K by the Hilbert Schmidt orthogonalization procedure :

$$\begin{aligned} e^1 &= \begin{cases} Ke'_1 / |Ke'_1|_{H_K}, & \text{if } |Ke'_1|_{H_K} \neq 0, \\ 0, & \text{otherwise,} \end{cases} \\ &\dots \\ e^{k+1} &= \begin{cases} (Ke'_{k+1} - \sum_{i=1}^k (Ke'_{k+1}, e^i)_{H_K} e^i) / d^{k+1}, & \text{if } d^{k+1} \neq 0, \\ 0, & \text{otherwise,} \end{cases} \\ &\dots, \end{aligned}$$

where $d^{k+1} = |Ke'_{k+1} - \sum_{i=1}^k (Ke'_{k+1}, e^i)_{H_K} e^i|_{H_K}$.

REMARK 2.1. From the definition of (e^k) it follows that for each k , there exists $\tilde{e}^k \in E'$ such that

$$(2.2) \quad e^k = Ke^k, \quad k = 1, 2, \dots$$

Now we present an explicit construction of a canonical isomorphism of the dual space H'_K onto H_K .

For $F \in H'_K$, define a linear functional on E' , written $F \circ K$, by the formula $F \circ K(e') = F(Ke')$, $e' \in E'$.

LEMMA 2.2. a) If $h \in H_K$, then there exists a unique $F \in H'_K$ such that $h = F \circ K$

b) For $F \in H'_K$, $F \circ K \in H_K$. Moreover,

$$|F|_{H'_K} = |F \circ K|_{H_K}$$

$$(F \circ K, Ke')_{H_K} = F \circ K(e') \text{ for all } e' \in E',$$

$$|F \circ K|_{H_K}^2 = \sum_n F \circ K(\tilde{e}^n)^2$$

i.e. $F \rightarrow F \circ K$ is a canonical isomorphism from H'_K onto H_K .

Proof:

a) If $h \in H_K$, then for all $e' \in E'$

$$(2.3) \quad \langle h, e' \rangle = (h, Ke')_{H_K}.$$

Define $F(Ke') = \langle h, e' \rangle$. Since (2.3) holds, $h = F \circ K$ and F is a continuous linear functional on KE' with respect to the H_K -norm. Indeed,

$$|F(Ke')| \leq |h|_{H_K} |Ke'|_{H_K}.$$

Since KE' is dense in H_K , $F \in H'_K$.

b) Let $F \in H'_K$. By Riesz theorem there exists $h \in H_K$ such that for all $e' \in E'$,

$$F(Ke') = F \circ K(e') = (h, Ke')_{H_K} = \langle h, e' \rangle.$$

Thus $h = F \circ K$ and

$$|h|_{H_K}^2 = \sup_{e'} |F \circ K(e')|^2 / \langle Ke', e' \rangle = \sum_n F \circ K(\tilde{e}^n)^2 = |F|_{H'_K}.$$

The statement is proved. ■

REMARK 2.2. Since the embedding of H_K into E is continuous, a continuous linear form on E is continuous on H_K , i.e. $E' \subset H'_K$.

COROLLARY 2.2. a) If $f' \in E'$, we have $f' \circ K = Kf'$ and $|f'|_{H'_K} = |Kf'|_{H_K}$. Also,

$$(2.4) \quad f' \circ K(e') = \langle f', Ke' \rangle_{E', E} = \langle e', Kf' \rangle_{E', E} = (Ke', Kf')_{H_K} = \langle f', Ke' \rangle_{H'_K, H_K}.$$

b) the sequence (\tilde{e}^n) defined by (2.2) is an orthogonal basis in H'_K , i.e. for $F \in H'_K$ we have an expansion $F = \sum_n F \circ K(\tilde{e}^n) \tilde{e}^n$ in H'_K , and E' is a dense subspace in H'_K .

c) K on E' can be continuously extended on H'_K , and it defines a canonical isometry from H'_K onto H_K . Also, for all $F \in H'_K$, $e' \in E'$, $h \in H_K$

$$(2.5) \quad \langle e', KF \rangle_{E', E} = \langle F, Ke' \rangle_{H'_K, H_K} = (KF, Ke')_{H_K} = F \circ K(e'),$$

$$(KF, h)_{H_K} = \langle F, h \rangle_{H'_K, H_K}, \langle F, KF \rangle_{H'_K, H_K} = |KF|_{H_K}^2 = |F|_{H'_K}^2.$$

Proof:

a) Indeed, for each $e' \in E'$ we have

$$f' \circ K(e') = \langle f', Ke' \rangle = \langle e', Kf' \rangle.$$

Thus $f' \circ K = Kf'$ and obviously $|f'|_{H'_K} = |Kf'|_{H_K}$.

b) By Lemma 2.2 for each $e' \in E'$ we have $(F \circ K, Ke')_{H_K} = F \circ K(e')$. Therefore

$$F \circ K = \sum_n (F \circ K, K\tilde{e}^n)_{H_K} K\tilde{e}^n = \sum_n F \circ K(\tilde{e}^n) \tilde{e}^n \circ K,$$

i.e. $F = \sum_n F \circ K(\tilde{e}^n) \tilde{e}^n$ in H'_K . Since $\tilde{e}^n \in E'$, it follows obviously that E' is dense in H'_K .

c) By (2.4)

$$\langle e', Kf' \rangle_{E', E} = (Ke', Kf')_{H_K} = \langle f', Ke' \rangle_{H'_K, H_K}.$$

By a) a sequence $f'_n \in E'$ is Cauchy in H'_K if and only if Kf'_n is Cauchy in H_K . If $f'_n \rightarrow F$ in H'_K then (Kf'_n) converges to an element $g \in H_K$. We denote $g = KF$ and obtain (2.5) by continuity. ■

REMARK 2.3. We notice here that the set of restrictions $E'|_{H_K} = \{e'|_{H_K} : e' \in E'\}$ is isomorphic to $E'/\text{Ker } K$ and $K : E'|_{H_K} \rightarrow KE'$ is an algebraic isomorphism. It follows from the Corollary 2.2 that K can be continuously extended to canonical isomorphism from H'_K onto H_K .

Let $M_t = (M_t(y'))_{y' \in E'}$ be a locally square integrable cylindrical martingale in E with covariance operator function Q_s and quadratic variation $\int_0^t Q_s d\lambda_s$. Now we can rewrite (2.1) as

$$M_t(y') M_t(y'') - \int_0^t (Q_s y', Q_s y'')_{H_{Q_s}} d\lambda_s \in \mathcal{M}_{loc}(\mathbb{F}, \mathbf{P}).$$

We say that $(H_{Q_s})_{s \in \mathbb{R}_+} ((H'_{Q_s})_{s \in \mathbb{R}_+})$ is the family of **covariance spaces (adjoint covariance spaces)** of M .

When the context allows no room for ambiguity, H_{Q_s} and H'_{Q_s} will often be abbreviated H_s and H'_s , respectively.

Let $L(Q)$ be the set of all vector fields $f = f_s = f(s, \omega)$ such that $f_s \in H_s$ and $(f_s, Q_s y')_{H_s}$ are $\mathcal{P}(\mathbb{F})$ -measurable for each $y' \in E'$.

Let $T = \{e'_1, \dots\}$ be a countable weakly dense subset of E' . We define a sequence of E -valued $\mathcal{P}(\mathbb{F})$ -measurable functions:

$$(2.6) \quad \begin{aligned} e_s^1 &= \begin{cases} Q_s e'_1 / |Q_s e'_1|_{H_s}, & \text{if } Q_s e'_1 \neq 0 \\ 0, & \text{if } Q_s e'_1 = 0 \end{cases}, \\ &\dots \\ e_s^{k+1} &= \begin{cases} (Q_s e'_{k+1} - \sum_{i=1}^k (Q_s e'_{k+1}, e_s^i) e_s^i / d_s^{k+1}, & \text{if } d_s^{k+1} \neq 0, \\ 0, & \text{if } d_s^{k+1} = 0 \end{cases}, \\ &\dots \end{aligned}$$

where $d_s^{k+1} = |Q_s e'_{k+1} - \sum_{i=1}^k (Q_s e'_{k+1}, e_s^i) e_s^i|_{H_s}$. It follows from the definition of the sequence (e_s^n) that for each n there exists E' -valued $\mathcal{P}(\mathbb{F})$ -measurable function \tilde{e}_s^n such that

$$(2.7) \quad e_s^n = Q_s \tilde{e}_s^n.$$

Since according to Corollary 2.1, $Q_s T$ is a dense subset of H_s , then (2.6) is the Hilbert-Schmidt orthogonalization procedure. Thus (e_s^n) form a basis of H_s for each s , and we arrive at the following statement.

COROLLARY 2.3. Let $f \in L(Q)$. For each s , we have the following expansion in $H_{Q_s} = H_s$:

$$(2.8) \quad f_s = \sum_k (f_s, e_s^k)_{H_s} e_s^k, \quad \text{and } |f_s|_{H_s}^2 = \sum_k (f_s, e_s^k)_{H_s}^2,$$

and so $|f_s|_{H_s}$ is a predictable function.

If f'_s is a predictable E' -valued function, then for each s , we have the expansion in $H'_{Q_s} = H'_s$

$$f'_s = \sum_k (Q_s f'_s, e_s^k)_{H_s} \tilde{e}_s^k, \quad \text{and } |f'_s|_{H'_s}^2 = \sum_k (f'_s, Q_s \tilde{e}_s^k)_{H_s}^2.$$

We conclude this section by exploring the concrete structure of covariance and adjoint covariance spaces for several important particular cases.

EXAMPLE 2.6. Let us consider the Schwartz pair $(\mathcal{D}'(\mathcal{O}), \mathcal{D}(\mathcal{O}))$ and the CBM $(W_t(f))_{f \in \mathcal{D}(\mathcal{O})}$ introduced in Example 2.1. Obviously in this case the covariance space $H_s = L^2(\mathcal{O})$, and the adjoint covariance space, H'_s , is naturally identified with H_s .

EXAMPLE 2.7. Let E and Z be separable Hilbert spaces. We identify them with their duals. Let M_t be an E -valued LSI martingale. Then $((M_t, y))_{y \in E}$ is a LSI cylindrical martingale in E , with the quadratic variation $\int_0^t Q_s^M d\langle M, M \rangle_s$.

Assume that there is a predictable operator $\sigma = \sigma_s : Z \rightarrow E$, such that $\sigma\sigma^* = Q_s$. Let \tilde{H}_s be the closure of σ^*E in Z , and pr_s be the orthogonal projection of Z on \tilde{H}_s . (To motivate this assumption one can think about the martingale problem associated with the Ito equation $dx_t = \sigma(x_t)dw_t$ in R^d where σ is a $d \times d_1$ -matrix and w is a d_1 -dimensional Brownian motion.)

PROPOSITION 2.3. (i) The mapping $h \rightarrow \sigma_s h$ is an isometry from \tilde{H}_s onto H_s . The covariance space $H_s = \sigma_s Z$, and for $g \in Z$, $|\sigma_s g|_{H_s} = |pr_s g|_Z$.

(ii) The mapping $F \rightarrow (\sigma_s^{-1})^* F$ is an isometry from \tilde{H}_s onto H'_s . In particular, for every $F \in \tilde{H}_s$, $|F|_E = |(\sigma_s^{-1})^* F|_{H'_s}$.

Proof:

For each $y \in E$, $|Q_s y|_{H_s}^2 = (\sigma_s \sigma_s^* y, y)_E = (\sigma_s^* y, \sigma_s^* y)_Z = |\sigma_s^* y|_Z$, i.e. σ_s is an isometric map from $\sigma_s^* E$ onto $Q_s E$. By continuity, we obtain that σ_s is an isometry from \tilde{H}_s onto H_s . This implies of course that $\sigma_s \tilde{H}_s = H_s$. Moreover, for each $x \in \tilde{H}_s$, $|\sigma_s x|_{H_s}^2 = |x|_Z^2$. To complete the proof it suffices to show that $\sigma_s \tilde{H}_s = \sigma_s Z$. Let $z \in Z$ and $z \perp \tilde{H}_s$, i.e. $0 = (\sigma_s^* y, z)_Z = (y, \sigma_s z)_Y$ for each $y \in E$. Then $\sigma_s z = 0$, and (i) is proved. The second part of the statement is pretty straightforward and we leave it to the interested reader. ■

EXAMPLE 2.8. Everywhere in this example we use the notation and the assumptions of Example 2.4. Let $(M_t)_{f \in B_0(U)}$ be a LSI measure-valued cylindrical martingale in E with the covariance operator function $Q_t(du \otimes dt) = q_t(du)d\lambda_t$. It is readily checked that the covariance and the adjoint covariance spaces for $(M_t)_{f \in B_0(U)}$ can be characterized as follows:

H_s is the space of all measures μ_s on $(U, B_0(U))$ absolutely continuous with respect to q_s equipped with the norm $|\mu_s|_{H_s}^2 = \int_U \left(\frac{d\mu_s}{dq_s}\right)^2 q_s(du)$;

$$H'_s = L^2(U, B_0, q_s).$$

2.3. Ito Stochastic Integrals. Now we construct the Ito stochastic integral for the class $\tilde{L}_{loc}^2(Q)$ of all $\mathcal{P}(\mathbb{F})$ -measurable E' -valued functions f such that

$$\int_0^t \langle Q_s f_s, f_s \rangle d\lambda_s < \infty \quad \text{P-a.s. for all } t > 0.$$

We start from the set of simple functions $S_b = \{f \in \tilde{L}_{loc}^2(Q) : f = \sum_1^n f_s^k h_k, f^k$ are $\mathcal{P}(\mathbb{F})$ -measurable bounded scalar functions, $h_k \in E', k = 1, \dots, n, n \geq 1\}$. For $f = \sum_1^n f_s^k h_k \in S_b$ we define the Ito integral by

$$\mathcal{I}_t(f) = \int_0^t f_s dM_s = \sum_1^n \int_s^t f_s^k dM_s(h_k).$$

We see immediately that the map $f \rightarrow \mathcal{I}(f)$ defined on S_b (with values in $\mathcal{M}_{loc}^2(\mathbb{F}, \mathbf{P})$) is linear up to evanescence and for each $f \in S_b$

$$(2.9) \quad \langle \mathcal{I}(f) \rangle_t = \int_0^t \langle Q_s f_s, f_s \rangle d\lambda_s.$$

PROPOSITION 2.4. *The map $f \rightarrow \mathcal{I}(f)$ defined on S_b has a further extension to the set $\tilde{L}_{loc}^2(Q)$ (still denoted $f \rightarrow \mathcal{I}_t(f) = \int_0^t f_s dM_s$) such that:*

1. $\mathcal{I}(f) \in \mathcal{M}_{loc}^2(\mathbb{F}, \mathbf{P})$ and (2.9) holds;
2. $f \rightarrow \mathcal{I}(f)$ is linear up to evanescence;
3. If $f^n, f \in \tilde{L}_{loc}^2(Q)$ and for each $t > 0$ $\int_0^t \langle Q_s(f_s^n - f_s), f_s^n - f_s \rangle d\lambda_s \rightarrow 0$ in probability, then for each $t > 0$

$$\sup_{s \leq t} |\mathcal{I}_s(f^n) - \mathcal{I}_s(f)| \rightarrow 0 \text{ in probability,}$$

as $n \rightarrow \infty$.

Moreover, this extension is unique (up to evanescence).

Proof:

1⁰. Firstly, we extend the Ito integral to

$$\tilde{L}^2(Q) = \left\{ f \in \tilde{L}_{loc}^2 : E \int_0^\infty \langle Q_s f_s, f_s \rangle d\lambda_s < \infty \right\}.$$

Let $S = \{f \in \tilde{L}^2(Q) : f = \sum_1^n f_s^k h_k, f^k \mathcal{P}(\mathbb{F})\text{-measurable functions, } h_k \in E', k = 1, \dots, n, n \geq 1\}$. Fix $f_s = \sum_1^p f_s^k h_k \in S$ and define

$$g_s^n = f_s 1_{\{\max_{1 \leq k \leq p} |f_s^k| \leq n\}}.$$

Obviously $g^n \in S_b \cap \tilde{L}^2(Q)$ and by Lebesgue dominated convergence theorem

$$E \int_0^\infty \langle Q_s(f_s - g_s^n), f_s - g_s^n \rangle d\lambda_s \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Now it follows from (2.9) that

$$E \sup_t |\mathcal{I}_t(g^n) - \mathcal{I}_t(g^m)|^2 \rightarrow 0, \quad \text{as } n, m \rightarrow \infty.$$

Thus we can extend linearly \mathcal{I} to S so that for each $f \in S$, $\mathcal{I}_t(f) \in \mathcal{M}^2(\mathbb{F}, P)$ and (2.9) holds. Now fix $f \in \tilde{L}^2(Q)$. Then for almost every s , $Q_s f_s \in H_s$, and by

Corollary 2.3 (see (2.8)) $Q_s f_s = \sum_k (Q_s f_s, e_s^k)_{H_s} e_s^k$. Let $g_s^N = \sum_1^N (Q_s f_s, e_s^k)_{H_s} e_s^k$. By the Lebesgue denominated convergence theorem

$$(2.10) \quad \mathbf{E} \int_0^\infty |Q_s f_s - g_s^N|_{H_s}^2 d\lambda_s \rightarrow 0$$

as $N \rightarrow \infty$. By the definition of g_s^N and e_s^k (see (2.6)) it follows that there exists $f^N \in S$ such that $g_s^N = Q_s f_s^N$. Therefore we can write (2.10) as

$$\mathbf{E} \int_0^\infty |Q_s f_s - Q_s f_s^N|_{H_s}^2 d\lambda_s = \mathbf{E} \int_0^\infty \langle Q_s(f_s - f_s^N), f_s - f_s^N \rangle d\lambda_s \rightarrow 0, \quad \text{as } N \rightarrow \infty.$$

Thus $\mathcal{I}_t(f^N)$ is a Cauchy sequence and we can find $\mathcal{I}(f) \in M^2(\mathbb{F}, \mathbf{P})$ such that (2.9) holds and

$$\mathbf{E} \sup_t |\mathcal{I}_t(f) - \mathcal{I}_t(f^N)|^2 \rightarrow 0, \quad \text{as } N \rightarrow \infty.$$

Obviously this extension to $\tilde{L}^2(Q)$ is linear (up to evanescence) and unique by the property 3.

2°. In order to extend \mathcal{I} to $\tilde{L}_{loc}^2(Q)$ we localize. Fix $f \in \tilde{L}_{loc}^2(Q)$. Then there exists a sequence of stopping times (τ_m) such that $f_s 1_{\{s \leq \tau_m\}} \in \tilde{L}^2(Q)$ for each m , $\mathbf{P}(\tau_m < t) \xrightarrow{m \rightarrow \infty} 0$ for each $t > 0$ and we can find $\mathcal{I}(f) \in \mathcal{M}_{loc}^2(\mathbb{F}, \mathbf{P})$ such that $\mathcal{I}_{t \wedge \tau_m}(f) = \mathcal{I}_t(f 1_{\{\cdot \leq \tau_m\}})$. The properties 1,2 of the extension are obvious.

3°. Finally, we prove that 3. holds for $\mathcal{I}(f)$, $f \in \tilde{L}_{loc}^2(Q)$. Let $\int_0^t \langle Q_s(f_s^n - f_s), f_s^n - f_s \rangle d\lambda_s \rightarrow 0$ in probability as $n \rightarrow \infty$ for all $t > 0$. There exists a sequence (τ_n) of stopping times such that for each $t > 0$

$$\mathbf{E} \int_0^{\tau_n} \langle Q_s(f_s^n - f_s), f_s^n - f_s \rangle d\lambda_s + \mathbf{P}(\tau_n < t) \rightarrow 0$$

as $n \rightarrow \infty$, and for each n $\mathbf{E} \int_0^{\tau_n} \langle Q_s f_s, f_s \rangle d\lambda_s < \infty$. Thus $\mathbf{E} \sup_t |\mathcal{I}_{t \wedge \tau_n}(f_n) - \mathcal{I}_{t \wedge \tau_n}(f)|^2 \rightarrow 0$, as $n \rightarrow \infty$. Since $\mathbf{P}(\tau_n < t) \xrightarrow{n \rightarrow \infty} 0$ for each $t > 0$, we derive easily that $\sup_{s \leq t} |\mathcal{I}_s(f_n) - \mathcal{I}_s(f)| \rightarrow 0$ in probability as $n \rightarrow \infty$ for each $t > 0$. Then the statement follows. ■

2.4. Completing the Set of Integrands. Now we shall complete the set of integrands $\tilde{L}_{loc}^2(Q)$. Let $\hat{L}(Q)$ be the set of all functions $F_s = F_s(\omega)$ such that $F_s \in H'_s = H'_{Q_s}$, and $F_s \circ Q_s(e')$ is $\mathcal{P}(\mathbb{F})$ -measurable for all $e' \in E'$. By Corollary 2.2, the latter is equivalent to $\langle Q_s F_s, e' \rangle$ is $\mathcal{P}(\mathbb{F})$ -measurable for all $e' \in E'$. Recall that $L(Q)$ is the set of E -valued functions $h_s = h(s, \omega)$ such that for every s , $h_s \in H_s$ and $\langle h_s, e' \rangle$ is $\mathcal{P}(\mathbb{F})$ -measurable for all $e' \in E'$. Define

$$\hat{L}_{loc}^2(Q) = \{F \in \hat{L}(Q) : \int_0^t |F_s|_{H'_s}^2 d\lambda_s < \infty \text{ P-a.s. for all } t > 0\}.$$

and

$$L_{loc}^2(Q) = \{h \in L(Q) : \int_0^t |h_s|_{H_s}^2 d\lambda_s < \infty \text{ P-a.s. for all } t > 0\}.$$

LEMMA 2.3. a) The inclusion $\tilde{L}_{loc}^2(Q) \subset \hat{L}_{loc}^2(Q)$ holds, and for each $F \in \hat{L}_{loc}^2(Q)$ there exists a sequence $F^n \in \tilde{L}_{loc}^2(Q)$ such that for each $t > 0$

$$\int_0^t |F_s^n - F_s|_{H'_s}^2 d\lambda_s \rightarrow 0$$

in probability, as $n \rightarrow \infty$.

b) If $h \in L_{loc}^2(Q)$, then there exists a unique $F \in \hat{L}_{loc}^2(Q)$ such that $F_s \circ Q_s = h_s$,

$$|h_s|_{H_s}^2 = |F_s|_{H'_s}^2.$$

Proof:

a) We have obviously $\tilde{L}_{loc}^2(Q) \subset \hat{L}_{loc}^2(Q)$ by Corollary 2.2. Let (e_s^n) and (\tilde{e}_s^n) be sequences defined by (2.6) and (2.7). Then (e_s^n) is a basis in $H_s = H_{Q_s}$, (\tilde{e}_s^n) is a sequence of E' -valued predictable processes such that $Q_s \tilde{e}_s^n = e_s^n$, and (\tilde{e}_s^n) is an orthogonal basis of H'_s by Corollary 2.2. Let

$$F_s^n = \sum_{k=1}^n F \circ Q_s(\tilde{e}_s^k) \tilde{e}_s^k.$$

Then the statement follows immediately by Corollary 2.2.

b) For all $e' \in E'$

$$\langle h_s, e' \rangle = (h_s, Q_s e')_{H_s}.$$

So $F_s(Q_s e') = \langle h_s, e' \rangle$ is uniquely and correctly defined. Also,

$$|F_s|_{H'_s} = |h_s|_{H_s},$$

and $F \in \hat{L}_{loc}^2(Q)$ since $h \in L_{loc}^2(Q)$. ■

Now we show that there exists a natural extension of the Ito integral.

PROPOSITION 2.5. The map $F \rightarrow \mathcal{I}_t(F)$ defined on $\tilde{L}_{loc}^2(Q)$ has a further extension to the set $\hat{L}_{loc}^2(Q)$ (still denoted $F \rightarrow \mathcal{I}_t(F) = \int_0^t F_s dM_s$) such that:

1. $\mathcal{I}(F) \in \mathcal{M}_{loc}^2(\mathbb{F}, \mathbf{P})$ and $\langle \mathcal{I}(F) \rangle_t = \int_0^t |F_s|_{H'_s}^2 d\lambda_s = \int_0^t |F_s \circ Q_s|_{H_s}^2 d\lambda_s$;
2. $F \rightarrow \mathcal{I}(F)$ is linear up to evanescence;
3. If $F^n, F \in \hat{L}_{loc}^2(Q)$ and $\int_0^t |F_s^n - F_s|_{H'_s}^2 d\lambda_s \rightarrow 0$ in probability for each $t > 0$, then for each $t > 0$

$$\sup_{s \leq t} |\mathcal{I}_s(F^n) - \mathcal{I}_s(F)| \rightarrow 0 \text{ in probability, as } n \rightarrow \infty.$$

Moreover, this extension is unique (up to evanescence).

Proof:

1°. Let $F \in \hat{L}_{loc}^2(Q)$. By Lemma 2.3 there exists a sequence $F^n \in \tilde{L}_{loc}^2(Q)$ such that for each $t > 0$

$$\int_0^t |F_s^n - F_s|_{H'_s}^2 d\lambda_s \rightarrow 0$$

in probability, as $n \rightarrow \infty$. Thus there exists increasing sequences of stopping times $(\tau_{N,p}), (\tau_p)$ such that $\tau_{N,p} \leq \tau_p$ for each N, p and for each $t > 0$

$$\mathbf{P}(\tau_p < t) \xrightarrow{p \rightarrow \infty} 0, \quad \mathbf{P}(\tau_{N,p} < \tau_p) \xrightarrow{n \rightarrow \infty} 0,$$

$$\mathbf{E} \int_0^{\tau_p} |F_s|_{H'_s}^2 d\lambda_s < \infty, \quad \int_0^{\tau_{N,p}} |F_s^n|_{H'_s}^2 d\lambda_s \leq \int_0^{\tau_p} |F_s|_{H'_s}^2 d\lambda_s + 1.$$

Then for each p

$$\mathbf{E} \sup_t |\mathcal{T}_{t \wedge \tau_{N,p}}(F^N) - \mathcal{T}_{t \wedge \tau_{M,p}}(F^M)|^2 \rightarrow 0,$$

as $N, M \rightarrow \infty$. Thus the existence of the extension satisfying 1,2 follows immediately.

2°. Now we prove property 3 of the extension. Let $F^n, F \in \hat{L}_{loc}^2(Q)$ and for each $t > 0$

$$\int_0^t |F_s^n - F_s|_{H'_s}^2 d\lambda_s \rightarrow 0 \text{ in probability, as } n \rightarrow \infty.$$

Then there exists a sequence of stopping times (τ_n) such that for each $t > 0$ $\mathbf{P}(\tau_n < t) + \mathbf{E} \int_0^{\tau_n} |F_s^n - F_s|_{H'_s}^2 d\lambda_s \xrightarrow{n \rightarrow \infty} 0$. Then

$$\begin{aligned} \mathbf{E} \sup_t |\mathcal{I}_{t \wedge \tau_n}(F^n) - \mathcal{I}_{t \wedge \tau_n}(F)|^2 &\leq \\ &\leq C \mathbf{E} \int_0^{\tau_n} |F_s^n - F_s|_{H'_s}^2 d\lambda_s \rightarrow 0, \quad \text{in probability, as } n \rightarrow \infty. \end{aligned}$$

Thus the property 3 holds for the extension which is obviously unique. ■

2.5. Normalized Ito Integrals. By part b) of Lemma 2.3 the spaces $\hat{L}_{loc}^2(Q)$ and $L_{loc}^2(Q)$ are isomorphic. Making use of this isomorphism, we will define now an integral of a function $h \in L_{loc}^2(Q)$. This integral turned out to be quite useful in the study of absolute continuity of measures corresponding to solutions of infinite dimensional martingale problems (see [[29]]). By Lemma 2.3 for each $h \in L_{loc}^2(Q)$, there exists a unique $F \in \hat{L}_{loc}^2(Q)$ such that $h_s = F_s \circ Q_s$.

DEFINITION 2.4. The normalized integral of $h \in L_{loc}^2(Q)$ is defined by

$$\mathcal{R}_t(h) = \int_0^t h_s * dM_s = \mathcal{I}_t(F) = \int_0^t F_s dM_s,$$

where $h_s = F_s \circ Q_s$.

PROPOSITION 2.6. The map $h \rightarrow \mathcal{R}(h) = \int_0^t h_s * dM_s$ defined on the set $L_{loc}^2(Q)$ has the following properties:

1. $\mathcal{R}(h) \in \mathcal{M}_{loc}^2(\mathbb{F}, \mathbf{P})$, $\langle \mathcal{R}(h) \rangle_t = \int_0^t |h_s|_{H_s}^2 d\lambda_s$;
2. $\mathcal{R}(h)$ is linear up to evanescence;
3. If $h_n, h \in L_{loc}^2(Q)$ and for each $t > 0$ $\int_0^t |h_s^n - h_s|_{H_s}^2 d\lambda_s \rightarrow 0$ in probability, as $n \rightarrow \infty$, then for each $t > 0$

$$\sup_{s \leq t} |\mathcal{R}_s(h_n) - \mathcal{R}_s(h)| \rightarrow 0 \text{ in probability, as } n \rightarrow \infty.$$

Proof:

All the properties follows immediately from the definition of this integral and Proposition 2.5. ■

2.6. Special Stochastic Integrals. In this section we discuss the structure of the integrand spaces of the Ito integral $\mathcal{I}_t(f)$ and the normalized integral $\mathcal{R}_t(f)$ in several important cases. These include stochastic integrals with respect to Hilbert-valued and measure-valued martingales and stochastic flows.

2.6.1. Stochastic Integral with Respect to Cylindrical Brownian Motion. This integral is one of the simplest possible examples of the stochastic integral developed above. In this case the covariance spaces $H_s = H'_s = L^2(\mathcal{O})$ (see Example 2.6), and so $L^2_{loc}(Q) = \hat{L}^2_{loc}(Q)$ is the space of all $\mathcal{P}(\mathbb{F})$ -measurable functions f_s , taking values in $L^2(\mathcal{O})$ such that $\mathbf{P}\left\{\int_0^t \int_{\mathcal{O}} |f_s(x)|^2 dx ds < \infty\right\} = 1$.

Obviously, in this case the Ito stochastic integral and the normalized stochastic integrals coincide.

2.6.2. Stochastic Integrals with Respect to Hilbert Space-valued Martingales. Let E be a separable Hilbert space and M_t be an E -valued LSI martingale. Then the covariance operator $Q_t = Q_t^M$ and $\lambda_t = \langle M \rangle_t$. Here and thereafter in this Section we use the same notation and assumptions as in Examples 2.3 and 2.7. For the sake of simplicity it will be assumed that in the Example 2.7, $Z = E$ and $\sigma_s = \sigma_s^* = Q_s^{1/2}$.

To describe the Ito and the normalized stochastic integrals, it suffices to characterize the spaces $L^2_{loc}(Q)$, $\hat{L}^2_{loc}(Q)$. In our case, for every $t > 0$, there exists a CONS $\{e_k(t)\}_{k \geq 1}$ of eigenvectors of Q_t , and

$$Q_t = \sum_k \lambda_k(t)(e_k(t), \cdot)_E e_k(t)$$

where λ_k is the eigenvalue corresponding to e_k (see [2]). We can and will assume that the eigenfunctions e_k and the eigenvalues λ_k are $\mathcal{P}(\mathbb{F})$ -measurable.

PROPOSITION 2.7. (i) H_s is the set of all $\mathcal{P}(\mathbb{F})$ -measurable functions $h_s = \sum_k \lambda_k^{1/2}(s) h_k(s) e_k(s)$, where h_k are scalar functions such that $\sum_k h_k^2(s) < \infty$. The norm $|h_s|_{H_s}^2 = \sum_k 1_{\{\lambda_k(s) > 0\}} h_k^2(s)$. $L^2_{loc}(Q)$ is the set of all functions $h_s \in H_s$ such that for every k , $\lambda_k(s) h_k(s) e_k(s)$ is $\mathcal{P}(\mathbb{F})$ -measurable, and

$$\mathbf{P}\left\{\int_0^t \sum_k h_k^2(s) 1_{\{\lambda_k(s) > 0\}} d\langle M \rangle_s < \infty, \text{ for all } t > 0\right\} = 1$$

(ii) H'_s is the set of all $\mathcal{P}(\mathbb{F})$ -measurable functions

$$F(s) = \sum_k 1_{\{\lambda_k(s) > 0\}} \lambda_k^{-1/2}(s) F_k(s) e_k(s)$$

such that $|F_s|_{H'_s}^2 = \sum_k 1_{\{\lambda_k(s) > 0\}} F_k^2(s) < \infty$. $\hat{L}^2_{loc}(Q)$ is the set of all functions $F_s \in H'_s$ such that for every k , $\lambda_k(s) F_k(s) e_k(s)$ is $\mathcal{P}(\mathbb{F})$ -measurable, and

$$\mathbf{P}\left\{\int_0^t \sum_k F_k^2(s) 1_{\{\lambda_k(s) > 0\}} d\langle M \rangle_s < \infty, \text{ for all } t > 0\right\} = 1.$$

(iii) The duality between these spaces H_s and H'_s is given by

$$\langle F(s), h(s) \rangle_{H_s, H'_s} = \sum_{\lambda_k > 0} h_k(s) F_k(s)$$

Proof:

By Proposition 2.3, the covariance space $H_s = Q_s^{1/2} E$. Obviously, \tilde{H}_s , the closure of $Q_s^{1/2} E$ in E , is the set of all $\mathcal{P}(\mathbb{F})$ -measurable E -valued functions h_s , such that $h_s = \sum_k 1_{\{\lambda_k(s) > 0\}} h_k(s) e_k(s)$ and $\sum_k h_k^2(s) < \infty$. It is readily checked now that the Proposition follows from the definitions of the spaces $L^2_{loc}(Q)$ and $\hat{L}^2_{loc}(Q)$, Proposition 2.3 and Corollary 2.2. ■

We know from Corollary 2.2 that the operator Q_s can be extended continuously from E to H'_s so that Q_s provides a canonical isometry between H_s and H'_s . In this particular case the extension procedure is quite transparent..

Indeed, for $F(s) = \sum_k 1_{\{\lambda_k(s)>0\}} \lambda_k^{-1/2}(s) F_k(s) e_k(s) \in H'_s$, we define

$$(2.11) \quad Q_s F(s) = \sum_k 1_{\{\lambda_k(s)>0\}} \lambda_k^{-1/2}(s) F_k(s) Q_s e_k(s) = \sum_k 1_{\{\lambda_k(s)>0\}} \lambda_k^{1/2}(s) F_k(s) e_k(s).$$

It follows from the proposition that (2.11) is the desired canonical isometry.

The standard relation $\mathcal{R}_t(h) = \mathcal{I}_t(Q^{-1}h)$ between the Ito and the normalized stochastic integrals can be written in the form

$$(2.12) \quad \int_0^t \left(\sum_k \lambda_k^{1/2}(s) h_k(s) e_k(s) \right) * dM_s = \int_0^t \left(\sum_k 1_{\{\lambda_k(s)>0\}} \lambda_k^{-1/2}(s) h_k(s) e_k(s) \right) dM_s$$

2.6.3. Stochastic Integrals with Respect to Measure-valued Martingales. We will use the notation and the assumptions of Example 2.4 throughout this section.

PROPOSITION 2.8. (i) For each s , H_s is the set of all measures on $(U, \mathcal{B}(U))$ of the form $f(x)q_s(dx)$ such that $\int_U (f(x))^2 q_s(dx) < \infty$, and H'_s is the set of all measurable functions f on $(U, \mathcal{B}(U))$ such that $\int_U (f(x))^2 q_s(dx) < \infty$.

(ii) $L_{loc}^2(Q)$ is the set of all $\mathcal{P}(\mathbb{F})$ -measurable measures on $(U, \mathcal{B}(U))$ of the form $f(s, x)q_s(dx)$ (f is a $\mathcal{P}(\mathbb{F}) \otimes \mathcal{B}(U)$ -measurable function) such that.

$$\begin{aligned} \int_0^t \int_U |f(s, x)|^2 q_s(dx) d\lambda_s &< \infty \text{ P-a.s., } \forall t, \\ |f_s q_s|_{H_s}^2 &= \int_U |f(s, x)|^2 q_s(dx). \end{aligned}$$

(iii) $\hat{L}_{loc}^2(Q)$ is the set of all $\mathcal{P}(\mathbb{F}) \otimes \mathcal{B}(U)$ -measurable functions f such that

$$\begin{aligned} \int_0^t \int_U |f(s, x)|^2 q_s(dx) d\lambda_s &< \infty \text{ P-a.s., } \forall t, \\ |f_s|_{H'_s}^2 &= \int_U |f(s, x)|^2 q_s(dx). \end{aligned}$$

Proof:

The set $Q_s E'$ consists of all measures of the form $f(u)q_s(du)$, $f \in E'$. We will treat it as the space of classes of equivalent measures, two measures $f(u)q_s(du)$ and $g(u)q_s(du)$ are equivalent if $f = g$ q_s -a.s. This vector space is endowed with the norm

$$|f(u)q_s(du)|_{H_s}^2 = \int_{\Omega} f^2(u)q_s(du) = \langle Q_s f, f \rangle$$

becomes a Banach space isometrically isomorphic to $L^2(U, \mathcal{B}(U), q_s) = H'_s$, and hence complete. The rest of the proof is straightforward. ■

PROPOSITION 2.9. Let $B_1(U)$ be a linear vector sub-space of $B_0(U)$ and the embedding is dense in the weak topology $\sigma(B_0(U), E)$. The statement of (2.1) and Example 2.8 remain true if $B_1(U)$ is chosen for E' instead of $B_0(U)$.

This can be proved by reference to Corollary 4 of Proposition 21 in [10].

The following example illustrates possible applications of Proposition 2.9.

EXAMPLE 2.9. Since $\mathcal{B}(U)$ is countable, there is a countable set \mathcal{K} of measurable subsets K of U so that $K \subset U_n$ for some n and $\sigma(1_K, K \in \mathcal{K}) = \mathcal{B}(U)$. Let $B_1(U)$ be the vector space of finite linear combinations of indicators $1_K, K \in \mathcal{K}$. Obviously this space is dense in $B_0(U)$ even in the uniform convergence topology, and hence Proposition 2.9 applies. So the stochastic integral is defined for cylindrical martingale measures restricted to finite linear combinations of indicators. The structure of the spaces H_s , H'_s and $L^2_{loc}(Q)$, $\hat{L}^2_{loc}(Q)$ in this case is the same as above.

The setting discussed in this example includes the orthogonal martingale measures studied by Walsh [28].

We remark that in the setting of this section, for $h \in L^2_{loc}(Q)$,

$$(2.13) \quad \int_0^t h_s * dM_s = \int_0^t (dh_s / dq(s)) dM_s.$$

This formula follows in a simple way from Proposition 2.8 and Example 2.8.

2.6.4. Integrals with Respect to Stochastic Flows. In this subsection we generalize some results of H. Kunita in [3] concerning the integrals with respect to stochastic flows.

Let U be a domain in \mathbf{R}^d . Assume for each $x \in U$ we are given $M_t(x) \in \mathcal{M}_{loc}^2(\mathbb{F}, \mathbf{P})$. Suppose that there exists a $\mathcal{P}(\mathbb{F}) \otimes \mathcal{B}(U) \otimes \mathcal{B}(U)$ -measurable function $Q_s(x, y)$ and an increasing $\mathcal{P}(\mathbb{F})$ -measurable process λ_t such that

$$(2.14) \quad M_t(x) M_t(y) - \int_0^t Q_s(x, y) d\lambda_s \in \mathcal{M}_{loc}(\mathbb{F}, \mathbf{P}).$$

We assume that Q_s is separately continuous on $U \times U$, continuous on the diagonal and locally bounded. In addition, we assume that Q_s is symmetric and non-negative definite, i.e. $\forall(\xi_n), \forall(x_j)$,

$$\sum_{n,j} \xi_n \xi_j Q_s(x_n, x_j) \geq 0.$$

Let E be the set of all real valued functions on U with the topology of simple convergence and E' its dual space, i.e., the space of finite combinations of Dirac measures. If $\mu \in E'$, we have $\mu = \sum_x c_x \delta_x$ and only a finite number of $c_x \neq 0$. For $\mu = \sum_x c_x \delta_x$ we can define

$$M_t(\mu) = \sum_x c_x M_t(\delta_x) = \sum_x c_x M_t(x).$$

Obviously, (2.14) yields

$$M_t(\mu) M_t(\mu') - \int_0^t \sum_{x,y} c_x c'_y Q_s(x, y) d\lambda_s \in \mathcal{M}_{loc}(\mathbb{F}, \mathbf{P}),$$

where

$$\mu = \sum_x c_x \delta_x, \quad \mu' = \sum_x c'_x \delta_x.$$

For $\mu = \sum_x c_x \delta_x$, we define $\bar{Q}_s \mu(x) = \sum_y c_y Q_s(y, x)$, i.e., $\bar{Q}_s \in \mathcal{L}^+(E)$ and $\forall \mu' \in E'$

$$M_t(\mu) M_t(\mu') - \int_0^t \langle \mu', \bar{Q}_s \mu \rangle_{E', E} d\lambda_s \in \mathcal{M}_{loc}(\mathbb{F}, \mathbf{P}) .$$

If f is an U -valued $\mathcal{P}(\mathbb{F})$ -measurable function such that for each $t > 0$,

$$\int_0^t Q_s(f_s, f_s) d\lambda_s < \infty \text{ P-a.s.},$$

then $\delta_{f_s} \in \tilde{L}_{loc}^2(Q)$. Indeed,

$$\int_0^t \langle \delta_{f_s}, \bar{Q}_s \delta_{f_s} \rangle_{E', E} d\lambda_s = \int_0^t Q_s(f_s, f_s) d\lambda_s .$$

Thus the integral $\int_0^t \delta_{f_s} dM_s$ is defined by Proposition 2.4 and it coincides with the integral based on martingales with spatial parameters considered in [3]. As it was proved in [30] the class of integrands can be essentially enlarged depending on the smoothness of Q_s .

3. Martingale Problems

3.1. Generalities and Definitions. In studying stochastic differential equations one has to differentiate between strong and weak solutions. Loosely speaking, the relation (or the difference) between these two is similar to the one between a random variable and its law. To motivate the definitions and the procedures which are discussed below in this section let us start with a simple example.

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space with the right-continuous filtration of σ -algebras $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$. Denote by $\mathcal{P}(\mathbb{F})$ the \mathbb{F} -predictable σ -algebra. Let U be a smooth bounded domain in \mathbf{R}^d . Consider the scale of Sobolev spaces $\mathcal{H}^s = \mathcal{H}^s(U)$, $s \in \mathbf{R}$, generated by $L^2(U)$ and the operator $I - \Delta$, where Δ is Dirichlet Laplacian (see Appendix). The norm and the scalar product in $\mathcal{H}^s(H^0 = L^2(U))$ will be denoted by $|\cdot|_s$ and $(\cdot, \cdot)_s$, respectively. The duality between \mathcal{H}^1 and \mathcal{H}^{-1} will be denoted by $\langle \cdot, \cdot \rangle$.

Let us assume that for all $v, u \in L^2(U)$,

$$|f(v)|_0^2 \leq C_1 + C_2 |v|_1^2$$

and

$$(3.1) \quad |f(v) - f(u)|_0 \leq C |v - u|_0 .$$

Let W_Q be an $L^2(U)$ -valued Brownian motion on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ with the nuclear covariance operator $Q = (1 - \Delta)^{-\alpha}$ for some $\alpha > d/2$. Consider the equation

$$(3.2) \quad \begin{aligned} du(t, x) &= (\Delta u(t, x) + f(u(t, x))) dt + dW_Q(t), \quad (t, x) \in (0, \infty) \times \mathbf{R}^d, \\ u(t, x)|_{\partial G} &= 0 \text{ and } u(0, x) = u_0(x) \in \mathcal{H}^0 \end{aligned}$$

It is a standard fact (see e. g. [51], [47]) that there is a $\mathcal{P}(\mathbb{F})$ -measurable function $u \in D = D_{[0, \infty)}(\mathcal{H}^{-1})$, the set of all \mathcal{H}^{-1} -valued weakly continuous functions

on $[0, \infty)$, such that $\mathbf{P}(\int_0^t |u(s)|_1^2 ds < \infty \text{ for all } t > 0) = 1$ and for every $\varphi \in \mathcal{H}^1$,

$$(3.3) \quad \begin{aligned} & \langle u(t), \varphi \rangle - (u_0, \varphi)_0 - \int_0^t (\langle \Delta u(s), \varphi \rangle + (f(u(s)), \varphi)_0) ds \\ &= (W_Q(t), \varphi)_0 \end{aligned}$$

\mathbf{P} -a.s.

It also follows from [51] that u is adapted with respect to the filtration $(\mathcal{F}_t^{W_Q})_{t \geq 0}$ generated by W^Q and so u is a *strong* solution to equation (3.2).

Denote by X_t the canonical process on D , $X_t = X_t(w) = w_t$ for $w \in D$. Let $\mathcal{D} = \sigma(X_s, s \geq 0)$, $\mathcal{D}_t = \sigma(X_s, s \leq t)$, and $\mathbb{D} = (\mathcal{D}_{t+})_{t > 0}$. Denote by $\mathcal{P} = \mathcal{P}(\mathbb{D})$ the σ -algebra of predictable sets with respect to \mathbb{D} . Write $B_t = \int_0^t |X_s|_1^2 ds$. Define the probability measure \mathbf{P}' on \mathcal{D} so that $\mathbf{P}'(A) = \mathbf{P}(u \in A)$ for any $A \in \mathcal{D}$.

Obviously, the measure \mathbf{P}' satisfies the following relations:

- (a) $\mathbf{P}'(X_0 \in A) = \mathbf{P}(u_0 \in A)$ for any $A \in \mathcal{D}$,
- (b) $\mathbf{P}'(X : B_t < \infty) = 1$, for all $t > 0$,
- (c) For every $\varphi, \psi \in \mathcal{H}^1$,

(3.4)

$$W'_t(\varphi) = \langle X_t, \varphi \rangle - (X_0, \varphi)_0 - \int_0^t (\langle \Delta X_s, \varphi \rangle + (f(X_s), \varphi)_0) ds \in \mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P}'),$$

and

$$(3.5) \quad W'_t(\varphi)W'_t(\psi) - (Q\varphi, \psi)t \in \mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P}').$$

Note that (b) yields $\int_0^t |\Delta X_s|_{-1} ds < \infty$ \mathbf{P}' -a.s., and so the integrals in (3.4) are defined.

Now let us drop the assumption (3.1) and assume only that f is measurable. Then we can not anymore claim the existence of a strong solution to equation (3.2). However, as it will be shown below (see Subsection 3.2.3) there still exists a probability measure \mathbf{P}' on (D, \mathcal{D}) with the properties (a),(b) and (c). We will call this measure a solution to the martingale problem associated with equation (3.2).

Of course, existence of a strong solution implies existence of a solution to the associated martingale problem. This statement is not in general reversible.. However, existence of a solution to the martingale problem (a),(b),(c) yields that there exist a (possibly different) stochastic basis $(\Omega', \mathbb{F}', \mathbf{P}')$, an $L^2(U)$ -valued Brownian motion \tilde{W}_Q with covariance operator Q on this basis, and a $\mathcal{P}(\mathbb{F}')$ -measurable process \tilde{u} in $L^2(U)$ so that \tilde{u} satisfies equation (3.3) where W_Q is replaced by W'_Q , \mathbf{P}' -a.s.

Indeed, by (c), $(W'_t(\varphi))_{\varphi \in \mathcal{H}^1}$ is a cylindrical Brownian motion in \mathcal{H}^{-1} on the stochastic basis $(D, \mathcal{D}, \mathbb{D}, \mathbf{P}')$ with the nuclear covariance operator Q . Then, there exists an $L^2(U)$ -valued Brownian motion W'_Q such that $\mathbf{P}'(W'_t(\varphi) = (W'_Q(t), \varphi))$ for all $t \geq 0$ and $\varphi \in \mathcal{H}^1$ (see Example 2.2). Thus, $\tilde{u}_s = X_s$ is a solution of equation (3.2), with $W'_Q(t)$ substituted for $W_Q(t)$, on the stochastic basis $(D, \mathcal{D}, \mathbb{D}, \mathbf{P}')$. The function, \tilde{u} , is usually referred to as a *weak* solution of equation (3.2).

Obviously, any weak solution of equation (3.2) generates a measure on $(D, \mathcal{D}, \mathbb{D})$ which solves the associated martingale problem (a),(b),(c). Thus, we see that existence of a solution to the martingale problem (a), (b), (c) is equivalent to existence of a weak solution of the associated stochastic differential equation.

Using arguments similar to those applied in the case of equation (3.2) and factorizing LSI cylindrical martingales as stochastic integrals with respect to a cylindrical Brownian motion (see Lemma 3.2) one can prove that the above "equivalence principle" is essentially universal.

Now let us introduce a general definition of a martingale problem in a Schwartz pair.

Let E and E' be topological spaces like in Section 1 and $D = D_{[0,\infty)}(E_w)$ be the set of all E -valued weakly continuous functions on $[0,\infty)$. Denote by X_t the canonical process on D , $X_t = X_t(w) = w_t$ for $w \in D$. Let $\mathcal{D} = \sigma(X_s, s \geq 0)$, $\mathcal{D}_t = \sigma(X_s, s \leq t)$, and $\mathbb{D} = (\mathcal{D}_{t+})_{t \geq 0}$. Denote by $\mathcal{P} = \mathcal{P}(\mathbb{D})$ the σ -algebra of predictable sets with respect to \mathbb{D} .

Everywhere below we will use the following notation. Let (Y, \mathcal{Y}) and (Z, \mathcal{Z}) be two measurable spaces. If ξ is a measurable function from Y to Z and ν is probability measure on (Y, \mathcal{Y}) , we denote by $\xi(\nu)$ the measure on \mathcal{Z} defined by $\xi(\nu)(A) = \nu(\xi \in A)$. If ρ is a nonnegative function on Z integrable with respect to $\xi(\nu)$, then $\rho\xi(\nu)$ stands for a measure $\rho\xi(\nu)(A) = \int_A \rho d\xi(\nu)$.

Let $Q : [0, \infty) \times D \rightarrow \mathcal{L}^+(E)$ be a kernel such that for any $y', y'' \in E'$, $\langle Qy', y'' \rangle$ is $\mathcal{P}(\mathbb{D})$ -measurable. Let $\lambda = \lambda_t$ be a continuous non-decreasing $\mathcal{P}(\mathbb{D})$ -measurable real-valued process, and μ_0 be a probability measure on E . Suppose also that we are given a process $A : [0, \infty) \times D \rightarrow L^*(E', \mathbf{R})$, the space of linear (not necessarily continuous) operators on E' , such that for every $y \in E'$, $A_t(y)$ is a continuous $\mathcal{P}(\mathbb{D})$ -measurable process of bounded variation. Let B_t be an increasing \mathbb{D} -adapted $[0, \infty]$ -valued process. We say that $(\mu_0, dA, Qd\lambda, B)$ defines a martingale problem.

DEFINITION 3.1. A probability measure \mathbf{P} on (D, \mathcal{D}) is a solution of the martingale problem $(\mu_0, dA, Qd\lambda, B)$ if

- (i) $X_0(\mathbf{P}) = \mu_0$, $\mathbf{P}(B_t < \infty) = 1$ for all $t > 0$;
- (ii) For every $y', y'' \in E'$, $M_t(y') =$

$$\langle X_t, y' \rangle - \langle X_0, y' \rangle - A_t(y') \in \mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P}),$$

and

$$M_t(y')M_t(y'') - \int_0^t \langle Q_s y', y'' \rangle d\lambda_s \in \mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P}).$$

The set of all solutions to $(\mu_0, dA, Qd\lambda, B)$ is denoted $\mathcal{S}(\mu_0, dA, Qd\lambda, B)$.

The role of B_t in the Definition 3.1 is to specify the support of a solution of the martingale problem. The utility of this process is demonstrated in the example in the beginning of this Section. However, in many cases, in particular in the important example below, there is no need in sharper requirements on the support of a solution. In such cases we will take $B = 0$ and write simply $(\mu_0, dA, Qd\lambda)$ and $\mathcal{S}(\mu_0, dA, Qd\lambda)$.

EXAMPLE 3.1. (Super-Brownian Motion). Let us take $E' = C_0^\infty(\mathbf{R}^d)$ and $E = M_F(\mathbf{R}^d)$, the space of finite measures on \mathbf{R}^d endowed with the topology $\sigma(M_F(\mathbf{R}^d), C_b^2(\mathbf{R}^d))$. For $s \in \mathbf{R}_+$ and $f \in C_b^2(\mathbf{R}^d)$, set $d\lambda_s = ds$, and $Q_s f(u) = \gamma f(u)X_s(du)$, or equivalently,

$$\langle Q_s f, g \rangle = \gamma \int_{\mathbf{R}^d} f(u)g(u) X_s(du) \text{ for all } f, g \in C_b^2(\mathbf{R}^d)$$

where $\gamma > 0$ is fixed. Clearly, $(M_F(\mathbf{R}^d), C_b^2(\mathbf{R}^d))$ is a Schwartz pair. For an arbitrary $\mu_0 \in M_F(\mathbf{R}^d)$, consider the martingale problem

$$(3.6) \quad X_0(P) = \mu_0.$$

For every $f, g \in E'$,

$$(3.7) \quad M_t(f) = \int_{\mathbf{R}^d} f(u) X_t(du) - \int_{\mathbf{R}^d} f(u) X_0(du) - \frac{1}{2} \int_0^t \int_{\mathbf{R}^d} \Delta f(u) X_s(du) ds$$

belongs to $\mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P})$ and

$$(3.8) \quad M_t(f) M_t(g) - \gamma \int_0^t \int_{\mathbf{R}^d} f(u) g(u) X_s(du) ds \in \mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P}).$$

It is well known (see Chapter 2, Section 1.3) that the super-Brownian motion is the only solution to this martingale problem.

3.2. Absolute Continuity and Singularity of Solutions. This section is concerned with absolute continuity of measures which solve martingale problems in topological spaces. We will establish criteria generalizing Girsanov's theorem, characterize all the measures absolutely continuous with respect to a solution of a given martingale problem, find necessary and sufficient conditions for equivalence of two solutions, and apply these results in studying uniqueness and existence of solutions to martingale problems.

Everywhere below we will write $\mu \prec \nu$ and $\mu \sim \nu$ if the measure μ is absolutely continuous with respect to ν and mutually absolutely continuous (equivalent) with ν , respectively.

3.2.1. Girsanov's Theorem and other Criteria of Absolute Continuity. We begin with a version of Girsanov's theorem convenient for our setting.

PROPOSITION 3.1. (Girsanov's Theorem [13]) *Let us suppose that the probability \mathbf{P} is in $S(\mu_0, dA, Qd\lambda, B)$, that $\tilde{\mu}_0$ is a measure absolutely continuous with respect to μ_0 and tha $h \in L_{loc}^2(Qd\lambda, P)$. Write $\rho_0 = d\tilde{\mu}_0/d\mu_0$,*

$$\rho_t = \rho_0 \exp \left\{ \int_0^t h_s * dM_s - \frac{1}{2} \int_0^t |h_s|_s^2 d\lambda_s \right\},$$

and assume

$$(3.9) \quad \mathbf{E}^P \rho_t = 1 \text{ for all } t > 0.$$

Let $\tilde{\mathbf{P}}$ be a measure on (D, \mathcal{D}) such that for each $t > 0$

$$\tilde{\mathbf{P}}|_{\mathcal{D}_t} = \rho_t \mathbf{P}|_{\mathcal{D}_t}.$$

Then $\tilde{\mathbf{P}} \in S(\tilde{\mu}_0, dA + hd\lambda, Q d\lambda, B)$.

Proof:

To prove the proposition, we need only to establish that for all $y, y' \in E'$,

- (a) the process $\tilde{M}_t(y) = \langle X_t, y \rangle - \langle X_0, y \rangle - \langle A_t, y \rangle - \int_0^t \langle h_s, y \rangle d\lambda_s \in \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}});$
- (b)

$$\tilde{M}_t(y) \tilde{M}_t(y') - \int_0^t \langle Q_s y', y \rangle d\lambda_s \in \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}}).$$

The proofs of (a) and (b) are very much alike, so we restrict ourselves to the former one. Obviously, to establish (a) it suffices to show that $\rho_t \tilde{M}_t(y) \in \mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P})$, and that is what we demonstrate below.

Using the representations $\rho_t = \rho_0 + \int_0^t \rho_s d\mathcal{R}_s(h)$ where $\mathcal{R}_t(h) = \int_0^t h_u * dM_u$ and $M_t(y) = \int_0^t Q_s y * dM_s$, $y \in E'$, and standard properties of the normalized stochastic integral (Proposition 2.6), we obtain

$$(3.10) \quad \begin{aligned} \langle \rho, M(y) \rangle_t &= \langle \int_0^t Q_s y * dM_s, \int_0^t \rho_s h_s * dM_s \rangle_t = \\ &\int_0^t (Q_s y, \rho_s h_s)_{H_s} d\lambda_s = \int_0^t \rho_s \langle h_s, y \rangle d\lambda_s \end{aligned}$$

The latter equality is due to the fact that for any $y \in E'$, $g \in H_s$, $(Q_s y, g)_{H_s} = \langle y, h_s \rangle$.

Now differentiating the product $\rho_t \tilde{M}_t(y)$ by Ito formula for local martingales (see [13]) and using (3.10) we arrive at

$$\rho_t \tilde{M}_t(y) = \int_0^t \rho_s dM_s(y) + \int_0^t \rho_s \tilde{M}_s(y) d\mathcal{R}_s(h) \in \mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P}).$$

This concludes the proof. ■

A well known sufficient condition for (3.9) is given by the Novikov Criterion.

PROPOSITION 3.2. *If $\mathbf{E}^P \exp\{\frac{1}{2} \int_0^t |h_s|_s^2 ds\} < \infty$ then $\mathbf{E}^P \rho_t = 1$ for all $t > 0$.*

Proof:

Since for $h \in L_{loc}^2(Q)$, $\mathcal{R}_t(h) = \int_0^t h_s * dM_s$ is a scalar LSI martingale and $\langle \mathcal{R}(h) \rangle_t = \int_0^t |h_s|_s^2 d\lambda_s$, the statement follows from the standard result for scalar martingales. ■

Unfortunately, unless h is bounded, the assumption $\mathbf{E}^P \rho_t = 1$ as well as Novikov's sufficient condition is usually difficult to verify; this is especially true for the infinite-dimensional case. A natural way to somewhat reduce this difficulty is to "localize" condition (3.9). Below, in Theorem 3.1, we give sufficient conditions for such localization. Everywhere below \prec stands for "absolutely continuous with respect to".

THEOREM 3.1. *Let $\mathbf{P}, \tilde{\mu}_0, \mu_0$ and M be like in Proposition 3.1 and $h : [0, \infty) \times D \rightarrow E$ be a function from class $L(Q)$. Let (τ_n) , $n \in \mathbb{N}$, be an increasing sequence of $(\mathcal{D}_t)_{t \geq 0}$ -stopping times. Assume that for any n , τ_n is bounded and $h^n := h_s 1_{(s < \tau_n)} \in L_{loc}^2(Qd\lambda, \mathbf{P})$. Write*

$$(3.11) \quad \rho_t = \rho_0 \exp \left\{ \int_0^t h_s * dM_s - \frac{1}{2} \int_0^t |h_s|_s^2 d\lambda_s \right\}.$$

Set $\rho^n = \rho_{\tau_n}$ and $\mathbf{P}^n = \rho^n X_{\wedge \tau_n}(\mathbf{P})$. Assume that for each n , $\mathbf{E}^P \rho^n = 1$ and $\lim_{n \rightarrow \infty} \mathbf{P}^n(\tau_n < t) = 0$ for each $t > 0$.

Then there exists a probability measure $\tilde{\mathbf{P}}$ on (D, \mathcal{D}) with these properties:

- 1) $h \in L_{loc}^2(Qd\lambda, \tilde{\mathbf{P}})$ and $\tilde{\mathbf{P}}(A) = \mathbf{P}^n(A)$ for $A \in \mathcal{D}_{\tau_n}$;
- 2) for every $t > 0$, $\tilde{\mathbf{P}}|_{\mathcal{D}_t} \prec \mathbf{P}|_{\mathcal{D}_t}$ and $d\tilde{\mathbf{P}}/d\mathbf{P}|_{\mathcal{D}_t} = \rho_t$;
- 3) $\tilde{\mathbf{P}} \in \mathcal{S}(\tilde{\mu}_0, dA + hd\lambda, Qd\lambda, B)$.

Proof:

To begin with we remark that the integrals in (3.11) are defined. Indeed, since $h_s 1_{(s < \tau_n)} \in L_{loc}^2(Qd\lambda, \mathbf{P})$ for every n , then $\int_0^\infty |h_s|_{H_s} d\lambda_s < \infty$ \mathbf{P} -a.e.

Now note that for $A \in \mathcal{D}$,

$$(3.12) \quad \mathbf{P}^n(A) = \mathbf{P}^{n+m}(X_{\cdot \wedge \tau_n} \in A),$$

Indeed,

$$\begin{aligned} \mathbf{P}^n(A) &= \mathbf{E}^{\mathbf{P}}[1_{\{X_{\cdot \wedge \tau_n} \in A\}} \rho_{\tau_n}^n] = \\ &= \mathbf{E}^{\mathbf{P}}[1_{\{X_{\cdot \wedge \tau_n} \in A\}} \mathbf{E}^{\mathbf{P}}[\rho^{n+m} | \mathcal{D}_{\tau_n}]] = \\ &= \mathbf{P}^{n+m}(X_{\cdot \wedge \tau_n} \in A). \end{aligned}$$

This implies of course that $\mathbf{P}^n(A) = \mathbf{P}^{n+m}(A)$ for any $A \in \mathcal{D}_{\tau_n}$. So, for $A \in \mathcal{D}_t$, $t > 0$

$$\begin{aligned} |\mathbf{P}^n(A) - \mathbf{P}^{n+m}(A)| &= |\mathbf{P}^{n+m}(X_{\cdot \wedge \tau_n \wedge t} \in A) - \mathbf{P}^{n+m}(X_{\cdot \wedge t} \in A)| \leq \\ &\leq \mathbf{P}^{n+m}(\tau_n \leq t) = \mathbf{P}^n(\tau_n \leq t). \end{aligned}$$

Since by the assumption, $\lim_{n \rightarrow \infty} \mathbf{P}^n(\tau^n < t) = 0$, $\{\mathbf{P}^n|_{\mathcal{D}_t}\}$ is a Cauchy sequence (in variation) and there exists a probability measure $\tilde{\mathbf{P}}$ on (D, \mathcal{D}) such that for each $t > 0$ $\text{Var}|\mathbf{P}^n|_{\mathcal{D}_t} - \tilde{\mathbf{P}}|_{\mathcal{D}_t}| \rightarrow 0$ as $n \rightarrow \infty$. Since τ_n are bounded, this together with (3.12) give

$$(3.13) \quad \mathbf{P}^n(A) = \tilde{\mathbf{P}}(A), \text{ where } A \in \mathcal{D}_{\tau_n}.$$

Now we prove that for every $t > 0$, $\tilde{\mathbf{P}}(\int_0^t |h_s|_s^2 d\lambda_s < \infty) = 1$. For an arbitrary ϵ , we choose n_ϵ such that $\tilde{\mathbf{P}}(\tau_{n_\epsilon} < t) = \mathbf{P}^{n_\epsilon}(\tau_{n_\epsilon} < t) < \epsilon$. Obviously

$$\begin{aligned} (3.14) \quad \tilde{\mathbf{P}}\left(\int_0^t |h_s|_s^2 d\lambda_s > N\right) &\leq \tilde{\mathbf{P}}\left(\int_0^{\tau_{n_\epsilon}} |h_s|_s^2 d\lambda_s > N/2\right) + \\ &+ \tilde{\mathbf{P}}\left(\int_{\tau_{n_\epsilon} \wedge t}^t |h_s|_s^2 d\lambda_s > N/2\right) \leq \mathbf{P}^{n_\epsilon}\left(\int_0^{\tau_{n_\epsilon}} |h_s|_s^2 d\lambda_s > N/2\right) + \epsilon. \end{aligned}$$

Since $h^n \in L^2_{loc}(Qd\lambda, \mathbf{P})$ and $\mathbf{P}^n \prec X_{\cdot \wedge \tau_n}(\mathbf{P})$, we have that $h^n \in L^2_{loc}(Q, \mathbf{P}^n)$. This and (3.14) yield

$$\limsup_{N \rightarrow \infty} \tilde{\mathbf{P}}\left(\int_0^t |h_s|_s^2 d\lambda_s > N\right) \leq \epsilon.$$

Due to the arbitrariness of ϵ , the latter implies that $h \in L^2_{loc}(Q, \tilde{\mathbf{P}})$.

By Girsanov theorem, $\mathbf{P}^n \in \mathcal{S}(\tilde{\mu}_0, dA_{\cdot \wedge \tau_n} + h^n d\lambda, Qd\lambda_{\cdot \wedge \tau_n})$. Therefore, by (3.13) we have $\tilde{\mathbf{P}} \in \mathcal{S}(\tilde{\mu}_0, dA + hd\lambda, Qd\lambda)$.

Now we prove 2). Let $A \in \mathcal{D}_t$ and $\mathbf{P}(A) = 0$. Obviously,

$$\begin{aligned} \mathbf{P}^n(A) &= \mathbf{E}^{\mathbf{P}}[1_{\{X_{\cdot \wedge \tau_n} \in A\}} \rho^n] = \\ &= \mathbf{E}^{\mathbf{P}}[1_{\{(X_{\cdot \wedge \tau_n} \in A) \cap (\tau_n < t)\}} \rho^n] + \int_A \rho^n d\mathbf{P} \leq \mathbf{P}^n(\tau_n < t). \end{aligned}$$

Thus $\lim_{n \rightarrow \infty} \mathbf{P}^n(A) = 0$. On the other hand

$$\tilde{\mathbf{P}}(A) \leq \tilde{\mathbf{P}}(X_{\cdot \wedge \tau_n} \in A) + \tilde{\mathbf{P}}(\tau_n < t) = \mathbf{P}^n(A) + \mathbf{P}^n(\tau_n < t).$$

This obviously implies that $\tilde{\mathbf{P}}(A) = 0$, and the local absolute continuity is proved.

Now we will prove that ρ_t is a $(\mathcal{D}_t, \mathbf{P})$ -martingale. Clearly,

$$1 = \mathbf{E}^{\mathbf{P}} \rho_{t \wedge \tau_n} = \mathbf{E}^{\mathbf{P}} 1_{\{\tau_n < t\}} \rho^n + \mathbf{E}^{\mathbf{P}} 1_{\{\tau_n \geq t\}} \rho_t.$$

By the Monotone Convergence Theorem, $\mathbf{E}^{\mathbf{P}} 1_{\{\tau_n \geq t\}} \rho_t \rightarrow \mathbf{E}^{\mathbf{P}} \rho_t$. On the other hand,

$$\mathbf{E}^{\mathbf{P}} 1_{\{\tau_n < t\}} \rho^n = \mathbf{P}^n(\tau_n < t) \rightarrow 0.$$

Thus, $\mathbf{E}^{\mathbf{P}} \rho_t = 1$ for all $t \geq 0$, and so ρ_t is a $(\mathcal{D}_t, \mathbf{P})$ -martingale. Now, 3) holds by Girsanov's theorem, and the proof is complete. ■

To illustrate possible applications of Theorem 3.1, let us consider a martingale problem slightly more general than (3.6)-(3.8). This problem, first studied by Dawson (see [58] and references therein and Chapter 2), is related to measure-valued processes with nonlinear birth and death rates.

Let $\theta : \mathbf{R}_+ \times D \times \mathbf{R}^d \rightarrow \mathbf{R}$ be an $\mathcal{P}(\mathbb{D}) \times \mathcal{B}(\mathbf{R}^d)$ -measurable function and $Q_s f(u) = \gamma f(u) X_s(du)$ for some $\gamma > 0$. Consider the martingale problem

$$(3.15) \quad X_0(\tilde{\mathbf{P}}) = \mu_0$$

$$(3.16) \quad \begin{aligned} M_t(f) &= \langle X_t, f \rangle - \langle X_0, f \rangle - \frac{1}{2} \int_0^t \int_{\mathbf{R}^d} (\Delta f(u) + \theta_s(u)) X_s(du) ds \\ &\in \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}}) \text{ for all } f \in C_b^2(\mathbf{R}^d), \end{aligned}$$

and for all $f, g \in C_b^2(\mathbf{R}^d)$,

$$(3.17) \quad M_t(f) M_t(g) - \gamma \int_0^t \int_{\mathbf{R}^d} f(u) g(u) X_s(du) ds \in \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}}).$$

COROLLARY 3.1. (Dawson [58]). Write

$$\rho_t = \exp \left\{ \int_0^t \int_{\mathbf{R}^d} \theta_s(u) dM_s - \frac{\gamma}{2} \int_0^t \int_{\mathbf{R}^d} |\theta_s(u)|^2 X_s(du) ds \right\}.$$

Assume that $\theta \leq \theta_0$ for some $\theta_0 < \infty$,

$$(3.18) \quad \int_0^t \int_{\mathbf{R}^d} |\theta_s(u)|^2 X_s(du) ds < \infty \quad \forall t > 0, \mathbf{P}-a.e.,$$

and $\sup\{|\theta_s(X, u)| : s \in \mathbf{R}_+, \langle 1, X \rangle \leq n, u \in \mathbf{R}^d\} < \infty$ for every $n \in \mathbb{N}$. Then ρ_t is a $(\mathcal{D}_t, \mathbf{P})$ -martingale, and the measure $\tilde{\mathbf{P}}$ defined by $\tilde{\mathbf{P}}|_{\mathcal{D}_t} = \rho_t \mathbf{P}|_{\mathcal{D}_t}$ is a solution of martingale problem (3.15)-(3.17) with $B = 0$.

Proof:

Set $h_s = \theta_s(u) \gamma X_s(du)$. Owing to (3.18), it follows from Propositions 2.8, 2.9 that $\theta \in \hat{L}_{loc}^2(Q)$ and $h \in L_{loc}^2(Q)$, and

$$(3.19) \quad |h_s|_{H_s}^2 = |\theta(s)|_{H'_s}^2 = \gamma \int_{\mathbf{R}^d} |\theta_s(u)|^2 X_s(du).$$

By the definition of the normalized integral (Definition 2.4),

$$\rho_t = \exp \left\{ \int_0^t h_s * dM_s - \frac{1}{2} \int_0^t |h_s|_{H_s}^2 ds \right\}.$$

Set $\tau_n = \inf\{t : \sup_{s \leq t} \langle X_s, 1 \rangle + \int_0^t \langle X_s, 1 \rangle ds > n\} \wedge n$. Write $h^n := h_s 1_{(s < \tau_n)}$, $\rho^n = \rho_{\tau_n}$ and $\mathbf{P}^n = \rho_\infty^n X_{\cdot \wedge \tau_n}(\mathbf{P})$. By Proposition 3.2, for every n , $\mathbf{E}^P \rho^n = 1$. Owing to Theorem 3.1, it suffices now to verify that $\lim_{n \rightarrow \infty} \mathbf{P}^n(\tau_n < t) = 0$.

Note that by Girsanov's theorem, $\mathbf{P}^n \in \mathcal{S}(\mu_0, dA_{\cdot \wedge \tau_n} + h^n dl, \gamma X_{\cdot \wedge \tau_n} dl_{\cdot \wedge \tau_n})$ where $dA_s(f) = \frac{1}{2} \langle X_s, \Delta f \rangle$ and $dl_s = ds$. In particular, we have that $\mathbf{P}^n - a.e.$

$$(3.20) \quad \langle X_{t \wedge \tau_n}, 1 \rangle = \langle \mu_0, 1 \rangle + \int_0^t 1_{(s < \tau_n)} \langle X_{s \wedge \tau_n}, \theta(s) \rangle ds + M_{t \wedge \tau_n}(1),$$

and

$$(3.21) \quad \langle M_{\cdot \wedge \tau_n}(1) \rangle_t = \gamma \int_0^{t \wedge \tau_n} \langle X_{s \wedge \tau_n}, 1 \rangle ds.$$

Clearly, equation (3.20) yields

$$\begin{aligned} \mathbf{E}^n \langle X_{t \wedge \tau_n}, 1 \rangle &= \langle \mu_0, 1 \rangle + \mathbf{E}^n \int_0^t 1_{(s < \tau_n)} \langle X_{s \wedge \tau_n}, \theta(s) \rangle ds \\ &\leq \langle \mu_0, 1 \rangle + \mathbf{E}^n \int_0^t \theta_0 1_{(s < \tau_n)} \langle X_{s \wedge \tau_n}, 1 \rangle ds. \end{aligned}$$

Then, by Gronwall lemma,

$$(3.22) \quad \mathbf{E}^n \langle X_{t \wedge \tau_n}, 1 \rangle \leq \langle \mu_0, 1 \rangle \exp[\theta_0 t \wedge \tau_n].$$

On the other hand, by (3.21) and Burkholder-Davis-Gundy inequality,

$$(3.23) \quad \mathbf{E}^n \sup_{s \leq t} |M_{s \wedge \tau_n}(1)| \leq 3\gamma \mathbf{E}^n \int_0^{t \wedge \tau_n} \langle X_{s \wedge \tau_n}, 1 \rangle ds \leq 3\gamma \langle \mu_0, 1 \rangle [\exp\{\theta_0 t \wedge \tau_n\} - 1].$$

Obviously, (3.20), (3.22), and (3.23) give

$$(3.24) \quad \mathbf{E}^n \sup_{s \leq t} \langle X_{s \wedge \tau_n}, 1 \rangle \leq \langle \mu_0, 1 \rangle (1 + 3\gamma + \theta_0) [\exp\{\theta_0 t \wedge \tau_n\} - 1] < \infty.$$

Now, from (3.22)-(3.24) follows that

$$\begin{aligned} \mathbf{P}^n(\tau_n < t) &\leq \frac{2}{n} \mathbf{E}^n \int_0^{t \wedge \tau_n} \langle X_{s \wedge \tau_n}, 1 \rangle ds \\ &+ \frac{2}{n} \mathbf{E}^n \sup_{s \leq t} \langle X_{s \wedge \tau_n}, 1 \rangle \leq const/n, \end{aligned}$$

and we are done. ■

3.2.2. Absolute Continuity, Singularity and Uniqueness of Solutions.

DEFINITION 3.2. The martingale problem $(\mu_0, dA, Qd\lambda, B)$ is said to satisfy condition U_{loc} if for any \mathbb{D} -predictable stopping time τ , the martingale problem $(\mu_0, dA_{\cdot \wedge \tau}, Qd\lambda_{\cdot \wedge \tau}, B_{\cdot \wedge \tau})$ has no more than one solution.

Of course U_{loc} is stronger assumption than "simple" uniqueness. However, uniqueness complemented by some simple measurability condition implies U_{loc} . A simple condition of this type is given in Proposition 3.3 below.

For $s > 0$ and $w, w' \in D$ such that $w'_s = w_s, B_s(w) < \infty$, we define the trajectory

$$\mathcal{J}_s(w, w')_r = \begin{cases} w_r, & \text{if } r \leq s \\ w'_r, & \text{if } r \geq s. \end{cases}$$

Let $A_t^{w,s} = A_t(\mathcal{J}_s(w, w')), \lambda_t^{w,s} = \lambda_t(\mathcal{J}_s(w, w')), Q_t^{w,s} = Q_t(\mathcal{J}_s(w, w')), B_t^{w,s} = B_t(\mathcal{J}_s(w, w'))$, $t \geq s$.

We say that $(dA, Q d\lambda, B)$ satisfies the condition *MF* if there exists a measurable family $(\mathbf{P}_{s,w})$ of probability measures on (D, \mathcal{D}) such that

$$\mathbf{P}_{s,w}(w'_r = w_r \ \forall r \leq s) = 1, \mathbf{P}(B_t^{w,s} < \infty) = 1 \text{ for all } t > 0 \text{ and}$$

$$X_t - X_s - (A_t^{w,s} - A_s^{w,s}) = M_t^s \in \mathcal{M}_{loc}^c((\mathcal{D}_{t+})_{t \geq s}, \mathbf{P}_{s,w}),$$

$$M_t^s(y') M_t^s(y'') - \int_s^t \langle y', Q_r^{w,s} y'' \rangle d\lambda_r^{w,s} \in \mathcal{M}_{loc}^c((\mathcal{D}_{t+})_{t \geq s}, \mathbf{P}_{s,w}).$$

PROPOSITION 3.3. *If the solution to the problem $(\mu_0, dA, Q d\lambda, B)$ is unique and if $(dA, Q d\lambda, B)$ satisfies the condition *MF*, then the condition *U_{loc}* is satisfied.*

Proof:

Fix a \mathbb{D} -predictable stopping time τ . Obviously τ is a $(\mathcal{D}_t)_{t \geq 0}$ -stopping time and $\mathcal{D}_\tau = \sigma(X_{u \wedge \tau}, u \geq 0)$ (see [26]). Suppose that the probability measures \mathbf{P}^1 and \mathbf{P}^2 are in $\mathcal{S}(\mu_0, dA_{\cdot \wedge \tau}, Q d\lambda_{\cdot \wedge \tau}, B_{\cdot \wedge \tau})$. Let $(P_{s,w})$ be the measurable family of probability measures existing according to *MF*. Define two measures on $(\mathcal{D}, \mathcal{P})$ by

$$\begin{aligned} \nu^i(f) &= \int f(\mathcal{J}_{\tau(w)}(w, w')) \mathbf{P}_{\tau(w), w}(dw') \mathbf{P}^i(dw), \\ i &= 1, 2, f \in \mathcal{D}. \end{aligned}$$

Then obviously $\nu^i|_{\mathcal{D}_\tau} = \mathbf{P}^i|_{\mathcal{D}_\tau}$, $i = 1, 2$. Fix $y, y' \in E'$, $m > 0$. Let $\tau_m = \inf(t : |M_t(y)| + |S_t(y, y')| \geq m) \wedge m$, $r \leq s$, $A \in \mathcal{D}_r$. If $N_t = M_{t \wedge \tau_m}(y)$ or $N_t = S_{t \wedge \tau_m}(y, y')$ we have $N_s - N_r = N_{s \vee \tau} - N_{r \vee \tau} + N_{s \wedge \tau} - N_{r \wedge \tau}$. Then

$$\begin{aligned} \nu^i[(N_s - N_r)1_A] &= \\ &= \int \int \mathbf{P}_{\tau(w), w}(dw') 1_A(\mathcal{J}_{\tau(w)}(w, w')) (N_{s \wedge \tau(w)}(w) - N_{r \wedge \tau(w)}(w)) \mathbf{P}^i(dw) + \\ &\quad + \int \int \mathbf{P}_{\tau(w), w}(dw') (N_{s \vee \tau(w)}(w) - N_{r \vee \tau(w)}(w)) 1_A(\mathcal{J}_{\tau(w)}(w, w')) \mathbf{P}^i(dw), \quad i = 1, 2. \end{aligned}$$

Obviously $1_A(\mathcal{J}_{\tau(w)}(w, w'))$ is a $\mathcal{D}_{\tau \wedge r}$ -measurable function of w and a $\mathcal{D}_{\tau \vee r}$ -measurable function of w' . Thus for $i = 1, 2$,

$$\int 1_A(\mathcal{J}_{\tau(w)}(w, w')) (N_{s \wedge \tau(w)}(w) - N_{s \wedge \tau(w)}(w')) \mathbf{P}^i(dw) = 0$$

and

$$\int 1_A(\mathcal{J}_{\tau(w)}(w, w')) (N_{s \vee \tau(w)}(w) - N_{s \vee \tau(w)}(w')) \mathbf{P}_{\tau(w), w}(dw') = 0.$$

So, $\nu^i[(N_s - N_r)1_A] = 0$. Since y, y' and $m > 0$ are arbitrary we derive that $\nu^i \in \mathcal{S}(\mu_0, dA, Q d\lambda)$, $i = 1, 2$, i.e. $\nu^1 = \nu^2$. Therefore $\mathbf{P}^1|_{\mathcal{D}_\tau} = \mathbf{P}^2|_{\mathcal{D}_\tau}$, and we are done. \blacksquare

THEOREM 3.2. (*Mikulevicius, Rozovskii [29]*) Let \mathbf{P} and $\tilde{\mathbf{P}}$ be probability measures on (D, \mathcal{D}) so that $\mathbf{P} \in \mathcal{S}(\mu_0, dA, Qd\lambda, B)$ and $\tilde{\mathbf{P}} \in \mathcal{S}(\tilde{\mu}_0, dA + hd\lambda, Qd\lambda, B)$ where $h \in L^2_{loc}(Qd\lambda, \tilde{\mathbf{P}})$. Assume that $\mu_0 \sim \tilde{\mu}_0$ and $(\mu_0, dA, Qd\lambda, B)$ satisfies U_{loc} . Then $\tilde{\mathbf{P}}|_{\mathcal{D}_t} \prec \mathbf{P}|_{\mathcal{D}_t}$ for each $t > 0$.

Proof:

Let $\tilde{M}_t := X_t - X_0 - A_t - \int_0^t h_s d\lambda_s$. By the assumptions for every $y', y'' \in E'$, $y'(\tilde{M}_t) \in \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}})$ and

$$y'(\tilde{M}_t)y''(\tilde{M}_t) - \int_0^t \langle y', Qy'' \rangle d\lambda_s \in \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}}).$$

For $n \in \mathbb{N}$, set $\tau_n = \inf(t : \int_0^t |h_s|_{H_s}^2 d\lambda_s \geq n) \wedge n$. Write $\rho_0 = d\mu_0/d\tilde{\mu}_0$,

$$\rho_t^n = \rho_0 \exp \left\{ - \int_0^{t \wedge \tau_n} h_s * d\tilde{M}_s - \frac{1}{2} \int_0^{t \wedge \tau_n} |h_s|_s^2 ds \right\},$$

and $\rho^n = \rho_{\tau_n}^n$.

Define $\mathbf{P}^n = \rho^n X_{\cdot \wedge \tau_n}(\tilde{\mathbf{P}})$. By Girsanov's Theorem (Theorem 3.1), \mathbf{P}^n is a solution to the martingale problem $(\mu_0, dA_{\cdot \wedge \tau_n}, Qd\lambda_{\cdot \wedge \tau_n}, B_{\cdot \wedge \tau_n})$ and

$$(3.25) \quad X_{\cdot \wedge \tau_n}(\tilde{\mathbf{P}}) \prec X_{\cdot \wedge \tau_n}(\mathbf{P}) \quad \forall n \in \mathbb{N}.$$

Let $A \in \mathcal{D}_t$ and $\mathbf{P}(A) = 0$. Then for each $n \in \mathbb{N}$,

$$\tilde{\mathbf{P}}(A) \leq \tilde{\mathbf{P}}(\tau_n \leq t) + \tilde{\mathbf{P}}(A \cap \{\tau_n > t\}).$$

Since τ_n is a $(\mathbb{D}, \tilde{\mathbf{P}})$ -predictable stopping time and due to (3.25) we have

$$\tilde{\mathbf{P}}(A \cap \{\tau_n > t\}) = X_{\cdot \wedge \tau_n}(\tilde{\mathbf{P}})(A \cap \{\tau_n > t\}).$$

On the other hand, since $(\mu_0, dA, Qd\lambda, B)$ satisfies U_{loc} ,

$$X_{\cdot \wedge \tau_n}(\mathbf{P})(A \cap \{\tau_n > t\}) = \mathbf{P}(A \cap \{\tau_n > t\}) \leq \mathbf{P}(A) = 0$$

So, $\tilde{\mathbf{P}}(A \cap \{\tau_n > t\}) = 0$ and

$$\tilde{\mathbf{P}}(A) \leq \lim_{n \rightarrow \infty} \tilde{\mathbf{P}}(\tau_n \leq t) = 0.$$

■

In fact the assumptions of the Theorem also guarantee uniqueness of a solution of the martingale problem $(\tilde{\mu}_0, dA + hd\lambda, Qd\lambda, B)$. Specifically, the following result holds.

THEOREM 3.3. (*Mikulevicius, Rozovskii [29]*). Suppose that $(\mu_0, dA, Qd\lambda, B)$ meets the condition U_{loc} , $\mathbf{P} \in \mathcal{S}(\mu_0, dA, Qd\lambda, B)$ and $\tilde{\mu}_0 \sim \mu_0$. Let $\tilde{\mathbf{P}}, \tilde{\mathbf{P}}' \in \mathcal{S}(\tilde{\mu}_0, dA + hd\lambda, Qd\lambda, B)$ and $h \in L^2_{loc}(Qd\lambda, \tilde{\mathbf{P}}) \cap L^2_{loc}(Qd\lambda, \tilde{\mathbf{P}}')$, i.e. $\int_0^t |h_s|_{H_s}^2 d\lambda_s < \infty$ $\tilde{\mathbf{P}} + \tilde{\mathbf{P}}'$ -a.s. for all $t > 0$. Then $\tilde{\mathbf{P}} = \tilde{\mathbf{P}}'$.

Proof:

Let us define $\tilde{M}_t, \rho_t^n, \rho^n$, and ρ_0 as in Theorem 3.2. By the assumptions, for every $y', y'' \in E'$ $y'(\tilde{M}_t) \in \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}}) \cap \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}}')$ and, $y'(\tilde{M}_t)y''(\tilde{M}_t) - \int_0^t \langle y', Qy'' \rangle d\lambda_s \in \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}}) \cap \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}}')$.

For $n \in \mathbb{N}$, set $\tau_n = \inf(t : \int_0^t |h_s|_{H_s}^2 d\lambda_s \geq n) \wedge n$ and

$$\rho_t^n = \rho_0 \exp \left\{ - \int_0^{t \wedge \tau_n} h_s * d\tilde{M}_s - \frac{1}{2} \int_0^{t \wedge \tau_n} |h_s|_s^2 ds \right\},$$

where $\rho_0 = d\mu_0/d\tilde{\mu}_0$.

Define $\mathbf{P}_0^n = \rho^n X_{\cdot \wedge \tau_n}(\tilde{\mathbf{P}})$ and $\mathbf{P}_1^n := \rho^n X_{\cdot \wedge \tau_n}(\tilde{\mathbf{P}}')$. By Girsanov's Theorem (Theorem 3.1) \mathbf{P}_0^n and \mathbf{P}_1^n are solutions to the martingale problem

$$(\mu_0, dA_{\cdot \wedge \tau_n}, Qd\lambda_{\cdot \wedge \tau_n}, B_{\cdot \wedge \tau_n}).$$

Since the assumption U_{loc} is satisfied, so we have $\mathbf{P}_0^n = \mathbf{P}_1^n = X_{\cdot \wedge \tau_n}(\mathbf{P})$. Since $\rho^n > 0$ $\tilde{\mathbf{P}} + \tilde{\mathbf{P}}'$ - a.s., this implies that $\tilde{\mathbf{P}}|_{\mathcal{D}_{\tau_n}} = \tilde{\mathbf{P}}'|_{\mathcal{D}_{\tau_n}}$ for each n . Taking into account that for each $t > 0$ $\tilde{\mathbf{P}}(\tau_n < t) + \tilde{\mathbf{P}}'(\tau_n < t) \rightarrow 0$ as $n \rightarrow \infty$ we derive from the latter equality that $\tilde{\mathbf{P}} = \tilde{\mathbf{P}}'$. ■

Let $\mathbf{P} \in \mathcal{S}(\mu_0, dA, Qd\lambda, B)$. Our next goal is to characterize the structure of all measures absolutely continuous with respect to \mathbf{P} .

To begin with, we will establish an important representation for local martingales M_t (with respect to \mathbf{P}) having integrable suprema $M^* = \sup_t |M_t|$.

Denote $J^1(\mathbf{P}) = \{M \in \mathcal{M}_{loc}(\mathbb{D}, \mathbf{P}) : |M|_1 = \mathbf{E}^\mathbf{P} \sup_t |M_t| < \infty\}$. Let

$$\mathcal{N} = \left\{ M_t(y), S_t(y, y') = M_t(y)M_t(y') - \int_0^t \langle Q_s y, y' \rangle d\lambda_s, y, y' \in E' \right\}.$$

Since $\mathbf{P} \in \mathcal{S}(\mu_0, dA, Qd\lambda)$ all the processes $N \in \mathcal{N}$ belong to $\mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P})$. For $N_t \in \mathcal{N}$, denote by $\mathcal{Z}^1(N)$ the set of the integrals $\int_0^t f_s dN_s$, where f is $\mathcal{P}(\mathbb{D})$ -measurable real valued and

$$\mathbf{E}^\mathbf{P} \left(\int_0^\infty f_s^2 d\langle N \rangle_s \right)^{1/2} < \infty.$$

Let $\mathcal{L}^1(\mathbf{P}, \mathcal{N})$ be the closure in the norm $|\cdot|_1$ of the linear span of $\cup_{N \in \mathcal{N}} \mathcal{Z}^1(N)$. We will call $\mathcal{L}^1(\mathbf{P}, \mathcal{N})$ the stable subspace of $J^1(\mathbf{P})$ generated by \mathcal{N} . Since $J^1(\mathbf{P})$ is a Banach space with respect to the norm $|\cdot|_1$, $\mathcal{L}^1(\mathbf{P}, \mathcal{N}) \subset J^1(\mathbf{P})$.

PROPOSITION 3.4. ([30]). *The stable subspace of $J^1(\mathbf{P})$ generated by \mathcal{N} is given by*

$$\begin{aligned} \mathcal{L}^1(\mathbf{P}, \mathcal{N}) = & \left\{ \mathcal{R}_t(f) = \int_0^t f_s * dM_s : f \in L^2(Qd\lambda, \mathbf{P}) \right. \\ & \text{and } \mathbf{E}^\mathbf{P} \left[\left(\int_0^\infty |f_s|_{H_s}^2 d\lambda_s \right)^{1/2} \right] < \infty \left. \right\}, \end{aligned}$$

Proof:

It is readily checked that for $y, y' \in E'$

$$S_t(y, y') = \int_0^t M_s(y) dM_s(y') + \int_0^t M_s(y') dM_s(y).$$

Then by the definition of the normalized integral

$$S_t(y, y') = \int_0^t (M_s(y)Q_s y' + M_s(y')Q_s y) * dM_s,$$

where

$$M_t(y) = \int_0^t Q_s y * dM_s.$$

Consequently, if f is a real valued predictable function

$$\int_0^\infty f_s^2 d\langle M(y) \rangle_s = \int_0^\infty |f_s Q_s y|_{H_s}^2 d\lambda_s,$$

$$\int_0^\infty f_s^2 d\langle S(y, y') \rangle_s = \int_0^\infty f_s^2 |M_s(y)Q_s y' + M_s(y')Q_s y|_{H_s}^2 d\lambda_s.$$

Thus

$$\begin{aligned} \cup_{N \in \mathcal{N}} \mathcal{Z}^1(N) &\subset \mathcal{A} = \{\mathcal{R}_t(f) = \int_0^t f_s * dM_s : f \in L_{loc}^2(Qd\lambda, \mathbf{P}), \\ &\quad \mathbf{E}^{\mathbf{P}} \int_0^\infty |f_s|_{H_s}^2 d\lambda_s < \infty\}, \end{aligned}$$

and then obviously $\mathcal{L}^1(\mathbf{P}, \mathcal{N}) \subset \mathcal{A}$. Now let f be an arbitrary predictable function such that $\mathbf{E}^{\mathbf{P}}(\int_0^\infty |f_s|_{H_s}^2 d\lambda_s)^{1/2} < \infty$. Set $f_s^n = \sum_{k=1}^n (f_s, e_s^k) e_s^k$. By Burkholder's inequality

$$\begin{aligned} \mathbf{E}^{\mathbf{P}} \sup_t |\mathcal{R}_t(f) - \mathcal{R}_t(f^n)| &\leq \\ &\leq 3\mathbf{E}^{\mathbf{P}} \left(\int_0^\infty |f_s - f_s^n|_{H_s}^2 d\lambda_s \right)^{1/2} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Recalling now that the construction of the basis $\{e_s^k\}$ we can find that for some N $f^n = \sum_{k=1}^N g_s^k e'_k$, where g_s^k are real valued predictable functions and $e'_k \in E'$. Let

$$\tau_m = \inf(t > 0 : \sup_k \langle M(e'_k) \rangle_t = \sup_k \int_0^t |Q_s e'_k|_{H_s}^2 d\lambda_s \geq m), \quad f_s^{n,m} = \sum_{k=1}^N g_s^{k,m} e'_k,$$

where $g_s^{k,m} = 1_{\{s \leq \tau_m, \sup_k |g_s^k| \leq m\}} g_s^k$. Then by Lebesgue dominated convergence theorem

$$\mathbf{E}^{\mathbf{P}} \left(\int_0^\infty |f_s^{n,m} - f_s^n|_{H_s}^2 d\lambda_s \right)^{1/2} \xrightarrow{m \rightarrow \infty} 0.$$

Thus $\mathbf{E}^{\mathbf{P}} \sup_t |\mathcal{R}_t(f^{n,m}) - \mathcal{R}_t(f^n)| \xrightarrow{m \rightarrow \infty} 0$. On the other hand, for each k

$$\mathbf{E}^{\mathbf{P}} \left(\int_0^\infty (g_s^{k,m})^2 |Q_s e'_k|_{H_s}^2 d\lambda_s \right)^{1/2} < \infty,$$

and therefore $\int_0^t g_s^{k,m} dM_s(e'_k) \in \mathcal{Z}^1(M(e'_k))$. Since:

$$\mathcal{R}_t(f^{n,m}) = \sum_{k=1}^N \int_0^t g_s^{k,m} dM_s(e'_k) = \sum_{k=1}^N \mathcal{R}_t(g_s^{k,m} Q e'_k),$$

this concludes the proof. ■

THEOREM 3.4. *We have $J^1(\mathbf{P}) = \left\{ \mathcal{R}_t(f) : \mathbf{E} \left(\int_0^\infty |f_s|_{H_s}^2 d\lambda_s \right)^{1/2} < \infty \right\}$ whenever \mathbf{P} is the only solution of the martingale problem $(\mu_0, dA, Qd\lambda, B)$.*

Proof:

The statement is an immediate corollary of Proposition 3.4 and general results on representation of martingales. Indeed, according to Theorems 11.2, 4.6 and Proposition 4.5 in [13], if \mathcal{N} is an arbitrary system of continuous processes and ν is the unique probability measure on (D, \mathcal{D}) such that every $N \in \mathcal{N}$ belongs to $\mathcal{M}_{loc}^c(\mathbb{D}, \nu)$, then $J^1(\nu) = \{M \in \mathcal{M}_{loc}(\mathbb{D}, P) : |M|_1 < \infty\}$ coincides with its stable subspace generated by \mathcal{N} . Thus the statement follows by Proposition 3.4. ■

THEOREM 3.5. (*Mikulevicius, Rozovskii [29]*). *Let \mathbf{P} be the only solution of the martingale problem $(\mu_0, dA, Qd\lambda, B)$ and $\tilde{\mathbf{P}}|_{\mathcal{D}_t} \prec \mathbf{P}|_{\mathcal{D}_t}$ for all $t > 0$. Then there exists $h \in L^2_{loc}(Qd\lambda, \tilde{\mathbf{P}})$ such that $\tilde{\mathbf{P}} \in \mathcal{S}(\mu_0, dA + hd\lambda, Qd\lambda, B)$. Let $Z_t = d\tilde{\mathbf{P}}/d\mathbf{P}|_{\mathcal{D}_{t+}}$, then there exists a sequence of stopping (τ_n) such that for each $t > 0$, $\tilde{\mathbf{P}}(\tau_n < t) \rightarrow 0$ as $n \rightarrow 0$ and*

$$Z_{t \wedge \tau_n} = \exp \left\{ \int_0^{t \wedge \tau_n} h_s * dM_s - 1/2 \int_0^{t \wedge \tau_n} |h_s|_s^2 d\lambda_s \right\}.$$

Proof:

Since Z_t is a (\mathbf{P}, \mathbb{D}) -martingale, there exists a sequence of stopping times $T_n \uparrow \infty$ \mathbf{P} -a.e. such that for every n , $Z_{t \wedge T_n} \in J^1(\mathbf{P})$ (In fact, the above statement holds for any $L \in \mathcal{M}_{loc}(\mathbb{D}, \mathbf{P})$, see [13]).

Due to Theorem 3.4, for every n , there is a function $f_n(s) \in L^2(Qd\lambda, \mathbf{P})$ so that $Z_{t \wedge T_n} = Z_0 + \int_0^t f_n(s) * dM_s$. Without loss of generality we can assume that $T_0 = 0$. Set

$$f(s) = \sum_{n=0}^{\infty} f_n(s) I_{\{s \in [T_n, T_{n+1})\}}.$$

So $f(s) = f_n(s)$ for $s \leq T_n$ and $\mathbf{P}(\int_0^t |f_s|_{H_s}^2 d\lambda_s < \infty) = 1$. So we proved that there exists a function $f \in L^2_{loc}(Qd\lambda, \mathbf{P})$ such that for all $t \geq 0$

$$Z_t = Z_0 + \int_0^t f(s) * dM_s.$$

Set $\tau_n = \inf(t : Z_t \leq \frac{1}{n})$ and

$$h_s = \begin{cases} f_s/Z_s, & \text{if } Z_s \neq 0 \\ 0, & \text{if } Z_s = 0 \end{cases}$$

Then

$$dZ_{t \wedge \tau_n} = h_t Z_{t \wedge \tau_n} * dM_{t \wedge \tau_n}$$

or, in other words,

$$Z_{t \wedge \tau_n} = \exp \left\{ \int_0^t h_s * dM_s - 1/2 \int_0^t |h_s|_{H_s}^2 d\lambda_s \right\}.$$

Fix $T > 0$. Let $\mathbf{P}^n = Z_{T \wedge \tau_n} X_{\cdot \wedge \tau_n \wedge T}(\mathbf{P})$. By Girsanov's theorem

$$(3.26) \quad \mathbf{P}^n \in \mathcal{S}(\tilde{\mu}_0, dA_{\cdot \wedge \tau_n \wedge T} + h d\lambda_{\cdot \wedge \tau_n \wedge T}, Q d\lambda_{\cdot \wedge \tau_n \wedge T}, B_{\cdot \wedge \tau_n \wedge T}).$$

On the other hand it is readily checked that

$$(3.27) \quad \mathbf{P}^n(A) = \tilde{\mathbf{P}}(A) \text{ for every } A \in \mathcal{D}_{\tau_n \wedge T}.$$

Since $\tilde{\mathbf{P}}(\tau_n < T) \xrightarrow{n \rightarrow \infty} 0$, we have $\tilde{\mathbf{P}}(\inf_{s \leq T} Z_s > 0) = 1$. Since $T > 0$ is arbitrary $h \in L^2_{loc}(Qd\lambda, \tilde{\mathbf{P}})$.

It follows from (3.26) that

$$\tilde{M}_{t \wedge \tau_n \wedge T}(y) = \langle X_{t \wedge \tau_n \wedge T}, y \rangle - \langle X_0, y \rangle - A_{t \wedge \tau_n}(y) - \int_0^{t \wedge \tau_n \wedge T} \langle h_s, y \rangle d\lambda_s$$

and

$$\tilde{S}_{t \wedge \tau_n \wedge T}(y) = \tilde{M}_{t \wedge \tau_n \wedge T}(y)^2 - \int_0^{t \wedge \tau_n \wedge T} \langle Q_s y, y \rangle d\lambda_s$$

belong to $\mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P}^n)$. In other words, there exists a sequence of stopping times (S_n) such that $\mathbf{P}(S_n < T) \rightarrow 0$, as $n \rightarrow \infty$, and for any $s \leq t$, $A \in \mathcal{D}_{s+}$

$$(3.28) \quad \begin{aligned} \int_A \tilde{M}_{t \wedge \tau'_n}(y) d\mathbf{P}^n &= \int_A \tilde{M}_{s \wedge \tau'_n}(y) d\mathbf{P}^n, \\ \int_A \tilde{S}_{t \wedge \tau'_n}(y) d\mathbf{P}^n &= \int_A \tilde{S}_{s \wedge \tau'_n}(y) d\mathbf{P}^n, \end{aligned}$$

where $\tau'_n = T \wedge \tau_n \wedge S_n$. Due to (3.27), equalities (3.28) hold true if $\tilde{\mathbf{P}}$ is substituted for \mathbf{P}^n on both sides. So we proved that $\tilde{\mathbf{P}} \in \mathcal{S}(\tilde{\mu}_0, dA + h d\lambda, Q d\lambda, B)$. ■

REMARK 3.1. Let $I = \cup_n [0, \tau_n]$, where $\tau_n = \inf(t > 0 : \int_0^t |h_s|_s^2 d\lambda_s \geq n)$. If the assumptions of Theorem 3.5 are satisfied we have by Theorem III.5.19 in [12] that the density process

$$Z_t = \begin{cases} Z_0 \exp\{\int_0^t h_s * dM_s - \frac{1}{2} \int_0^t |h_s|_s^2 d\lambda_s\}, & \text{if } t \in I, \\ 0, & \text{if } t \notin I. \end{cases}$$

Moreover by Corollary IV.1.35 in [12] the associated Hellinger process of order $1/2$ is

$$h(1/2)_t = 1/8 \int_0^t |h_s|_s^2 d\lambda_s.$$

Let $(\Omega^n, \mathcal{F}^n, \mathbf{P}^n)$ and $(\Omega^n, \mathcal{F}^n, \tilde{\mathbf{P}}^n)$ be two sequences of probability spaces. We say that the sequences of measures (\mathbf{P}^n) and $(\tilde{\mathbf{P}}^n)$ are entirely separated if there is a sequence $n_k \rightarrow \infty$ as $k \rightarrow \infty$ and for each k a set $A^{n_k} \in \mathcal{F}^{n_k}$ such that $\mathbf{P}^{n_k}(A^{n_k}) \rightarrow 1$ and $\tilde{\mathbf{P}}^{n_k}(A^{n_k}) \rightarrow 0$ as $k \rightarrow \infty$. From Theorem V.2.4 in [12] we have immediately by the Remark 3.1 the following statement.

THEOREM 3.6. Assume \mathbf{P}^n is a unique solution to $(\mu_0, dA^n, Q^n d\lambda^n, B^n)$, $\tilde{\mathbf{P}}^n \in \mathcal{S}(\mu'_0, dA^n + h^n d\lambda^n, Q^n d\lambda^n, B^n)$, $h^n \in L_{loc}^2(Q^n d\lambda^n, \tilde{\mathbf{P}}^n)$, $\mu_0 \sim \mu'_0$ and $\tilde{\mathbf{P}}^n|_{\mathcal{D}_t} \prec \mathbf{P}^n|_{\mathcal{D}_t}$, for all $t > 0$ and $n \geq 1$. Let (τ_n) be a sequence of stopping times and denote $\tilde{\mathbf{P}}_{\tau_n}^n, \mathbf{P}_{\tau_n}^n$ the restrictions of $\tilde{\mathbf{P}}^n, \mathbf{P}^n$ to \mathcal{D}_{τ_n} .

The sequences of measures $(\tilde{\mathbf{P}}_{\tau_n}^n)$ and $(\mathbf{P}_{\tau_n}^n)$ are entirely separated if for each $N > 0$

$$\limsup_n \tilde{\mathbf{P}}^n\left(\int_0^{\tau_n} |h_s^n|_s^2 d\lambda_s \geq N\right) = 1.$$

For the farther consideration of singularity of measures the following observation is basic.

REMARK 3.2. Let (X^n) be a sequence of measurable functions from the measurable space (Ω, \mathcal{F}) to the other measurable space (Ω', \mathcal{F}') and $\tilde{\mathbf{P}}, \mathbf{P}$ be two probability measures on (Ω, \mathcal{F}) .

If the sequences $X^n(\tilde{\mathbf{P}})$ and $X^n(\mathbf{P})$ are entirely separated, then $\tilde{\mathbf{P}}$ and \mathbf{P} are singular.

Proof:

By the definition of the entire separation we can find a sequence of measurable sets $A_{n_k} \in \mathcal{F}'$ such that $X^{n_k}(\tilde{\mathbf{P}})(A_{n_k}) \rightarrow 1$ and $X^{n_k}(\mathbf{P})(A_{n_k}) \rightarrow 0$ as $k \rightarrow \infty$. Thus we have a sequence of measurable sets $A^k = \{X^{n_k} \in A_{n_k}\}$ in Ω and $\tilde{\mathbf{P}}(A^k) \rightarrow 1$ and $\mathbf{P}(A^k) \rightarrow 0$ as $k \rightarrow \infty$, and the statement follows. ■

Let us apply now Theorems 3.5 and 3.3 to give a martingale characterization of all measures absolutely continuous with respect to the law of super Brownian motion.

COROLLARY 3.2. *Consider the Schwartz pair $(M_F(\mathbf{R}^d), C_b^2(\mathbf{R}^d))$ introduced in Example 3.1. Let \mathbf{P} be a solution of martingale problem (3.6)-(3.8) and $\tilde{\mathbf{P}}|_{\mathcal{D}_t} \prec \mathbf{P}|_{\mathcal{D}_t}$ for all $t > 0$. Then there exists a $\mathcal{P}(\mathbb{D}) \times \mathcal{B}(\mathbf{R}^d)$ -measurable function θ so that*

$$(3.29) \quad \tilde{\mathbf{P}}\left(\int_0^t \int_{\mathbf{R}^d} |\theta_s(X, u)|^2 X_s(du) ds < \infty \text{ for all } t < \infty\right) = 1,$$

and $\tilde{\mathbf{P}}$ is a unique solution to the martingale problem

$$(3.30) \quad X_0(\tilde{\mathbf{P}}) = \mu_0,$$

and for every $f, g \in E'$,

$$(3.31) \quad \begin{aligned} M_t(f) &= \int_{\mathbf{R}^d} f(u) X_t(du) - \int_{\mathbf{R}^d} f(u) X_0(du) - \\ &- \frac{1}{2} \int_0^t \int_{\mathbf{R}^d} (\Delta f(u) + \theta_s(X, u)) X_s(du) ds \in \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}}) \end{aligned}$$

and

$$(3.32) \quad M_t(f) M_t(g) - \gamma \int_0^t \int_{\mathbf{R}^d} f(u) g(u) X_s(du) ds \in \mathcal{M}_{loc}^c(\mathbb{D}, \tilde{\mathbf{P}}).$$

Let $Z_t = d\tilde{\mathbf{P}} / d\mathbf{P}|_{\mathcal{D}_{t+}}$. Then there exists a sequence of stopping (τ_n) such that for each $t > 0$, $\tilde{\mathbf{P}}(\tau_n < t) \rightarrow 0$ as $n \rightarrow 0$ and

$$Z_{t \wedge \tau_n} = \exp \left\{ \int_0^{t \wedge \tau_n} \theta_s dM_s - \gamma/2 \int_0^{t \wedge \tau_n} \int_{\mathbf{R}^d} |\theta_s(u)|^2 X_s(du) ds \right\}.$$

Proof:

Since $C_b^2(\mathbf{R}^d)$ is weakly dense in $B(\mathbf{R}^d)$, the space of bounded Borel functions on \mathbf{R}^d , we have by Propositions 2.8 and 2.9, where $q_s = \gamma X_s$ and $\lambda_s = s$, that $L_{loc}^2(Q, \tilde{\mathbf{P}})$ is the set of all $\mathcal{P}(\mathbb{D})$ -measurable measures on $\mathcal{B}(\mathbf{R}^d)$ of the form $\theta_s(X, u) \gamma X_s(du)$, where $\theta_s(X, u)$ are $\mathcal{P}(\mathbb{D}) \times \mathcal{B}(\mathbf{R}^d)$ -measurable functions, so that (3.29) holds and

$$|\theta_s \gamma X_s|_{H_s}^2 = |\theta_s|_{H'_s}^2 = \gamma \int_{\mathbf{R}^d} |\theta_s(X, u)|^2 X_s(du).$$

By Definition 2.4, $\int_0^t \theta_s \gamma X_s * dM_s = \int_0^t \theta_s dM_s$, and so

$$\begin{aligned} Z_{t \wedge \tau_n} &= \exp \left\{ \int_0^{t \wedge \tau_n} \theta_s \gamma X_s * dM_s - 1/2 \int_0^{t \wedge \tau_n} |\theta_s \gamma X_s|_s^2 ds \right\} \\ &= \exp \left\{ \int_0^{t \wedge \tau_n} \theta_s dM_s - \gamma/2 \int_0^{t \wedge \tau_n} \int_{\mathbf{R}^d} |\theta_s(u)|^2 X_s(du) ds \right\}. \end{aligned}$$

■

From Theorems 3.5 and 3.2 one can derive the following criterion for mutual absolute continuity of solutions to martingale problems.

COROLLARY 3.3. (*Mikulevicius, Rozovskii [29]*) Let \mathbf{P}_i be a solution to the martingale problem $(\mu_0, dA^i, Qd\lambda, B)$, $i = 1, 2$. Assume that condition U_{loc} holds at least for one of the above martingale problems. Then \mathbf{P}_1 and \mathbf{P}_2 are locally mutually absolutely continuous (i.e. $\mathbf{P}_1|_{\mathcal{D}_t} \sim \mathbf{P}_2|_{\mathcal{D}_t}$ for each $t > 0$) if and only if the following assumption holds:

(A) There is a process $h_s \in L^2_{loc}(Qd\lambda, \mathbf{P}_i)$ for $i = 1, 2$ so that for every $f \in E'$, $\mathbf{P}_1 + \mathbf{P}_2$ -a.s. $A.(f) = A^1.(f) - A^2.(f) \prec \lambda$ and $dA_t(f)/d\lambda_t = \langle h_t, f \rangle$ for $d\lambda d(\mathbf{P}_1 + \mathbf{P}_2)$ -a.e.

Moreover, if (A) holds,

$$\begin{aligned} d\mathbf{P}_1/d\mathbf{P}_2|_{\mathcal{D}_{t+}} &= \\ &= \exp\left\{\int_0^t h_s * dM_s - 1/2 \int_0^t |h_s|_s^2 d\lambda_s\right\} \end{aligned}$$

for all $t > 0$.

This result generalizes Theorem 7.19 in [16] and Theorem 3 in [52]. Corollaries 3.3 and 3.2 yield the following characterization of the measures locally equivalent to the law of super-Brownian motion.

COROLLARY 3.4. Let \mathbf{P} be a solution of martingale problem (3.6)-(3.8). A probability measure $\tilde{\mathbf{P}}|_{\mathcal{D}_t} \sim \mathbf{P}|_{\mathcal{D}_t}$ for all $t > 0$ if and only if there exists a $\mathcal{P}(\mathbb{D}) \times \mathcal{B}(\mathbf{R}^d)$ -measurable function θ so that

$$(3.33) \quad \int_0^t \int_{\mathbf{R}^d} |\theta_s(X, u)|^2 X_s(du) ds < \infty \text{ for all } t < \infty$$

$\mathbf{P} + \tilde{\mathbf{P}}$ -a.e. and $\tilde{\mathbf{P}}$ is a solution of martingale problem (3.30)-(3.32). In this case

$$d\tilde{\mathbf{P}}/d\mathbf{P}|_{\mathcal{D}_t} = \exp\left\{\int_0^t \theta_s dM_s - \gamma/2 \int_0^t \int_{\mathbf{R}^d} |\theta_s(u)|^2 X_s(du) ds\right\}.$$

Note that Corollary 3.1 gives a sufficient condition for (3.33) to hold $\tilde{\mathbf{P}} - a.s.$

3.2.3. Existence via Absolute Continuity. Girsanov's formula has been used routinely in stochastic ordinary differential equations to prove existence of solutions with measurable but not necessarily continuous drift. It turns out that this technique can be extended to various types of SPDE's. Corollary 3.1 is a good example of the absolute continuity approach to proving existence of a solution to an SPDE. Another typical example is given below; for more examples of this sort see [29], [59], and references therein.

Let U be a smooth bounded domain in \mathbf{R}^d . Consider the scale of Sobolev spaces $\mathcal{H}^s = \mathcal{H}^s(U)$, $s \in \mathbf{R}$, generated by $L^2(U)$ and the operator $I - \Delta$ where Δ is the Dirichlet Laplacian (see Appendix). The norm and the scalar product in $\mathcal{H}^s(H^0 = L^2(U))$ will be denoted by $|\cdot|_s$ and $(\cdot, \cdot)_s$, respectively. The duality between \mathcal{H}^1 and \mathcal{H}^{-1} will be denoted by $\langle \cdot, \cdot \rangle$.

For $i = 1, 2$, consider the following boundary problems:

$$\begin{aligned} (3.34) \quad du(t) &= A_t^i(u(t))dt + dW_t \\ u(0) &= u_0, u(t)|_{\partial U} = 0 \end{aligned}$$

where $A_t^1(u) = \sum_{i=1}^d \partial_{x_i}^2 u$, $A_t^1(u) = \sum_{i=1}^d \partial_{x_i}^2 u + f(t, x, u)$, and W_t is a standard cylindrical Brownian motion in $L^2(U)$. For $X \in D = D_{[0, \infty)}(\mathcal{H}^{-1-\alpha})$, write

$B_t(X) = \int_0^t |Q^{1/2}X_s|_1^2 ds$. Note that $B_t(X)$ is well defined for all $X \in D$, however it is finite only for $X \in H^{1-\alpha}$ almost everywhere.

PROPOSITION 3.5. *Let $\alpha > d/2$ and $u_0 \in \mathcal{H}^{-\alpha}$. Assume that there exist constants C_1, C_2 such that*

$$(3.35) \quad |f(t, x, v)|_0 \leq C_1 + C_2 |v|_{1-\alpha}^2 \text{ for all } v \in \mathcal{H}^{1-\alpha}.$$

Then there exist unique solutions \mathbf{P}^1 and \mathbf{P}^2 to the martingale problems:

$$(\delta_{u_0}, dA_t^1, Idt, B_t) \quad \text{and} \quad (\delta_{u_0}, dA_t^2, Idt, B_t),$$

respectively, in the Schwartz pair $(\mathcal{H}^{-1-\alpha}, \mathcal{H}^{1-\alpha})$. These solutions are locally equivalent, i.e. $\mathbf{P}^2|_{\mathcal{D}_t} \sim \mathbf{P}^1|_{\mathcal{D}_t}$, for all $t < \infty$.

Proof:

It is a standard fact that $Q = (I - \Delta)^{-\alpha}$ is a nuclear operator in \mathcal{H}^0 , and so by Example 2.2, $W^Q = Q^{1/2}W$ is an \mathcal{H}^0 -valued Brownian motion with covariance operator Q . Let $v_0 = Q^{1/2}u_0$. It was shown in Section 3.1 that the martingale problem $(\delta_{v_0}, dA_t^1, Qdt, B_t \circ Q^{-1/2})$ in the Schwartz pair $(\mathcal{H}^{-1}, \mathcal{H}^1)$ has a solution \mathbf{P}^1 supported by $L_{loc}^2((0, \infty); \mathcal{H}^1)$, and this solution is unique. Write $g(t, x, X_t) = Q^{1/2}f(t, x, Q^{-1/2}X_t)$ where $X \in L_{loc}^2((0, \infty); \mathcal{H}^1)$. It is readily checked (cf. Example 2.7) that the covariance space H_s for W^Q is \mathcal{H}^α and the adjoint covariance space $H'_s = \mathcal{H}^{-\alpha}$. Then

$$\begin{aligned} \int_0^t |g(s, x, X_s)|_{H_s}^2 ds &= \int_0^t |g(s, x, X_s)|_\alpha^2 ds \\ &= \int_0^t |f(s, x, Q^{-1/2}X_s)|^2 ds \leq C_1 + C_2 \int_0^t |X_s|_1^2 ds. \end{aligned}$$

For $X \in D = D_{[0, \infty)}(\mathcal{H}_w^{-1})$, write

$$\tau_n = \inf \left(t > 0 : \int_0^t |X_s|_1^2 ds \geq n \right),$$

$$\begin{aligned} \rho_t^n &= \exp \left\{ \int_0^{t \wedge \tau_n} g(s, x, X_s) * dW_s^Q - \frac{1}{2} \int_0^{t \wedge \tau_n} |g(s, x, X_s)|_{H_s}^2 ds \right\} \\ &= \exp \left\{ \int_0^{t \wedge \tau_n} Q^{1/2}f(s, x, Q^{-1/2}X_s) dW_s^Q - \frac{1}{2} \int_0^{t \wedge \tau_n} |f(s, x, Q^{-1/2}X_s)|^2 ds \right\} \end{aligned}$$

Let $\mathbf{P}^{2,n}$ be a probability measure on D such that $\mathbf{P}^{2,n}|_{\mathcal{D}_t} = \rho_t^n X_{\cdot \wedge \tau_n}(\mathbf{P}^1)|_{\mathcal{D}_t}$. By Girsanov's theorem we have that $\mathbf{P}^{2,n}$ is a solution to the martingale problem $(\delta_{v_0}, d\bar{A}_{t \wedge \tau_n}, Qd(t \wedge \tau_n), B_{t \wedge \tau_n} \circ Q^{-1/2})$ where $\bar{A}_{t \wedge \tau_n}(X) = \int_0^{t \wedge \tau_n} (\Delta X_s + g(s, x, X_s)) ds$. As we explained in Section 3.1 this implies that the canonical stochastic process X satisfies the evolution equation in \mathcal{H}^{-1}

$$X_{t \wedge \tau_n} = v_0 + \int_0^{t \wedge \tau_n} (\Delta X_s + g(s, x, X_s)) ds + Q^{1/2}W_{t \wedge \tau_n}, \quad \mathbf{P}^{2,n}\text{-a.s.}$$

It follows from [51] that in fact $X_{t \wedge \tau_n}$ is continuous in \mathcal{H}^0 $\mathbf{P}^{2,n}$ -a.s. and the following formula Ito holds

$$(3.36) \quad \begin{aligned} |X_{t \wedge \tau_n}|^2 &= |v_0|^2 + 2 \int_0^{t \wedge \tau_n} (\langle \Delta X_s, X_s \rangle + (g(s, x, X_s), X_s)) ds + \\ &\quad + 2 \int_0^{t \wedge \tau_n} X_s Q^{1/2} dW_s + \int_0^{t \wedge \tau_n} \text{tr } Q ds, \quad \mathbf{P}^{2,n}\text{-a.s.} \end{aligned}$$

Making use of the elementary inequality $ab \leq \frac{\varepsilon}{2}a^2 + \frac{1}{2\varepsilon}b^2$ and (3.35), we obtain

$$(3.37) \quad \begin{aligned} (g(s, x, X_s), X_s) &\leq \varepsilon |g(s, x, X_s)|^2 + C(\varepsilon) |X_s|^2 \\ &\leq \varepsilon(C_1 + C_2 |X_s|_1^2) + C |X_s|^2. \end{aligned}$$

Since by the definition of the norm $|\cdot|_1$ (see Appendix), $\langle \Delta X_s, X_s \rangle = -|X_s|_1^2 + |X_s|_0^2$, by choosing ε to be sufficiently small we derive from (3.37) that there exist constants $\delta > 0$ and $K < \infty$ so that for all s ,

$$(3.38) \quad \langle \Delta X_s, X_s \rangle + (g(s, x, X_s), X_s) \leq -\delta |X_s|_1^2 + K |X_s|^2.$$

Now it follows from (3.36), (3.38) and Burkholder inequality and Gronwall lemma that

$$\sup_n \int_D \left(\sup_s |X_{s \wedge \tau_n}|^2 + \int_0^{\tau_n} |X_s|_1^2 ds \right) d\mathbf{P}^{2,n} < \infty.$$

This implies of course that $\lim_{n \rightarrow \infty} \mathbf{P}^{2,n}(\tau_n < t) = 0$ for any t . Then by Theorem 3.1 there exists a probability measure $\tilde{\mathbf{P}}^2$ on (D, \mathcal{D}) supported by $L_{loc}^2((0, \infty); \mathcal{H}^1)$ so that $g \in L_{loc}^2(Qds, \tilde{\mathbf{P}}^2)$ and $\tilde{\mathbf{P}}^2 \in \mathcal{S}(\delta_{v_0}, d\bar{A}_t, Qdt, B_t \circ Q^{-1/2})$. In addition for every $t > 0$, $\tilde{\mathbf{P}}^2|_{\mathcal{D}_t} \prec \mathbf{P}^1|_{\mathcal{D}_t}$. Obviously, condition U_{loc} holds for the martingale problem $(\delta_{v_0}, d\bar{A}_t^1, Qdt, B_t \circ Q^{-1/2})$. Hence, by Theorem 3.3, $\tilde{\mathbf{P}}^2$ is the only solution of $(\delta_{v_0}, d\bar{A}_t, Qdt, B_t \circ Q^{-1/2})$.

Owing to (3.38), $g \in L_{loc}^2(Qds, \mathbf{P}^1)$ and so by Corollary 3.3, $\tilde{\mathbf{P}}^2|_{\mathcal{D}_t} \sim \mathbf{P}^1|_{\mathcal{D}_t}$ for every $t > 0$.

Now note that X is a weak solution to equation

$$dX_t = \bar{A}_t(X_t)dt + Q^{1/2}dW_t$$

$$X_0 = v_0, \quad X_{t \wedge \partial U} = 0$$

where $\bar{A}_t^2(X_t) = \Delta X_t + g(t, x, X_t)$ if and only if $Q^{-1/2}X_t$ is a weak solution to (3.34) for $i = 2$. Then existence and uniqueness a measure $\mathbf{P}^2 \in \mathcal{S}(\delta_{u_0}, A_t^2 dt, Idt, B_t)$ follows from existence and uniqueness of $\tilde{\mathbf{P}}^2 \in \mathcal{S}(\delta_{v_0}, \bar{A}_t dt, Qdt, B_t \circ Q^{-1/2})$. Similarly, $\mathbf{P}^2|_{\mathcal{D}_t} \sim P^1|_{\mathcal{D}_t}$ for every t , because $\tilde{\mathbf{P}}^2$ and \mathbf{P}^1 are locally equivalent. ■

3.3. Existence of Solutions: Compactness Method. Let $U \subset V \subset H$ be separable Banach spaces with dense injections, and let us assume that U and H are Hilbert spaces. Identifying H and its dual we have $U \subset V \subset H \subset V' \subset U'$. Assume that the norm in V is given by:

$$|u|_V = |u|_{1,V} + |u|_{2,V},$$

Denote V_1 (respectively V_2) the completion of the space V with respect to $|u|_{1,V}$ (respectively $|u|_{2,V}$) norm. Let us assume also that both spaces are reflexive. Notice

that V' is the set of all $l = g + h$, $g \in V'_1$, $h \in V'_2$ such that $|g|_{V'_1} + |h|_{V'_2} < \infty$ and the norm of l is defined by

$$|l|_{V'} = \inf_{l=g+h} (|g|_{V'_1} + |h|_{V'_2}).$$

REMARK 3.3. If the embedding $U \subset H$ is compact, then there is an orthonormal basis (e_i) in H such that its elements, e_i , are orthogonal in U as well.

Indeed, let I be the isomorphism from U' onto U . The unit ball B of H has an image $I(B)$ in U which is included in the unit ball of U . Therefore it is compact in H , and I defines an self-adjoint compact operator in H . Thus there exists an orthonormal basis (e_i) in H such that $I(e_i) = \lambda_i e_i$ with $\lambda_i \neq 0$. Also we have for each $u \in U$,

$$\lambda_i (u, e_i)_U = (u, I(e_i))_U = (u, e_i)_H = \langle u, e_i \rangle.$$

In particular for $u = e_j$, $j \neq i$,

$$(e_j, e_i)_U = (e_j, e_i)_H = 0.$$

Let $a : [0, \infty) \times V \rightarrow U'$, $Q : [0, \infty) \times V \rightarrow \mathcal{L}^+(U')$ be measurable functions. Denote by $D_{[0, \infty)}(U')$ the space of continuous U' -valued functions. Let $\mathcal{D}_t = \sigma(X_s, s \leq t)$, $\mathcal{D}_{t+} = (\mathcal{D}_s)_{s \geq t}$, $X_s = X_s(w) = w_s$, $w \in D_{[0, \infty)}(U')$.

In applications to SPDE's it is convenient to modify the general Definition 3.1 as follows.

DEFINITION 3.3. A probability measure \mathbf{P} on $D_{[0, \infty)}(U')$ is said to be a solution of the martingale problem (x, a, Q) ($x \in H$) if $X_s \in V$ $ds d\mathbf{P}$ -a.e. and $\int_0^t |a(s, X_s)|_{U'} ds < \infty$ \mathbf{P} -a.s. for all $t > 0$, and for all $u, v \in U$

$$X_t - x - \int_0^t a(s, X_s) ds = M_t \in \mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P}),$$

$$M_t(u)M_t(v) - \int_0^t \langle Q(s, X_s)u, v \rangle ds \in \mathcal{M}_{loc}^c(\mathbb{D}, \mathbf{P})$$

The set of all solutions to (x, a, Q) is denoted by $\mathcal{S}(x, a, Q)$.

REMARK 3.4. Let $B_t = \int_0^t |a(s, X_s)|_{U'} 1_{\{X_s \in V\}} ds + D_t^{-1}$, where

$$D_t = \begin{cases} 1, & \text{if } \int_0^t 1_{\{X_s \in V\}} ds = t \\ 0, & \text{otherwise} \end{cases}$$

Then $\mathbf{P} \in \mathcal{S}(x, a, Q)$ is a solution of the martingale problem

$$(\delta_x, a(t, X_t) 1_{\{X_t \in V\}} dt, Q(t, X_t) 1_{\{X_t \in V\}} dt, B)$$

in the sense of Definition 3.1. Note that the process B_t reducing (x, a, Q) to a martingale problem in the sense of Definition 3.1 can be chosen in many different ways.

For $u \in U$, denote $\varphi^u(s, X_s) = i \langle a(s, X_s), u \rangle - 1/2 \langle Q(s, X_s)u, u \rangle$

LEMMA 3.1. Let \mathbf{P} be a measure on $D_{[0, \infty)}(U')$ such that $X_s \in V$ $ds d\mathbf{P}$ -a.e. and

$$\int_0^t |a(s, X_s)|_{U'} ds < \infty \quad \mathbf{P}\text{-a.s. for all } t > 0.$$

Then $\mathbf{P} \in \mathcal{S}(x, a, Q)$ if and only if for each u from a dense vector subspace $\tilde{U} \subset U$

$$M_t^u = e^{i\langle X_t, u \rangle} - \int_0^t e^{i\langle X_s, u \rangle} \varphi^u(s, X_s) ds \in \mathcal{M}_{loc}^c(\mathbb{D}, P).$$

Proof:

If $\mathbf{P} \in \mathcal{S}(x, a, Q)$, then $M_t^u \in \mathcal{M}_{loc}^c(\mathbb{D}, P)$ by Ito formula. If $M^u \in \mathcal{M}_{loc}^c(\mathbb{D}, P)$ for all u , then by [12] for all u , $\langle M_t, u \rangle \in \mathcal{M}_{loc}^c(\mathbb{D}, P)$ and

$$M_t(u)M_t(v) - \int_0^t \langle Q(s, X_s)u, v \rangle ds \in \mathcal{M}_{loc}^c(\mathbb{D}, P).$$

According to the following statement, if \mathbf{P} is a solution to the martingale problem, then X_t is a weak solution of an evolution equation in U' .

LEMMA 3.2. *Let M_t be a U' -valued continuous local martingale on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with a right continuous filtration of σ -algebras $\mathbb{F} = (\mathcal{F}_t)$ such that for each $u \in U$*

$$\langle M_t, u \rangle^2 - \int_0^t \langle Q_s u, u \rangle ds \in \mathcal{M}_{loc}^c(\mathbb{F}, P),$$

where Q is a predictable $\mathcal{L}^+(U')$ -valued function. Assume that

$$\langle Q_s u, u \rangle = \int_E \langle \sigma_s(r), u \rangle^2 \kappa(dr),$$

where σ is U' -valued $\mathcal{P}(\mathbb{F}) \times \mathcal{E}$ -measurable function on $[0, \infty) \times \Omega \times E$ and κ is a nonnegative measure on the measurable space (E, \mathcal{E}) such that $Y = L^2(E, \mathcal{E}, \kappa)$ is separable.

Then there exists a cylindrical Wiener process W in Y (possibly in some extension of the probability space (Ω, \mathcal{F}, P)) such that

$$M_t = \int_0^t \sigma_s dW_s.$$

Proof:

Since M_t is U' -valued, we can assume that for all (s, ω)

$$\int_E |\sigma_s(r)|_{U'}^2 \kappa(dr) < \infty.$$

By the orthogonalization procedure (see (2.6)) we can find a sequence of predictable U -valued processes \tilde{u}_s^n such that $u_s^n = Q_s \tilde{u}_s^n$ is complete orthogonal system in $H_s = H_{Q_s}$. Moreover

$$(3.39) \quad |u_s^n|_{H_s}^2 \in \{1, 0\}.$$

As it was noticed, for each $u \in U$,

$$(3.40) \quad \langle M_t, u \rangle = \sum_n \int_0^t f_s^n dM_s^n,$$

where $f_s^n = \langle Q_s \tilde{u}_s^n, u \rangle$, $M_t^n = \int_0^t \tilde{u}_s^n dM_s$, i.e.

$$(3.41) \quad M_t^n \text{ are orthogonal and } \langle M^n \rangle_t = \int_0^t |u_s^n|_{H_s}^2 ds, n = 1, 2, \dots.$$

Under our assumptions

$$|u_s^n|_{H_s}^2 = |\langle \sigma_s, \tilde{u}_s^n \rangle|_Y^2,$$

and $y_s^n = \langle \sigma_s, \tilde{u}_s^n \rangle$, $n \geq 1$, is an orthogonal system in Y . Also, for each $u \in U$

$$\begin{aligned} |Q_s u - \sum_{n=1}^m f_s^n u_s^n|_{H_s}^2 &= \langle Q_s(u - \sum_{n=1}^m f_s^n \tilde{u}_s^n), u - \sum_{n=1}^m f_s^n \tilde{u}_s^n \rangle = \\ &= \int_E \langle \sigma_s(r), u - \sum_{n=1}^m f_s^n \tilde{u}_s^n \rangle^2 \kappa(dr) = \int_E (\langle \sigma_s(r), u \rangle - \sum_{n=1}^m f_s^n y_s^n)^2 \kappa(dr) \rightarrow 0, \end{aligned}$$

as $n \rightarrow \infty$, where

$$\begin{aligned} f_s^n &= \langle Q_s \tilde{u}_s^n, u \rangle = \int_E \langle \sigma_s(r), \tilde{u}_s^n \rangle \langle \sigma_s(r), u \rangle \kappa(dr) = \\ &= \int_E y_s^n(r) \langle \sigma_s(r), u \rangle \kappa(dr). \end{aligned}$$

Thus

$$(3.42) \quad \langle \sigma_s, u \rangle = \sum_{n=1}^{\infty} f_s^n y_s^n$$

in Y . For each (s, ω) , we define the projections

$$pr_s y = \sum_{n=1}^{\infty} (y_s^n, y)_Y y_s^n, \quad y \in Y,$$

onto the subspace of Y generated by $\{y_s^n, n \geq 1\}$. Now in order to complete the family $\{y_s^n, n \geq 1\}$ in Y , we fix a dense sequence $(\tilde{y}^n)_{n \leq 0}$ in Y . For $n \leq 0$, we define

$$\begin{aligned} y_s^0 &= \begin{cases} (\tilde{y}^0 - pr_s \tilde{y}^0) / |\tilde{y}^0 - pr_s \tilde{y}^0|_Y, & \text{if } |\tilde{y}^0 - pr_s \tilde{y}^0|_Y \neq 0 \\ 0, & \text{otherwise,} \end{cases} \\ \dots \\ y_s^{n-1} &= \begin{cases} (\tilde{y}^{n-1} - pr_s \tilde{y}^{n-1} - \sum_{i=n}^0 (\tilde{y}^{n-1}, y_s^i) y_s^i) / d_s^{n-1}, & \text{if } d_s^{n-1} \neq 0, \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

where $d_s^{n-1} = |\tilde{y}^{n-1} - pr_s \tilde{y}^{n-1} - \sum_{i=n}^0 (\tilde{y}^{n-1}, y_s^i) y_s^i|_Y$. So we construct the sequence (y_s^n) of predictable Y -valued functions such that for each (s, ω)

$\{y_s^n, n = 0, \pm 1, \pm 2, \dots\}$ is a complete orthogonal system in Y .

Let \tilde{W}_t^n , $n = 0, \pm 1, \pm 2, \dots$, be a sequence of independent standard Wiener processes (independent of \mathcal{F} as well). Here possibly the extension of probability space is needed. Define

$$W_t^n = M_t^n + \int_0^t 1_{\{|u_s^n|_{H_s}^2 = 0\}} d\tilde{W}_s^n, \quad n \geq 1,$$

$$W_t^n = \tilde{W}_t^n, \quad n \leq 0.$$

Since (3.39) and (3.41) hold, $(W_t^n)_{-\infty < n < \infty}$ is a sequence of independent standard Wiener processes. It allows to define a cylindrical Wiener process in Y by the equality

$$(3.43) \quad W_t(y) = \sum_{n=-\infty}^{\infty} \int_0^t g_s^n dW_s^n,$$

where $g_s^n = (y, y_s^n)_Y$, $y \in Y$. Indeed, since $\{y_s^n, n = 0, \pm 1, \pm 2, \dots\}$ is a complete orthogonal system in Y , $y = \sum_n g_s^n y_s^n$ and $|y|^2 = \sum_n (g_s^n)^2$. Thus

$$\langle W(y) \rangle_t = \sum_{n=-\infty}^{\infty} \int_0^t |g_s^n|^2 ds = |y|^2 t,$$

i.e. W is a cylindrical Wiener process in Y . We claim that

$$(3.44) \quad M_t = \int_0^t \sigma_s dW_s.$$

Indeed, by (3.40), (3.42)

$$\begin{aligned} \langle M_t, u \rangle &= \sum_{n=1}^{\infty} \int_0^t f_s^n dM_s^n = \sum_{n=1}^{\infty} \int_0^t f_s^n dW_s^n = \\ &= \int_0^t \langle \sigma_s, u \rangle dW_s \end{aligned}$$

and we are done. ■

Let us furnish $D_{[0,\infty)}(U')$ with the topology \mathcal{T}_1 of the uniform convergence on finite intervals. Let $D_{[0,\infty)}(H_w)$ be the set of H -valued weakly continuous functions with the topology \mathcal{T}_0 of the uniform weak convergence on finite intervals, written \mathcal{T}_0 . Let $L_{loc}^{p_1, p_2}(V)$ ($p_1, p_2 > 1$) be the set of V -valued functions f_s with finite seminorms

$$\int_0^t \left(|f_s|_{1,V}^{p_1} + |f_s|_{2,V}^{p_2} \right) ds < \infty, t > 0.$$

We denote by \mathcal{T}_2 the topology of weak convergence on finite intervals, i.e. the topology defined by the maps

$$f_s \rightarrow \int_0^t (\langle f_s, g_s + h_s \rangle) ds,$$

where $t > 0$, and g and h are V'_1 -valued (respectively V'_2 -valued) functions such that

$$\int_0^t \left(|g_s|_{V'_1}^{q_1} + |h_s|_{V'_2}^{q_2} \right) ds < \infty, t > 0,$$

$$1/p_1 + 1/q_1 = 1 = 1/p_2 + 1/q_2.$$

The set $L_{loc}^{p_1, p_2}(V)$ with the topology \mathcal{T}_2 on it will be denoted $L_{loc,w}^{p_1, p_2}(V)$. Consider on V a norm $|v|_c$ such that for the completion V_c of V with respect to this norm the embedding $V \subset V_c$ is compact (in particular, the case $|v|_c = |v|_H$ is not excluded), and the embedding $V_c \subset H$ is continuous. Let $L_{loc}^{p_1 \wedge p_2}(V_c)$ be the set of locally $L^{p_1 \wedge p_2}$ -integrable V_c -valued functions with the topology \mathcal{T}_3 of $L^{p_1 \wedge p_2}$ -convergence on finite intervals. We will need a characterization of compact subsets of the intersection of all these spaces.

LEMMA 3.3. (cf. [46]) *Assume that the embedding $V \subset H$ is compact. Let $\mathcal{Z} = D_{[0,\infty)}(U') \cap D_{[0,\infty)}(H_w) \cap L_{loc,w}^{p_1, p_2}(V) \cap L_{loc}^{p_1 \wedge p_2}(V_c)$ and \mathcal{T} be the supremum of corresponding topologies.*

Then $K \subset \mathcal{Z}$ is \mathcal{T} -relatively compact if for each $t > 0$

$$\sup_{x \in K} \sup_{s \leq t} |x_s|_H < \infty, \sup_{x \in K} \int_0^t \left(|x_s|_{1,V}^{p_1} + |x_s|_{2,V}^{p_2} \right) ds < \infty,$$

$$\lim_{\delta \rightarrow 0} \sup_{x \in K} \sup_{\substack{|t-s| \leq \delta, \\ s, t \leq T}} |x_t - x_s|_{U'} = 0.$$

Proof:

It can be assumed that K is closed in \mathcal{T} . The topologies T_0, T_1, T_2, T_3 are metrizable on K . Consider a sequence (x^n) in K . Since the embedding $H \subset V'$ is also compact, there exist a subsequence (x^{n_k}) , $x \in D_{[0,\infty)}(U') \cap L_{loc,w}^{p_1,p_2}(V)$ such that $x^{n_k} \rightarrow x$ in $D_{[0,\infty)}(U') \cap L_{loc,w}^{p_1,p_2}(V)$ with respect to the supremum of T_1 and T_2 . Since for each $u \in U$ $\langle x^{n_k}, u \rangle_H = \langle x^{n_k}, u \rangle$ is uniformly convergent on finite intervals to $\langle x., u \rangle$, we have

$$\sup_{s \leq t} \sup_{|u|_H \leq 1} \langle x_s, u \rangle \leq \sup_{x \in K} \sup_{s \leq t} |x_s|_H < \infty,$$

and x is H -valued. Thus $x^{n_k} \rightarrow x$ in $D_{[0,\infty)}(H_w)$ as well. Also, for each $t > 0$, $p > 0$

$$(3.45) \quad \int_0^t |x_s^{n_k} - x_s|_{U'}^p ds \rightarrow 0$$

as $k \rightarrow \infty$. Let $p_2 \geq p_1 > 1$. We claim that for each $\epsilon > 0$ there is a constant C_ϵ such that for all $u \in V$

$$(3.46) \quad |u|_c^{p_1} \leq \epsilon |u|_V^{p_1} + C_\epsilon |u|_{U'}^{p_1}.$$

Indeed, if (3.46) does not hold there exists $\epsilon > 0$ and a sequence $u_n \in U$ such that

$$|u_n|_c^{p_1} > \epsilon |u_n|_V^{p_1} + n |u_n|_{U'}^{p_1}.$$

Then for $v_n = |u_n|_c^{-1} u_n$ we have

$$1 > \epsilon |v_n|_V^{p_1} + n |v_n|_{U'}^{p_1}.$$

Thus (v_n) is a bounded sequence in V and $v_n \rightarrow 0$ in U' as $n \rightarrow \infty$. Since the embedding $V \rightarrow V_c$ is compact, $v_n \rightarrow 0$ in V_c . On the other hand $|v_n|_c = 1$ for all n and we have a contradiction. Thus (3.46) holds. Since

$$\sup_n \int_0^t \left(|x_s^n|_{1,V}^{p_1} + |x_s^n|_{2,V}^{p_2} + |x_s|_{1,V}^{p_1} + |x_s|_{2,V}^{p_2} \right) ds < \infty,$$

we have $\sup_n \int_0^t (|x_s^n|_V^{p_1} + |x_s|_V^{p_1}) ds < \infty$. It follows by (3.45), (3.46)

$$\int_0^t |x_s^{n_k} - x_s|_{V_c}^{p_1} ds \rightarrow 0$$

for each $t > 0$. ■

The following statement is true.

THEOREM 3.7. *Assume that the embedding $V \rightarrow H$ is compact and*

(i) there exists a measurable function $\tilde{a} : [0, \infty) \times V \rightarrow V'$ such that for all $u \in U$,

$$\langle a(s, u), u \rangle = \langle \tilde{a}(s, u), u \rangle;$$

(ii) there exist $\delta > 0$, measurable non-negative functions R_t and g_t such that R_t is increasing and g_t is locally integrable and for all $t, v \in V$,

$$2 \langle \tilde{a}(t, v), v \rangle + tr_H Q(t, v) + \delta(|v|_{1,V}^{p_1} + |v|_{2,V}^{p_2}) \leq g_t(1 + |v|_H^2),$$

$$tr_H Q(s, v) \leq g_t(1 + |v|_H^2) + R_t(|v|_{1,V}^{p_1} + |v|_{2,V}^{p_2}),$$

$$\text{where } tr_H Q(s, v) = \sum_{i=1}^{\infty} \langle Q(s, v) e_i, e_i \rangle;$$

(iii) for each $K > 0, T > 0, \epsilon > 0$ there exists an integrable function on $[0, T]$, written $g^{K, T, \epsilon}$, such that for all $t \leq T$,

$$\sup_{|v|_H \leq K} (|a(t, v)|_{U'} + \text{tr}_U Q(s, v)) \leq g_t^{K, T, \epsilon} + \epsilon(|v|_{1, V}^{p_1} + |v|_{2, V}^{p_2}),$$

where $\text{tr}_U Q(s, v) = \sum_i \langle Q(s, v) \tilde{e}_i, \tilde{e}_i \rangle$ and (\tilde{e}_i) is a CONS in U ;

(iv) for each $u \in \tilde{U}$, there are V' -valued function $f^u(s, v), h^u(s, v)$ on $[0, \infty) \times V$ and a measurable function g^u on $[0, \infty) \times V$ such that

$$\langle a(t, v), u \rangle = \langle f^u(t, v) + h^u(t, v), v \rangle + g^u(t, v),$$

$$\int_0^t \sup_{|v|_H \leq K} (|f^u(s, v)|_{V'_1}^{q_1} + |h^u(s, v)|_{V'_2}^{q_2}) ds < \infty,$$

and the mappings $v \rightarrow g^u(t, v), v \rightarrow \langle Q(t, v) u, u \rangle$ are continuous with respect to the V_c -norm on V , and $v \rightarrow f^u(s, v), v \rightarrow h^u(s, v)$ are continuous from V_c to V'_1 (respectively to V'_2).

Then, for each $u_0 \in H$, there exists $\mathbf{P} \in \mathcal{S}(u_0, a, Q)$ concentrated on

$$\mathcal{Z} = D_{[0, \infty)}(U') \cap D_{[0, \infty)}(H_w) \cap L_{loc, w}^{p_1, p_2}(V) \cap L_{loc}^{p_1 \wedge p_2}(V_c).$$

Proof:

In the first part of the proof, we construct the approximating sequence of measures. In the second, we prove the relative compactness of this sequence. Then we prove that any limiting point $\mathbf{P} \in \mathcal{S}(u_0, a, Q)$.

1⁰. Approximating sequence of measures. Let (e_n) be a CONS in H which is orthogonal in U . Such (e_n) exists by Remark 3.3. Define a sequence of functions on \mathbf{R}^n

$$a_l^i(s, y) = \left\langle a(s, \sum_{j=1}^m y^j e_j), e_i \right\rangle \varphi_l(y),$$

$$q_l^{ik}(s, y) = \left\langle Q(s, \sum_{j=1}^m y^j e_j) e_i, e_k \right\rangle \varphi_l(y), l = 1, 2, \dots, y = (y^j)_{1 \leq j \leq n} \in \mathbf{R}^n,$$

where φ_l is a smooth function on \mathbf{R}^n such that $0 \leq \varphi_l \leq 1, \varphi_l(y) = 1$, if $|y| \leq l$. It follows from the assumptions (ii)-(iv) that for each l , a_l^i, q_l^{ik} are bounded and continuous in y . Then by Theorem 2.31 of Chapter IX in [12] for each $y_0 = (y_l^j) \in \mathbf{R}^n$ there exists a measure $\tilde{\mathbf{P}}^{n,l}$ on $D_{[0, \infty)}(\mathbf{R}^n)$ of continuous \mathbf{R}^n -valued functions such that

$$Y_t^i - y_0^i - \int_0^t a_l^i(s, Y_s) ds = M_t^i \in \mathcal{M}_{loc}^c(\mathbb{D}^n, \tilde{\mathbf{P}}^{n,l}), i = 1, \dots, n,$$

$$\langle M^i, M^k \rangle_t = \int_0^t q_l^{ik}(Y_s) ds, i, k = 1, \dots, n,$$

where $Y_t = (Y_t^i)_{1 \leq i \leq n} = Y_t(w) = w_t, w \in D_{[0, \infty)}(\mathbf{R}^n), \mathbb{D}^n = (\mathcal{D}_{t+}^n), \mathcal{D}_t^n = \sigma(Y_s, s \leq t)$.

By Ito formula we have

$$\sum_{i=1}^m (Y_t^i)^2 = \sum_{i=1}^m (Y_0^i)^2 + \int_0^t [2 \left\langle a(s, \sum_{j=1}^m Y_s^j e_j), \sum_{i=1}^m Y_s^i e_i \right\rangle \varphi_l(Y_s) ds +$$

$$+ \sum_{i=1}^m \left\langle Q(s, \sum_{j=1}^m Y_s^j e_j) e_i, e_i \right\rangle \varphi_l(Y_s) ds + 2 \sum_{i=1}^n \int_0^t Y_s^i dM_s^i].$$

By (ii) and Gronwall's lemma for each $t > 0$

$$\sup_{n,l} \sup_{r \leq t} \left(\sum_{i=1}^n |Y_r^i|^2 + \int_0^t \left[\left| \sum_{i=1}^n Y_s^i e_i \right|_{1,V}^{p_1} + \left| \sum_{i=1}^n Y_s^i e_i \right|_{2,V}^{p_2} \right] ds \right) < \infty.$$

Also, we have the following estimate

$$\begin{aligned} \left\langle \sum_{i=1}^n \int_0^t Y_s^i dM_s^i \right\rangle &= \int_0^t \sum_{i,k=1}^n Y_s^k Y_s^i q_l^{ki} ds \leq \\ &\leq \sup_{s \leq t} \sum_{i=1}^n |Y_s^i|^2 \int_0^t \sum_{i=1}^n q_l^{ii}(s, Y_s) ds. \end{aligned}$$

Since

$$\mathbf{E} \sup_{s \leq t} \left| \sum_{i=1}^n \int_0^s Y_s^i dM_s^i \right| \leq C \mathbf{E} \left[\left\langle \sum_{i=1}^n \int_0^t Y_s^i dM_s^i \right\rangle^{1/2} \right]$$

and $\sum_{i=1}^n q_l^{ii}(s, Y_s) \leq \text{tr } Q(s, Y_s)$, this together with assumption (ii) implies that

$$(3.47) \quad \sup_{n,l} \tilde{\mathbf{P}}^{n,l} \sup_{s \leq t} \sum_{i=1}^n |Y_s^i|^2 < \infty.$$

Let $L > 0$, $\tau_L = \inf(t : |Y_t| > L)$, $Y_t^L = Y_{t \wedge T_L}$. Then for each $\eta, T > 0$ and a sequence (T_l) ($T_l \leq T$) of \mathbb{D}^n -stopping times and a sequence of numbers $1 > \delta_l \downarrow 0$

$$\begin{aligned} \limsup_l \tilde{\mathbf{P}}^{n,l} (|Y_{T_l+\delta_l} - Y_{T_l}| > \eta) &\leq \\ &\leq \limsup_l \tilde{\mathbf{P}}^{n,l} (|Y_{T_l+\delta_l} - Y_{T_l}| > \eta, \tau_L \geq T+1) + \\ &+ \limsup_l \tilde{\mathbf{P}}^{n,l} (\tau_L < T+1) \leq \limsup_l \tilde{\mathbf{P}}^{n,l} (|Y_{T_l+\delta_l}^L - Y_{T_l}^L| > \eta) + \\ &+ \limsup_l \tilde{\mathbf{P}}^{n,l} (\tau_L < T+1) \leq C/L^2. \end{aligned}$$

Thus $\lim_l \tilde{\mathbf{P}}^{n,l} (|Y_{T_l+\delta_l} - Y_{T_l}| > \eta) = 0$ and by Aldous criterion of the relative compactness of measures (see Theorem 4.5 of Chapter VI in [12]) the sequence $\{\tilde{\mathbf{P}}^{n,l}, l \geq 1\}$ is relatively compact. Denote $\tilde{\mathbf{P}}^n$ a limiting point of this sequence. It is standard to show that

$$\begin{aligned} Y_t^i - Y_0^i - \int_0^t \left\langle a(s, \sum_{j=1}^n Y_s^j e_j), e_i \right\rangle ds &= M_t^i \in \mathcal{M}_{loc}^c(\mathbb{D}^n, \tilde{\mathbf{P}}^n), i = 1, \dots, n, \\ \langle M^i, M^k \rangle_t &= \int_0^t \left\langle Q(s, \sum_{j=1}^n Y_s^j e_j) e_i, e_k \right\rangle ds, i, k = 1, \dots, n, \end{aligned}$$

Again by Ito formula we have

$$\begin{aligned} \sum_{i=1}^n (Y_t^i)^2 &= \sum_{i=1}^n (y_0^i)^2 + \int_0^t [2 \left\langle a(s, \sum_{j=1}^n Y_s^j e_j), \sum_{i=1}^n Y_s^i e_i \right\rangle ds + \\ &+ \sum_{i=1}^m \left\langle Q(s, \sum_{j=1}^n Y_s^j e_j) e_i, e_i \right\rangle ds + 2 \sum_{i=1}^n \int_0^t Y_s^i dM_s^i]. \end{aligned}$$

This together with the assumption (ii) for Q and (3.47) gives the estimate

$$(3.48) \quad \sup_n \tilde{\mathbf{P}}^n \left\{ \sup_{s \leq t} \sum_{i=1}^n |Y_s^i|^2 + \int_0^t \left(\left| \sum_{i=1}^n Y_s^i e_i \right|_{1,V}^{p_1} + \left| \sum_{i=1}^n Y_s^i e_i \right|_{2,V}^{p_2} \right) ds \right\} < \infty.$$

Let $\mathbf{P}^n = (\sum_{i=1}^n Y^i e_i) \left(\tilde{\mathbf{P}}^n \right)$ be the measure induced by $\sum_{i=1}^n Y^i e_i$ on $D_{[0,\infty)}(U')$. For each $u \in U$

$$(3.49) \quad M_t^{n,u} = e^{i\langle X_t, u \rangle} - \int_0^t e^{i\langle X_s, u \rangle} \varphi_n^u(s, X_s) ds \in \mathcal{M}_{loc}^c(\mathbf{P}^n, \mathbb{D}),$$

where $\varphi_n^u(s, X_s) = i \langle a(s, X_s), \Pi^n u \rangle - 1/2 \langle Q(s, X_s) \Pi^n u, \Pi^n u \rangle$, and Π^n is the projection of U onto the subspace generated by $(e_i)_{1 \leq i \leq n}$. Also,

$$(3.50) \quad \begin{aligned} X_t - \int_0^t \Pi^{n*} a(s, X_s) ds &= M_t^n \in \mathcal{M}_{loc}^c(\mathbf{P}^n, \mathbb{D}), \\ \langle M_t^n, u \rangle^2 - \int_0^t \langle Q(s, X_s) \Pi^n u, \Pi^n u \rangle ds &\in \mathcal{M}_{loc}^c(\mathbf{P}^n, \mathbb{D}), \quad u \in U. \end{aligned}$$

By (3.48) we have

$$(3.51) \quad \sup_n \mathbf{P}^n \left\{ \sup_{s \leq t} |X_s|_H^2 + \int_0^t (|X_s|_{1,V}^{p_1} + |X_s|_{2,V}^{p_2}) \right\} < \infty.$$

2⁰. Relative compactness of (\mathbf{P}^n) . For each $T > 0$, $\eta > 0$ and any sequence of stopping times (τ_n) such that $\tau_n \leq T$ and sequence of numbers $1 > \delta_n \downarrow 0$, we have by (3.51) and assumption (iii)

$$\mathbf{P}^n \left(\int_{\tau_n}^{\tau_n + \delta_n} \text{tr}_U Q(t, X_t) dt > \eta \right) \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

By Lenglart's inequality we have then

$$(3.52) \quad \mathbf{P}^n (|M_{\tau_n + \delta_n}^n - M_{\tau_n}^n|_{U'} > \eta) \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Let $K > 0$, $\epsilon > 0$, $T_K = \inf(t : |X_t|_H > K)$. Then by (iii)

$$\begin{aligned} (3.53) \quad & \left| \int_{\tau_n \wedge T_K}^{(\tau_n + \delta_n) \wedge T_K} \Pi^{n*} a(s, X_s) ds \right|_{U'} \leq \int_{\tau_n \wedge T_K}^{(\tau_n + \delta_n) \wedge T_K} |\Pi^{n*} a(s, X_s)|_{U'} ds \leq \\ & \leq \int_{\tau_n \wedge T_K}^{(\tau_n + \delta_n) \wedge T_K} |a(s, X_s)|_{U'} ds \leq \\ & \leq \int_{\tau_n \wedge T_K}^{(\tau_n + \delta_n) \wedge T_K} \left\{ g_t^{K,T+1,\epsilon} + \epsilon \left(|X_t|_{1,V}^{p_1} + |X_t|_{2,V}^{p_2} \right) \right\} dt. \end{aligned}$$

On the other hand by (3.51)

$$\sup_n \mathbf{P}^n (T_K < T + 1) \rightarrow \infty, \quad \text{as } K \rightarrow \infty,$$

and from (3.53) it follows then that

$$\mathbf{P}^n \left(\left| \int_{\tau_n}^{\tau_n + \delta_n} a(s, X_s) ds \right|_{U'} > \eta \right) \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

This and (3.52) implies that

$$\mathbf{P}^n (|X_{\tau_n + \delta_n} - X_{\tau_n}|_{U'} > \eta) \rightarrow 0$$

as $n \rightarrow \infty$. Since the embedding $H \rightarrow U'$ is compact, we have by Aldous criterion (see Theorem 3.2 in [45]) the relative compactness of measures $\{\mathbf{P}^n, n \geq 1\}$ on $D_{[0,\infty)}(U')$. Thus for each $T > 0$, $\eta > 0$

$$\lim_{\delta \rightarrow 0} \sup_n \mathbf{P}^n \left(\sup_{s,t \leq T} \frac{|s-t| \leq \delta}{s,t} |X_t - X_s|_{U'} > \eta \right) = 0.$$

Therefore the relative compactness of measures $\{\mathbf{P}^n, n \geq 1\}$ on $\mathcal{Z} = D_{[0,\infty)}(U') \cap D_{[0,\infty)}(H_w) \cap L_{loc,w}^{p_1, p_2}(V) \cap L_{loc}^{p_1 \wedge p_2}(V_c)$ with supremum topology \mathcal{T} follows by Lemma 3.3.

3⁰. Identification of the limit measures. It follows from the relative compactness that we can assume that a sequence of measures (\mathbf{P}^n) converges weakly to a measure \mathbf{P} on \mathcal{Z} . Let $u \in \tilde{U}$ and $M_t^u = M_t^u(X) = e^{i\langle X_t, u \rangle} - \int_0^t e^{i\langle X_s, u \rangle} \varphi^u(s, X_s) ds$, where

$$\varphi^u(s, X_s) = i \langle a(s, X_s), u \rangle - 1/2 \langle Q(s, X_s)u, u \rangle.$$

Let $\omega^n \rightarrow \omega$ in \mathcal{Z} . Then for each t

$$(3.54) \quad \begin{aligned} & \sup_n \{ \sup_{s \leq t} (|\omega_s^n|_H + |\omega_s|_H) + \\ & + \int_0^t \left(|\omega_s^n|_{1,V}^{p_1} + |\omega_s^n|_{2,V}^{p_2} + |\omega_s|_{1,V}^{p_1} + |\omega_s|_{2,V}^{p_2} \right) ds < \infty, \text{ and} \\ & \int_0^t |\omega_s^n - \omega_s|_{V_c}^2 ds + \sup_{s \leq t} |\langle \omega_s^n - \omega_s, u \rangle| \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

This and assumptions (iii), (iv) imply that the sequence $(M_s^u(\omega^n))$ is equicontinuous with respect to n on finite time intervals. Also, for each t

$$(3.55) \quad e^{i\langle \omega_t^n, u \rangle} \varphi^u(t, \omega_t^n) \rightarrow e^{i\langle \omega_t, u \rangle} \varphi^u(t, \omega_t).$$

Let $F_1(t, \omega_t^n) = \operatorname{Re} e^{i\langle \omega_t^n, u \rangle} \varphi^u(t, \omega_t^n)$, $F_2(t, \omega_t^n) = \operatorname{Im} e^{i\langle \omega_t^n, u \rangle} \varphi^u(t, \omega_t^n)$. It follows from (iii), (iv) that for each $K > \sup_n \{ \sup_{s \leq t} (|\omega_s^n|_H + |\omega_s|_H) \}$, $T > 0$ and for each $\epsilon > 0$ there exists an integrable function $g_t^{K,T,\epsilon}$ such that for all $t \leq T$

$$(3.56) \quad |F_1(t, \omega_t^n)| + |F_2(t, \omega_t^n)| \leq g_t^{K,T,\epsilon} + \epsilon \left(|\omega_t^n|_{1,V}^{p_1} + |\omega_t^n|_{2,V}^{p_2} \right).$$

Thus for $k = 1, 2$, $s \in [0, T]$

$$F_k(s, \omega_s^n) - \epsilon \left(|\omega_s^n|_{1,V}^{p_1} + |\omega_s^n|_{2,V}^{p_2} \right) \leq g_s^{K,T,\epsilon},$$

$$F_k(s, \omega_s^n) + \epsilon \left(|\omega_s^n|_{1,V}^{p_1} + |\omega_s^n|_{2,V}^{p_2} \right) \geq -g_s^{K,T,\epsilon}$$

and by Fatou lemma and (3.55) we have for each $t \leq T$

$$\begin{aligned} & \int_0^t F_k(s, \omega_s) ds \leq \int_0^t \liminf_n \left[F_k(s, \omega_s^n) + \epsilon \left(|\omega_s^n|_{1,V}^{p_1} + |\omega_s^n|_{2,V}^{p_2} \right) \right] ds \leq \\ & \leq \liminf_n \int_0^t F_k(s, \omega_s^n) ds + \epsilon \sup_n \int_0^t \left(|\omega_s^n|_{1,V}^{p_1} + |\omega_s^n|_{2,V}^{p_2} \right) ds, \\ & \limsup_n \int_0^t \left(F_k(s, \omega_s^n) - \epsilon \left(|\omega_s^n|_{1,V}^{p_1} + |\omega_s^n|_{2,V}^{p_2} \right) \right) ds \leq \\ & \leq \int_0^t F_k(s, \omega_s) ds. \end{aligned}$$

Since ϵ and T are arbitrary it follows by (3.54) that for each t

$$\int_0^t F_k(s, \omega_s) ds \rightarrow \int_0^t F_k(s, \omega_s) ds,$$

i.e.

$$|M_t^u(\omega^n) - M_t^u(\omega)| \rightarrow 0$$

as $n \rightarrow \infty$. Thus for each t

$$(3.57) \quad \sup_{s \leq t} |M_s^u(\omega^n) - M_s^u(\omega)| \rightarrow 0$$

Let $N > 0$, $\tau_N = \inf(t : |X_t|_H > N)$. We have obviously for each $t > 0$, $\eta > 0$, $N > 0$

$$\begin{aligned} & \mathbf{P}^n \left(\sup_{s \leq t} |M_s^{n,u} - M_s^u| > \eta \right) \leq \\ & \leq \mathbf{P}^n \left(|u - \Pi^n u|_U \int_0^{t \wedge \tau_N} |a(s, X_s)|_{U'} ds > \eta/2 \right) + \\ & + \mathbf{P}^n \left(2 |u - \Pi^n u|_U |u|_U \int_0^{t \wedge \tau_N} \text{tr}_U Q(s, X_s) ds > \eta/2 \right) + \\ & + \mathbf{P}^n (\tau_N < t). \end{aligned}$$

Thus by assumption (iii) and (3.51) for each t

$$\begin{aligned} & \limsup_n \mathbf{P}^n \left(\sup_{s \leq t} |M_s^{n,u} - M_s^u| > \eta \right) \leq \\ & \leq \limsup_n \mathbf{P}^n (\tau_N < t) \rightarrow 0, \end{aligned}$$

as $N \rightarrow \infty$, i.e. for each $t > 0$, $\eta > 0$

$$(3.58) \quad \lim_n \mathbf{P}^n \left(\sup_{s \leq t} |M_s^{n,u} - M_s^u| > \eta \right) = 0.$$

For $L > 0$ define

$$\tau_L = \inf(t : |M_t^u| > L).$$

Let $\tau^n = \inf(t : |M_t^{n,u} - M_t^u| > 1)$, $\tau_L^n = \tau_L \wedge \tau^n$. Then for each $t > 0$ we have by (3.58)

$$(3.59) \quad \mathbf{P}^n (\tau^n < t) \leq \mathbf{P}^n \left(\sup_{s \leq t} |M_s^{n,u} - M_s^u| > 1 \right) \rightarrow 0,$$

as $n \rightarrow \infty$. Also obviously

$$(3.60) \quad \sup_{n,s} |M_{s \wedge \tau_L^n}^{n,u}| \leq L + 1.$$

Let $\mathbf{P}(\tau_L = \tau_{L-}) = 1$, where $\tau_{L-} = \lim_{\delta \downarrow 0} \tau_{L-\delta}$. If $\tau_L(\omega) = \tau_{L-}(\omega)$ and $\omega^n \rightarrow \omega$ in \mathcal{Z} , we have by (3.57) for each $T > 0$

$$(3.61) \quad \sup_{s \leq T} |M_{s \wedge \tau_L(\omega^n)}^u(\omega^n) - M_{s \wedge \tau_L(\omega)}^u(\omega)| \rightarrow 0.$$

Let $s \leq t$ and f be a bounded \mathcal{D}_s -measurable \mathbf{P} -a.s continuous function. Then by (3.58)-(3.61)

$$0 = \mathbf{P}^n \left[f \left(M_{t \wedge \tau_L^n}^{n,u} - M_{s \wedge \tau_L^n}^{n,u} \right) \right] \rightarrow \mathbf{P} [f(M_{t \wedge \tau_L}^u - M_{s \wedge \tau_L}^u)].$$

The statement is proved. ■

3.4. Examples.

3.4.1. *Cauchy-Dirichlet problem.* Let W be a cylindrical Wiener process in a separable Hilbert space $Y = L^2(E, \mathcal{E}, \kappa)$ where κ is a non-negative measure on the measurable space (E, \mathcal{E}) . Let G be a smooth bounded domain in \mathbf{R}^d . Consider

$$(3.62) \quad \begin{cases} \partial_t u(t, x) = \partial_i (a^{ij}(t, x, u(t, x)) \partial_j u(t, x)) + b^i(t, x, u(t, x)) \partial_i u(t, x) + \\ + \partial_i [f^i(t, x, u(t, x))] + f(t, x, u(t, x)) + \\ (\partial_j [\sigma^j(t, x, u(t, x))] + g(t, x, u(t, x))) \partial_t W, \\ u(t, x)|_{\partial G} = 0, \\ u(0, x) = u_0(x), \end{cases}$$

where a^{ij} , b^i , f^i , f are measurable functions on $[0, \infty) \times G \times \mathbf{R}$, and σ^j , g are Y -valued measurable functions on $[0, \infty) \times G \times \mathbf{R}$. We assume always that the matrix (a^{ij}) is non-negative. The summation convention that repeated indices indicate summation from 1 to d is followed here as it will throughout the rest of the Section.

We will assume the following:

A1) a^{ij} , b^j are bounded and continuous in u , $|\sigma_u^j|_Y$ are bounded, and there is $\delta > 0$ such that for all $\xi \in \mathbf{R}^d$

$$[a^{ij}(t, x, u) - (1/2)(\sigma_u^j(t, x, u), \sigma_u^i(t, x, u))] \xi_i \xi_j \geq \delta |\xi|^2;$$

A2) f^j , f are continuous in u , and σ^j , g are continuous in u in Y -norm, also there exist $C > B > 0$, $m \geq 1$, $p < m$ such that

$$f(t, x, u) \leq C|u|^2 - B|u|^{2m} + L(t, x),$$

$$|f(t, x, u)| \leq C|u|^{2m-1} + H(t, x),$$

$$\begin{aligned} |\sigma_{x_j}^j(t, x, u)|_Y + |\sigma^j(t, x, u)|_Y + |f^j(t, x, u)| + |g(t, x, u)|_Y \\ \leq C(|u|^{p+1} + |u| + K(t, x)), \end{aligned}$$

and for each t

$$\int_0^t \int_G [H(t, x)^{2m/(2m-1)} + L(t, x) + K(t, x)^2] dt dx < \infty.$$

For the consideration of (3.62) we choose the following spaces:

$U = H_0^s(G)$, $s > d/2 + 1$, is the completion of $\tilde{U} = C_0^\infty(G)$ with respect to the Sobolev norm

$$|h|_s = \sum_{|k|=0}^s \left(\int_G (\partial^k h)^2 dx \right)^{1/2};$$

V is a Banach space $H_0^1(G) \cap L^{2m}(G)$ with the norm

$$|v|_V = |v|_{1,V} + |v|_{2,V},$$

where $|v|_{1,V} = |v|_{H_0^1}$, $|v|_{2,V} = |v|_{L^{2m}}$;

$$H = L^2(G);$$

V_c is the Banach space $L^2(G) \cap L^{2m-1/2}(G)$ with the norm

$$|v|_c = |v|_{L^2} + |v|_{L^{2m-1/2}}.$$

We define the operators for $u, \tilde{u} \in \tilde{U} = C_0^\infty(G), v \in V$

$$\begin{aligned} \langle a(s, v), u \rangle &= \int_G \{-a^{ij}(t, x, v) \partial_i v \partial_j u - \\ &\quad - f^j(t, x, v) \partial_j u + b^j(t, x, v) \partial_j v u + f(t, x, v) u\} dx \end{aligned}$$

$$\begin{aligned} \langle Q(s, v) u, \tilde{u} \rangle &= \\ &= \int_E \int_G (\partial_i [\sigma^i(s, x, v)] + g^i(s, x, v)) u(x) dx \\ &\quad \int_G (\partial_i [\sigma^i(s, x, v)] + g^i(s, x, v)) \tilde{u}(x) dx d\kappa, \end{aligned}$$

$$\begin{aligned} \langle Q(s, v) u, u \rangle &= \int_E \{\int_G (\partial_i [\sigma^i(s, x, v)] + g^i(s, x, v)) u(x) dx\}^2 d\kappa = \\ &= \int_E \{\int_G (-\sigma^i(s, x, v) \partial_i u(x) + g^i(s, x, v) u(x)) dx\}^2 d\kappa. \end{aligned}$$

We will need some estimates for these operators.

LEMMA 3.4. *If A1 and A2 hold, then for each $v \in V$,*

$$(3.63) \quad \begin{aligned} |a(t, v)|_V &\leq C(|\partial v|_{L^2} + |v|_{L^{2m}}^{2m-1} + |H(t, \cdot)|_{L^{2m/(2m-1)}} + \\ &\quad + |v|_{L^{2m}}^m + |K(t, \cdot)|_{L^2} + |v|_{L^2}), \end{aligned}$$

and

$$(3.64) \quad \begin{aligned} \text{tr}_U Q(s, v) &\leq C \int_G (|v|^2 + |v|^{2(p \vee 1)} + K(t, x)^2) dx, \\ \text{tr}_H Q(s, v) &\leq C (|v|_{L^2}^2 + |\partial v|_{L^2}^2 + |v|_{L^{2m}}^{2m} + |K(t, \cdot)|_{L^2}^2). \end{aligned}$$

Proof:

Applying Hölder inequality we have the following estimates:

$$(3.65) \quad \begin{aligned} \left| \int_G f^j(t, x, v) \partial_j u dx \right| &\leq \\ &\leq |\partial_j u|_{L^2} \left(\int_G f^j(t, x, v)^2 dx \right)^{1/2} \leq \\ &\leq C |\partial_j u|_{L^2} \left(\int_G (|v|^{2(p \vee 1)} + |v|^2 + K(t, x)^2) dx \right)^{1/2}; \end{aligned}$$

$$(3.66) \quad \begin{aligned} \left| \int_G f(t, x, v) u dx \right| &\leq \\ &\leq |u|_{L^{2m}} \left(\int_G |f(t, x, v)|^{2m/(2m-1)} dx \right)^{(2m-1)/2m} \leq \\ &\leq C |u|_{L^{2m}} \left(|v|_{L^{2m}}^{2m-1} + |H(t, \cdot)|_{L^{2m/(2m-1)}} \right) \end{aligned}$$

$$(3.67) \quad \begin{aligned} \left| \int_G b^j(t, x, v) \partial_j v u dx \right| &\leq \\ &\leq C |\partial_j v|_{L^2} |u|_{L^2}, \end{aligned}$$

$$(3.68) \quad \begin{aligned} & \left| \int_G a^{ij} (t, x, v) \partial_i v \partial_j u dx \right| \leq \\ & \leq C |\partial_i v|_{L^2} |\partial_j u|_{L^2}. \end{aligned}$$

Thus (3.63) holds. Notice that

$$\begin{aligned} \text{tr}_H Q(s, v) &= \int_E \int_G (\sigma_{x_i}^i (s, x, v) + \sigma_v (s, x, v) \partial_i v + g^i (s, x, v))^2 dx d\kappa, \\ \text{tr}_U Q(s, v) &= \int_E \left| \partial_i [\sigma^i (s, x, v)] + g^i (s, x, v) \right|_{U'}^2 d\kappa. \end{aligned}$$

Since

$$\begin{aligned} & \left| \partial_i \sigma^i (s, x, v) + g^i (s, x, v) \right|_{U'} \\ &= \sup_{|u|_U \leq 1} \int_G (\sigma^i (s, x, v) \partial_i u(x) + g^i (s, x, v) u(x)) dx \\ &\leq C \left(\int_G (u^2 + \partial_i u^2) dx \right)^{1/2} \left(\int_G (\sigma^i (s, x, v) + g^i (s, x, v))^2 dx \right)^{1/2}, \end{aligned}$$

we have the estimate

$$(3.69) \quad \begin{aligned} \text{tr}_U Q(s, v) &= \int_E \left| \partial_i \sigma^i (s, x, v) + g^i (s, x, v) \right|_{U'}^2 d\kappa \leq \\ &\leq C \int_G (|\sigma^i (s, x, v)|_Y^2 + |g^i (s, x, v)|_Y^2) dx \leq \\ &\leq C \int_G (|v|^2 + |v|^{2(p \vee 1)} + K(t, x)^2) dx. \end{aligned}$$

Also,

$$(3.70) \quad \begin{aligned} \text{tr}_H Q(s, v) &= \int_E \int_G (\sigma_{x_i}^i (s, x, v) + \sigma_v (s, x, v) \partial_i v + g^i (s, x, v))^2 dx d\kappa \leq \\ &\leq C \int_G \left(|\sigma_{x_i}^i (s, x, v)|_Y^2 + |\partial v|^2 + |g (s, x, v)|_Y^2 \right) dx \leq \\ &\leq C \left(|v|_{L^2}^2 + |\partial v|_{L^2}^2 + |v|_{L^{2m}}^{2m} + |K(t, \cdot)|_{L^2}^2 \right). \end{aligned}$$

■

PROPOSITION 3.6. *Let the assumptions A1, A2 hold. Then for each $u_0 \in H$, there exists a solution \mathbf{P} of the martingale problem (u_0, a, Q) supported by $\mathcal{Z} = D_{[0, \infty)}(U') \cap D_{[0, \infty)}(H_w) \cap L_{loc, w}^{2, 2m}(V) \cap L_{loc}^2(V_c)$.*

Proof:

By Lemma 3.4 the assumptions (ii), (iii) of the Theorem hold. Consider

$$\begin{aligned} \langle Q(s, v) u, u \rangle &= \int_E \{ \int_G (\partial_j [\sigma^j (s, x, v)] + g (s, x, v)) u(x) dx \}^2 d\kappa = \\ &= \int_E \{ \int_G (-\sigma^j (s, x, v) \partial_j u(x) + g (s, x, v) u(x)) dx \}^2 d\kappa = \\ &= \left| \int_G (-\sigma^j (s, x, v) \partial_j u(x) + g (s, x, v) u(x)) dx \right|_Y^2. \end{aligned}$$

If $v_n \rightarrow v$ in V_c , then $v_n \rightarrow v$ in Lebesgue measure on G and by generalized Minkowsky inequality

$$(3.71) \quad \begin{aligned} & \left| \langle Q(s, v_n) u, u \rangle^{1/2} - \langle Q(s, v) u, u \rangle^{1/2} \right| \leq \\ & \leq \int_G (|\sigma^j(s, x, v_n) - \sigma^j(s, x, v)|_Y |\partial_j u(x)| + |g(s, x, v_n) - g(s, x, v)|_Y |u(x)|) dx. \end{aligned}$$

Since $|\sigma^j(s, x, v_n) - \sigma^j(s, x, v)|_Y$ and $|g(s, x, v_n) - g(s, x, v)|_Y$ are uniformly integrable by A2 the continuity of $\langle Q(s, v) u, u \rangle$ follows. Define

$$\langle f^u(t, v), h \rangle = \int_G [-a^{ij}(t, x, v) \partial_i h \partial_j u + b^j(t, x, v) u \partial_j h] dx.$$

If $v_n \rightarrow v$ in V_c , then because of boundedness and continuity of a^{ij}, b^j we have by Cauchy inequality

$$|f^u(t, v_n) - f^u(t, v)|_{V'_1} = \sup_{|h|_{V'_1} \leq 1} |\langle f^u(t, v_n) - f^u(t, v), h \rangle| \rightarrow 0$$

as $n \rightarrow \infty$. Also, $\sup_v |f^u(t, v)|_{V'_1} < \infty$. Define

$$g^u(t, v) = \int_G (f(t, x, v) u - f^j(t, x, v) \partial_j u) dx.$$

If $v_n \rightarrow v$ in V_c , then $v_n \rightarrow v$ in Lebesgue measure on G and by A2 the sequences $f(t, x, v_n), f^j(t, x, v_n)$ are uniformly integrable. So, $g^u(t, v_n) \rightarrow g^u(t, v)$. Thus assumption (iv) of Theorem 3.7 is satisfied with $f^u(t, v)$ and $g^u(t, v)$ defined and considered above ($h^u = 0$). ■

3.4.2. Navier-Stokes equation. Consider the equation for $u = (u^l)_{1 \leq l \leq d}, p, \tilde{p}$

$$(3.72) \quad \left\{ \begin{array}{l} \partial_t u^l(t, x) = \partial_i (a^{l,ij}(t, x, u(t, x)) \partial_j u^l(t, x)) - u^j(t, x) \partial_j u^l(t, x) + \\ + \partial_l p(t, x) + \partial_j f^{l,j}(t, x, u(t, x)) + b^{l,j}(t, x, u(t, x)) \partial_j u^l + f^l(t, x, u(t, x)) + \\ (\partial_j [\sigma^{l,j}(t, x, u^l(t, x))] + g^l(t, x, u(t, x)) + \partial_l \tilde{p}(t, x)) \partial_t W, \quad l = 1, \dots, d, \\ u(t, x)|_{\partial G} = 0, \\ \operatorname{div} u(t, x) = 0 \text{ in } G, \\ u(0, x) = u_0(x) \in L^2(G). \end{array} \right.$$

We assume that $a^{l,ij}, b^{l,j}, f^{l,j}, f^l$ are measurable functions on $[0, \infty) \times G \times \mathbf{R}^d$, $\sigma^{l,j}, g^l$ are Y -valued measurable functions on $[0, \infty) \times G \times \mathbf{R}$ (respectively $[0, \infty) \times G \times \mathbf{R}^d$), and the matrices $(a^{l,ij})$ are non-negative.

In addition we assume the following:

B1) $a^{l,ij}, b^{l,j}$, are bounded and continuous in u , $|\sigma_u^{l,j}|_Y$ are bounded, and there is $\delta > 0$ such that for all $\xi \in \mathbf{R}^d$,

$$[a^{l,ij}(t, x, u) - (1/2)(\sigma_u^{l,j}(t, x, u^l), \sigma_u^{l,i}(t, x, u^l))_Y] \xi_i \xi_j \geq \delta |\xi|^2;$$

B2) $f^{l,j}$, f^l are continuous in u , $\sigma^{l,j}$, g^l are continuous in u in Y -norm, and there exist $C > B > 0$, $m \geq 1$, $p < m$ such that

$$\begin{aligned} f^l(t, x, u) u^l &\leq C|u|^2 - B|u|^{2m} + L(t, x), \\ |f^l(t, x, u)| &\leq C|u|^{2m-1} + H(t, x), \\ |\sigma_{x_j}^{l,j}(t, x, u)|_Y + |\sigma^{l,j}(t, x, u)|_Y + |f^{l,j}(t, x, u)| + |g^l(t, x, u)|_Y &\leq \\ &\leq C(|u|^{p+1} + |u| + K(t, x)), \end{aligned}$$

where

$$\int_0^t \int_G [H(t, x)^{2m/(2m-1)} + L(t, x) + K(t, x)^2] dt dx < \infty$$

for each t .

In studying (3.72) we specify the spaces introduced in the beginning of the Section as follows:

$$\tilde{U} = \left\{ h = (h^i)_{1 \leq i \leq d} : h^i \in C_0^\infty(G), \operatorname{div} h = 0 \right\};$$

$U = \tilde{H}^s(G)$, $s > d/2 + 1$, is a completion of \tilde{U} with respect to the Sobolev norm

$$|h|_s = \sum_{|k|=0}^s \left(\int_G \sum_{l=1}^d (\partial^k h^l)^2 dx \right)^{1/2};$$

V is a Banach space $\tilde{H}^1(G) \cap L^{2m}(G)$ with the norm

$$|v|_V = |v|_{1,V} + |v|_{2,V},$$

where $|v|_{1,V} = |v|_1$, $|v|_{2,V} = |v|_{L^{2m}}$;

H be a closure of \tilde{U} in $L^2(G)^d$;

V_c is the Banach space $H \cap L^{2m-1/2}(G)$ with the norm

$$|v|_c = |v|_{L^2} + |v|_{L^{2m-1/2}}.$$

Now we shall define the operators. For each $u \in U$,

$$\begin{aligned} \langle a(s, v), u \rangle &= \int_G \{-a^{l,ij}(t, x, v) \partial_i v^l \partial_j u^l - v^k(t, x) \partial_k v^l u^l + \\ &- f^{l,j}(t, x, v) \partial_j u^l + b^{l,j}(t, x, v) \partial_j v^l u^l + f^l(t, x, v) u^l\} dx, \end{aligned}$$

$$\langle Q(s, v) u, u \rangle = \int_E \left\{ \int_G (\partial_j [\sigma^{l,j}(t, x, v^l)] + g^l(t, x, u)) u^l dx \right\}^2 d\kappa.$$

LEMMA 3.5. If B1 and B2 hold, then for each $v \in V$,

$$\begin{aligned} (3.73) \quad |a(t, v)|_{U'} &\leq C(|\partial v|_{L^2} + |v|_{L^{2m}}^{2m-1} + |H(t, \cdot)|_{L^{2m/(2m-1)}} + \\ &+ |v|_{L^{2m}}^m + |K(t, \cdot)|_{L^2} + |v|_{L^2} + |v|_{L^2} |\partial v|_{L^2}), \end{aligned}$$

and

$$(3.74) \quad \begin{aligned} \text{tr}_U Q(s, v) &\leq C \int_G \left(|v|^2 + |v|^{2(p\vee 1)} + K(t, x)^2 \right) dx, \\ \text{tr}_H Q(s, v) &\leq C \left(|v|_{L^2}^2 + |\partial v|_{L^2}^2 + |v|_{L^{2m}}^{2m} + |K(t, \cdot)|_{L^2}^2 \right). \end{aligned}$$

Proof:

Similarly as in the proof of Lemma 3.4 applying Hölder inequality we have the following estimates:

$$\begin{aligned} & \left| \int_G f^{l,j}(t, x, v) \partial_j u^l dx \right| \leq \\ & \leq C |\partial u|_{L^2} \left(\int_G \left(|v|^{2(p\vee 1)} + |v|^2 + K(t, x)^2 \right) dx \right)^{1/2}; \\ & \left| \int_G f^l(t, x, v) u^l dx \right| \leq \\ & \leq C |u|_{L^{2m}} \left(|v|_{L^{2m}}^{2m-1} + |H(t, \cdot)|_{L^{2m/(2m-1)}} \right); \\ & \left| \int_G b^{l,j}(t, x, v) \partial_j v^l u^l dx \right| \leq \\ & \leq C |\partial v|_{L^2} |u|_{L^2}, \\ & \left| \int_G a^{l,ij}(t, x, v) \partial_i v^l \partial_j u^l dx \right| \leq \\ & \leq C |\partial v|_{L^2} |\partial u|_{L^2}. \end{aligned}$$

Also, by Sobolev embedding theorem ($s > d/2 + 1$)

$$\begin{aligned} & \left| \int_G v^k(t, x) \partial_k v^l u^l dx \right| \leq \\ & \leq C |u|_U |v|_{L^2} |\partial v|_{L^2}. \end{aligned}$$

Thus (3.73) holds. Notice that

$$\text{tr}_H Q(s, v) = \int_E \int_G \sum_l \left(\sigma_{x_j}^{l,j}(s, x, v) + \sigma_v^{l,j}(s, x, v) \partial_j v + g^l(s, x, v) \right)^2 dx d\kappa,$$

$$\text{tr}_U Q(s, v) = \int_E \left| (\partial_j [\sigma^{l,j}(s, x, v)] + g^l(s, x, v))_{1 \leq l \leq d} \right|_U^2 d\kappa.$$

Since

$$\begin{aligned} & \left| (\partial_j [\sigma^{l,j}(s, x, v)] + g^l(s, x, v))_l \right|_U = \\ & = \sup_{|u|_U \leq 1} \int_G (\sigma^{l,j}(s, x, v) \partial_j u^l(x) + g^l(s, x, v) u^l(x)) dx \leq \\ & \leq C \left(\int_G (|u|^2 + |\partial u|^2) dx \right)^{1/2} \left(\int_G \sum_l (|\sigma^l(s, x, v)| + g^l(s, x, v))^2 dx \right)^{1/2}, \end{aligned}$$

we have the estimate

$$\begin{aligned} \text{tr}_U Q(s, v) &\leq C \left(\int_G \sum_l (|\sigma^l(s, x, v)|_Y + |g^l(s, x, v)|_Y)^2 dx \right) \leq \\ &\leq C \int_G (|v|^2 + |v|^{2(p\vee 1)} + K(t, x)^2) dx. \end{aligned}$$

Also,

$$\begin{aligned} \text{tr}_H Q(s, v) &= \int_E \int_G \sum_l \left(\sigma_{x_i}^{l,j}(s, x, v) + \sigma_v^{l,j}(s, x, v^l) \partial_j v^l + g^l(s, x, v) \right)^2 dx d\kappa \leq \\ &\leq C \int_G \sum_l \left(|\sigma_{x_i}^l(s, x, v)|_Y^2 + |\partial v^l|^2 + |g^l(s, x, v)|_Y^2 \right) dx \leq \\ &\leq C \left(|v|_{L^2}^2 + |\partial v|_{L^2}^2 + |v|_{L^{2m}}^{2m} + |K(t, \cdot)|_{L^2}^2 \right). \end{aligned}$$

Thus (3.74) holds. ■

Now we prove the existence of solutions of the martingale problem.

PROPOSITION 3.7. *Let the assumptions B1, B2 hold. Then for each $u_0 \in H$, there exists a solution to the martingale problem (u_0, a, Q) and it is supported by $\mathcal{Z} = D_{[0, \infty)}(U') \cap D_{[0, \infty)}(H_w) \cap L_{loc, w}^{2, 2m}(V) \cap L_{loc}^2(V_c)$.*

Proof:

For all $u \in U$ the equality $\langle a(s, u), u \rangle = \langle \tilde{a}(s, u), u \rangle$ holds, where

$$\begin{aligned} \langle \tilde{a}(s, v), h \rangle &= \int_G \{-a^{l,ij}(t, x, v) \partial_i v^l \partial_j h^l - \\ &- f^{l,j}(t, x, v) \partial_j h^l + b^{l,j}(t, x, v) \partial_j v^l h^l + f^l(t, x, v) h^l\} dx. \end{aligned}$$

By Lemma 3.5 it follows that the assumptions (ii), (iii) are satisfied. Repeating with obvious changes the proof of Proposition 3.6 we show that $\langle Q(t, v) u, u \rangle$ is a continuous function with respect to V_c -norm. Define

$$\langle f^u(t, v), h \rangle = \int_G [-a^{l,ij}(t, x, v) \partial_i h^l \partial_j u^l + b^{l,j}(t, x, v) u^l \partial_j h^l] dx.$$

Again, as in the case of Proposition 3.6 (all the arguments are the same) we can prove the continuity of $f^u(t, v)$. Obviously, $\sup_v |f^u(t, v)|_{V'_1} < \infty$. Define

$$g^u(t, v) = \int_G (f^l(t, x, v) u^l - f^{l,j}(t, x, v) \partial_j u^l + v^k \partial_k u^l v^l) dx.$$

Obviously, $g^u(t, v)$ is continuous with respect to V_c -norm. Thus the assumption (iv) of Theorem 3.7 is satisfied with the functions $f^u(t, v)$, $g^u(t, v)$ defined above and $h^u = 0$, and the statement follows. ■

4. Equations of Stochastic Quantization

Let $U \subset \mathbf{R}^2$ be a bounded smooth domain and Δ be the Dirichlet Laplacian (see Appendix). The free Euclidean quantum field is given by the Gaussian measure μ on the space of distributions $\mathcal{D}'(U)$ with zero mean and the covariance operator $(I - \Delta)^{-\alpha_0}$, $\alpha_0 \geq 1$ (see [9], [7]). It is rather straightforward to show that μ is a unique invariant measure of the Markov process in $\mathcal{D}'(U)$ given by the equation

$$(4.1) \quad dX_t = -1/2(I - \Delta)^\varepsilon X_t dt + (I - \Delta)^\gamma dW_t,$$

where $\varepsilon > 0$, $\alpha_0 = \varepsilon - 2\gamma \geq 1$ and W is a standard cylindrical Wiener process in $L_2(U)$ (see [17]).

The polynomial interaction quantum field is given by the measure

$$\mu' = c_1 \exp \left\{ -\lambda \int_U :X(x)^{2m}: dx \right\} \mu(dX),$$

where c_1 is the normalizing constant and $:X(x)^{2m}:$ denotes the Wick product in $(\mathcal{D}'(U), \mu)$ of order $2m$. Parisi and Wu (see [38]) proposed a program to find a Markov process or a family of Markov processes so that μ' is their invariant measure. It was shown in [17] that this Markov process can be constructed as a stationary solution of the so called stochastic quantization equation with the initial distribution μ'

$$(4.2) \quad dX_t = -1/2(I - \Delta)^\varepsilon X_t dt - \lambda m(I - \Delta)^{2\gamma} :X^{2m-1}: dt + \\ +(I - \Delta)^\gamma dW_t ,$$

where $m = 2$. In the latter work the stationary solution was defined as a stationary measure on the space of trajectories Ω such that (4.2) holds for the canonical process X_t and $X_0(\mathbf{P}) = \mu'$. Under the assumption $\varepsilon \in (0, 1/10)$, $\alpha_0 = 1$, the Girsanov transform was used and the solution constructed by changing the measure of the corresponding stationary solution of (4.1). Since then there was an extensive study of the problem (see [18], [20], [31]-[42]). In the case of $\alpha_0 = 1$, $\varepsilon \in (0, 1)$, the approach based on the use of Dirichlet forms was taken in [18]. This argument was later elaborated in [20], and the case of the infinite volume was considered (see [31]-[37]). The rigorous proof of the result for $\varepsilon \in (0, 1/10)$ based on Girsanov method and Kazamaki criterion of the uniform integrability was presented in [21]. In [19] the change of measure method was extended to the case of $\varepsilon \in (0, 1)$. Unfortunately, we have been unable to reconstruct some parts of the proof of the uniform integrability of the sequence approximating the Green function of the Dirichlet Laplacian in [19].

In this Section we develop a new approach to the stochastic quantization equation that is based on the compactness method. We will demonstrate that this method allows to extend previously known existence of solutions results to the most interesting "white noise" case, $\gamma = 0$, $\varepsilon = 1$. We also prove that in this case the stationary solutions to equations (4.2) and (4.1) are singular. The latter implies of course that in the "white noise" case the Girsanov method is inapplicable.

The main result of this Section is as follows (see Theorem 4.3 and Propositions 4.2, 4.1, 4.3, 4.5 below).

THEOREM 4.1. *Let $0 \leq \alpha_0 - 1 \leq 2$, $\varepsilon > 0$. Then:*

- a) *There exists a stationary solution \mathbf{P} of (4.2) such that $X_0(\mathbf{P}) = \mu'$;*
- b) *If $2\gamma < \alpha_0 - 1$, the stationary solution is unique ergodic and strong Markov with respect to the natural filtration of σ -algebras (\mathcal{F}_t) , where $\mathcal{F}_t = \sigma(X_s, s \leq t)$. Moreover, $\mathbf{P}|_{\mathcal{F}_t} \sim \mathbf{P}_0|_{\mathcal{F}_t}$ for each $t > 0$, where \mathbf{P}_0 is the stationary solution of (4.1);*
- c) *If $\gamma = 0$, $\varepsilon = 1$, $\alpha_0 = 1$, the stationary solutions \mathbf{P} of (4.2) and \mathbf{P}_0 of (4.1) are singular on \mathcal{F}_t for each $t > 0$.*

Using absolute continuity arguments and stationary solution of (4.2) we prove the existence and uniqueness and Markov property of some non-stationary solutions of (4.2) (see Propositions 4.2, 4.1).

The outline of this section is as follows. In Subsection 1 we discuss the fundamental properties of Wick products generated by Gaussian random fields. Many technical results used in this subsection, including Hilbert space generated by the Dirichlet Laplacian and some properties of Green function for the $P(\varphi)_2$ quantum

field theory can be found in Appendix. In the second subsection we construct the stationary solutions and derive all the main results.

4.1. Wick Products of Gaussian Random Fields. Let $U \subset \mathbf{R}^2$ be a bounded smooth domain. For $s \geq 0$ we denote $H_0^s = H_0^s(U)$ the completion of $C_0^\infty(U)$ with respect to the $H^s(U)$ -norm. By definition $H^s(U)$ is the space of generalized functions $f \in \mathcal{D}'(U)$ which are restrictions of $g \in H^s(\mathbf{R}^d)$ (we write $f = Rg$ and call R a restriction operator). The corresponding norm is defined by

$$|f|_{H^s} = |f|_{H^s(U)} = \inf_{Rg=f} |g|_{H^s(\mathbf{R}^2)},$$

where $|g|_{H^s(\mathbf{R}^2)} = |\mathcal{F}^{-1}(1 + |\xi|^s)\mathcal{F}g|_{L_2(\mathbf{R}^2)}$, \mathcal{F} is a Fourier transform and \mathcal{F}^{-1} its inverse. Let (e_k) be the basis in $L_2(U)$ of the eigenfunctions of the Dirichlet Laplacian (see Appendix), i.e. $(-\Delta)e_k = \lambda_k e_k$, $e_k|_{\partial U} = 0$, $k = 1, 2, \dots$. It is a standard fact that $e_k \in H_0^1(U) \cap C^\infty(\bar{U})$ and $\lambda_k \sim k$ (see [11]). Introduce the spaces $\mathcal{H}^s(U) = \{f \in L_2(U) : f = \sum_k f_k e_k\text{, and }|f|_s = (\sum_k (1 + \lambda_k)^s f_k^2)^{1/2} < \infty\}$, $s \geq 0$. According to [27] $\mathcal{E} = \cap_{s \geq 0} \mathcal{H}^s$ (with the sets $V_{s,\varepsilon} = \{f : |f|_s \leq \varepsilon\}$, $s \geq 0$, $\varepsilon > 0$, as zero neighbourhood basis) is a countably Hilbert nuclear space, its dual $\mathcal{E}' = \cup_{s \geq 0} (\mathcal{H}^s)'$. We have here a Gelfand triple $\mathcal{E} \subset L^2(U) \subset \mathcal{E}'$.

Consider the following equation in U

$$(I - \Delta)^\alpha u = f.$$

If $f = \sum_k f_k e_k \in \mathcal{H}^s(U)$, then $u = (I - \Delta)^{-\alpha} f = \sum_k (1 + \lambda_k)^{-\alpha} f_k e_k \in \mathcal{H}^{s+2\alpha}(U)$ is a solution of it. According to Remark 5.3 in Appendix, it can be represented also using Green's function G_α by the formula

$$(I - \Delta)^{-\alpha} f = \int_U G_\alpha(x, y) f(y) dy$$

(see Appendix for the properties of the Green's functions).

Let μ be the Gaussian measure on \mathcal{E}' with the covariance operator $(I - \Delta)^{-\alpha_0}$ ($\alpha_0 \geq 1$) and zero mean, i.e. $\mu = N(0, (I - \Delta)^{-\alpha_0})$. Denote X the "canonical element" of \mathcal{E}' . Thus for $\varphi, \psi \in \mathcal{E}$

$$\mathbf{E} X(\varphi) X(\psi) = ((I - \Delta)^{-\alpha_0} \varphi, \psi)_0 = (\varphi, \psi)_{-\alpha_0},$$

where $(\varphi, \psi)_0 = \int_U \varphi \psi dx$.

REMARK 4.1. We notice here that for each $\alpha < \alpha_0 - 1$ X is μ -a.s. $\mathcal{H}^\alpha(U)$ -valued random variable.

Indeed, for each $\varphi, \psi \in \mathcal{E}$

$$\begin{aligned} \mathbf{E} X((I - \Delta)^\alpha \varphi) X(I - \Delta)^\alpha \psi &= ((I - \Delta)^{-\alpha_0} I - \Delta)^\alpha \varphi, (I - \Delta)^\alpha \psi)_0 = \\ &= ((I - \Delta)^{-\alpha_0 + \alpha} \varphi, \psi)_\alpha. \end{aligned}$$

By Theorem 3.3.4 in [5] for each natural number $N = 0, 1, 2, \dots$ there exists an extension operator $S = S^N$ mapping continuously $H^s(U)$ to $H^s(\mathbf{R}^2)$ for all $|s| < N$ and such that $RS = I$, where R is a restriction operator (see formula 3.3.4/6 in [5]). Moreover, there exists $c > 0$ such that $\text{supp } Sf \subset \{x : |x| < c\}$ for each $f \in H^s(U)$, $|s| < N$. Fix $N > \max\{2(\alpha_0 - 1), 2\}$ and consider the extension

$\bar{X} = SX = S^N X$. Since $X \in \mathcal{H}^\alpha(U)$ μ -a.s., if $\alpha < \alpha_0 - 1$, and $\mathcal{H}^\alpha(U) \subset H^{-1}(U)$, if $\alpha > -1$, it is correctly defined. Then

$$\mathbf{E}\bar{X}(\varphi)\bar{X}(\psi) = \int \int \bar{G}_{\alpha_0}(x, y)\varphi(x)\psi(y) dx dy,$$

where $\varphi, \psi \in \mathcal{D}(\mathbf{R}^2)$ and

$$\bar{G}_\alpha(x, y) = \bar{G}_{\alpha, N}(x, y) = S_x^N S_y^N G_\alpha(x, y)$$

Let $\bar{X}^\varepsilon(x) = \bar{X} * w_\varepsilon(x) = \langle \bar{X}, w_\varepsilon(x - \cdot) \rangle_{\mathcal{D}', \mathcal{D}}$, where $w_\varepsilon(x) = \varepsilon^{-2}w(x/\varepsilon)$, $w \in C_0^\infty(\mathbf{R}^2)$, $w \geq 0$, $\int w dx = 1$. Since \bar{X} has a compact support, $\bar{X}^\varepsilon \in C_0^\infty(\mathbf{R}^2)$ and

$$\mathbf{E}\bar{X}^\varepsilon(x)\bar{X}^{\varepsilon'}(y) = \bar{G}_{\alpha_0}^{\varepsilon, \varepsilon'}(x, y)$$

$$\mathbf{E}\bar{X}^\varepsilon(x)\bar{X}^{\varepsilon'}(y) = \bar{G}_{\alpha_0}^{\varepsilon, \varepsilon'}(x, y)$$

for each $x, y \in \mathbf{R}^2$, where

$$\bar{G}_\alpha^{\varepsilon, \varepsilon'}(x, y) = \bar{G}_{\alpha, N}^{\varepsilon, \varepsilon'}(x, y) = \int \int \bar{G}_{\alpha, N}(z, z') w_\varepsilon(x - z) w_{\varepsilon'}(y - z') dz dz' .$$

Now we will define Wick products of X as limits of \bar{X}^ε . For $\alpha > 0, p \geq 2$ we introduce the space $C^{\alpha, p}(U, \mu) = C^{\alpha, p}$ of random functions f on U with finite Hölder norm

$$|f|_{\alpha, p} = |f|_{\alpha, p; U} = \sup_U (\mathbf{E}^\mu |f(x)|^p)^{1/p} + \sup_U (\mathbf{E}^\mu |\partial^k f(x)|^p)^{1/p} +$$

$$+ \sup_{x \neq y} (\mathbf{E}^\mu |\partial^k f(x) - \partial^k f(y)|^p)^{1/p} |x - y|^{-\kappa} ,$$

where $\alpha = k + \kappa$, $k \in \mathbf{N}$, $\kappa \in [0, 1)$.

LEMMA 4.1. a) For each n there exists a unique \mathcal{E}' -valued random variable $:X^n : : - : X^n : |_\alpha^2 \rightarrow 0$, as $\varepsilon \rightarrow 0$, if $\alpha < \alpha_0 - 1$

$$\mathbf{E}^\mu |:(\bar{X}^\varepsilon)^n : - : X^n :|_\alpha^2 \rightarrow 0, \text{ as } \varepsilon \rightarrow 0, \text{ if } \alpha < 0,$$

$$|:(\bar{X}^\varepsilon)^n : - : X^n :|_{\alpha, 2} \rightarrow 0, \text{ as } \varepsilon \rightarrow 0, \text{ if } \alpha \geq 0,$$

$$\mathbf{E}^\mu |:X^n :|_\alpha^2 = \int_U \int_U G_{-\alpha}(x, y) G_{\alpha_0}(x, y)^n dx dy, \text{ if } \alpha < 0,$$

$$\mathbf{E}^\mu [\langle :X^n :, \varphi \rangle_{\mathcal{E}', \mathcal{E}} \langle :X^n :, \psi \rangle_{\mathcal{E}', \mathcal{E}}] = \int_U \int_U G_{\alpha_0}(x, y)^n \varphi(x) \psi(y) dx dy ; \varphi, \psi \in \mathcal{E};$$

b) for each $\alpha < \alpha_0 - 1$ and $p \geq 2$ there exists a constant C such that

$$(\mathbf{E}^\mu |:(\bar{X}^\varepsilon)^n : - : X^n :|_\alpha^p)^{2/p} \leq C \mathbf{E}^\mu |:(\bar{X}^\varepsilon)^n : - : X^n :|_\alpha^2, \alpha < 0,$$

$$|:(\bar{X}^\varepsilon)^n : - : X^n :|_{\alpha, p} \leq C |:(\bar{X}^\varepsilon)^n : - : X^n :|_{\alpha, 2}^2, \alpha \geq 0,$$

$$(\mathbf{E}^\mu |:X^n :|_\alpha^p)^{2/p} \leq C \mathbf{E}^\mu |:X^n :|_\alpha^2, \alpha < 0,$$

$$|:X^n :|_{\alpha, p} \leq C |:X^n :|_{\alpha, 2}, \alpha \geq 0;$$

c) there exist $\eta > 0$, $C > 0$ such that

$$\mathbf{E}^\mu [(\int_U |:(X^\varepsilon)^n :| dx - \int_U |:X^n :| dx)^2] \leq C \varepsilon^\eta .$$

Moreover, for each $p \geq 2$, there is a constant C such that

$$(\mathbf{E}^\mu[(\int_U :X^\varepsilon:n dx - \int_U :X^n:dx)^p])^{2/p} \leq C\mathbf{E}^\mu[(\int_U :X^\varepsilon:n dx - \int_U :X^n:dx)^2],$$

$$(\mathbf{E}^\mu[|\int_U :X^n:dx|^p])^{2/p} \leq C\mathbf{E}^\mu[|\int_U :X^n:dx|^2] = C \int_U \int_U G_{\alpha_0}(x, y)^n dx dy.$$

REMARK 4.2. Here

a)

$$\langle :X^n:, \varphi \rangle_{\varepsilon', \varepsilon} = \begin{cases} \langle :X^n:, \varphi \rangle_{H^{-\alpha}, H^\alpha}, & \text{if } \alpha < \alpha_0 - 1 = 0 \\ (\langle :X^n:, \varphi \rangle_0, & \text{if } \alpha_0 > 1; \end{cases}$$

b) If $\alpha_0 = 1$, $\int_U :X^n:dx$ is the limit of $\langle :X^n:, \varphi_m \rangle_{\varepsilon', \varepsilon}$, where φ_m is any sequence in $\mathcal{D}(U)$ such that $\sup_m |\varphi_m|$ is bounded and $\varphi_m \rightarrow 1$ dx -a.e.

Proof:

a) Let $\varepsilon, \varepsilon' \in (0, 1)$, $\alpha_0 = 1$. Fix $\alpha \in (0, 1)$ and denote $J^{\varepsilon, \varepsilon'}(x) =: \bar{X}^\varepsilon(x)^n : - : \bar{X}^{\varepsilon'}(x)^n :$. We notice here that

$$|J^{\varepsilon, \varepsilon'}|_{-\alpha}^2 = \int_U L^{\varepsilon, \varepsilon'}(x)^2 dx,$$

where $L^{\varepsilon, \varepsilon'}(x) = \int_U G_{\alpha/2}(x, y) J^{\varepsilon, \varepsilon'}(y) dy$.

Then by Minkowsky inequality

$$\begin{aligned} \mathbf{E}^\mu |J^{\varepsilon, \varepsilon'}|_{-\alpha}^p &= \mathbf{E}^\mu [(\int_U L^{\varepsilon, \varepsilon'}(x)^2 dx)^{p/2}] \leq \\ &\leq (\int_U (\mathbf{E}^\mu |L^{\varepsilon, \varepsilon'}(x)|^p)^{2/p} dx)^{p/2}. \end{aligned}$$

By supercontractivity principle (see Theorem 1.22 in [9])

$$\begin{aligned} (\int_U (\mathbf{E}^\mu |L^{\varepsilon, \varepsilon'}(x)|^p)^{2/p} dx)^{p/2} &\leq C(\int_U \mathbf{E}^\mu L^{\varepsilon, \varepsilon'}(x)^2 dx)^{p/2} = \\ &= C(\mathbf{E}^\mu \int_U L^{\varepsilon, \varepsilon'}(x)^2 dx)^{p/2} = C(\mathbf{E}^\mu |J^{\varepsilon, \varepsilon'}|_{-\alpha}^2)^{p/2}. \end{aligned}$$

Thus

$$(4.3) \quad (\mathbf{E}^\mu |J^{\varepsilon, \varepsilon'}|_{-\alpha}^p)^{1/p} \leq C(\mathbf{E}^\mu |J^{\varepsilon, \varepsilon'}|_{-\alpha}^2)^{1/2}.$$

On the other hand,

$$\begin{aligned} \mathbf{E}^\mu |J^{\varepsilon, \varepsilon'}|_{-\alpha}^2 &= \mathbf{E}^\mu \int_U \int_U G_\alpha(x, y) J^{\varepsilon, \varepsilon'}(x) J^{\varepsilon, \varepsilon'}(y) dx dy = \\ &= \int_U \int_U G_\alpha(x, y) (\bar{G}_{\alpha_0}^{\varepsilon, \varepsilon'}(x, y)^n - \bar{G}_{\alpha_0}^{\varepsilon, \varepsilon'}(x, y)^n - \bar{G}_{\alpha_0}^{\varepsilon', \varepsilon}(x, y)^n + \\ &\quad + \bar{G}_{\alpha_0}^{\varepsilon', \varepsilon'}(x, y)^n) dx dy \rightarrow 0, \text{ as } \varepsilon, \varepsilon' \rightarrow 0, \end{aligned}$$

by Lemma 5.2 in Appendix.

Similarly, by Minkowsky inequality and supercontractivity principle (Theorem 1.22 in [9]), if $\alpha_0 > 1$, $0 \leq \alpha < \alpha_0 - 1$,

$$(4.4) \quad |J^{\varepsilon, \varepsilon'}|_{\alpha, p} \leq C |J^{\varepsilon, \varepsilon'}|_{\alpha, 2}.$$

Let $\alpha = k + \kappa$, $k \in \mathbf{N}$, $\kappa \in [0, 1)$. Then

$$\begin{aligned} |J^{\varepsilon, \varepsilon'}|_{\alpha, p; U}^2 &\leq C |J^{\varepsilon, \varepsilon'}|_{\alpha, 2; \mathbf{R}^2}^2 = |J^{\varepsilon, \varepsilon'}|_{0, 2; \mathbf{R}^2}^2 + ||\partial_x^\kappa \partial_y^k J^{\varepsilon, \varepsilon'}|_{0, 2; \mathbf{R}^2}^2 = \\ &= \sup_x |\bar{G}^{\varepsilon, \varepsilon}(x, x)^n - \bar{G}^{\varepsilon, \varepsilon'}(x, x)^n - \bar{G}^{\varepsilon', \varepsilon}(x, x)^n + \bar{G}^{\varepsilon', \varepsilon'}(x, x)^n| + \sup_x \\ &\quad ||\partial_x|^\kappa \partial_y|^\kappa \partial_y^k \partial_x^k ((\bar{G}^{\varepsilon, \varepsilon}(x, y)^n - \bar{G}^{\varepsilon, \varepsilon'}(x, y)^n - \bar{G}^{\varepsilon', \varepsilon}(x, y)^n + \bar{G}^{\varepsilon', \varepsilon'}(x, y)^n)|_{y=x}|. \end{aligned}$$

The last term here can be estimated by

$$\begin{aligned} &\sup_{\eta, x} (||\partial_x|^\kappa \partial_y|^\kappa \partial_y^k \partial_x^k ((\bar{G}^{\varepsilon, \eta}(x, y)^n - \bar{G}^{0, \eta}(x, y)^n)|_{y=x} + \\ &\quad + ||\partial_x|^\kappa \partial_y|^\kappa \partial_y^k \partial_x^k ((\bar{G}^{\varepsilon', \eta}(x, y)^n - \bar{G}^{0, \eta}(x, y)^n)|_{y=x}). \end{aligned}$$

Thus $|J^{\varepsilon, \varepsilon'}|_{\alpha, 2} \rightarrow 0$, as $\varepsilon, \varepsilon' \rightarrow 0$, by Corollary 5.1 in Appendix. In any case for each $\varphi, \psi \in \mathcal{E}$ and $\alpha < \alpha_0 - 1$

$$\mathbf{E}^\mu | : (\bar{X}^\varepsilon)^n : |_\alpha^2 \rightarrow \mathbf{E}^\mu | : X^n : |_\alpha^2, \text{ if } \alpha < 0,$$

$$| : (\bar{X}^\varepsilon)^n : |_{\alpha, 2} \rightarrow | : X^n : |_{\alpha, 2}, \text{ if } \alpha \geq 0,$$

$$\mathbf{E}^\mu \langle : (\bar{X}^\varepsilon)^n : , \varphi \rangle_{\mathcal{E}', \mathcal{E}} \langle : (\bar{X}^\varepsilon)^n : , \psi \rangle_{\mathcal{E}', \mathcal{E}} = \mathbf{E}^\mu [((\bar{X}^\varepsilon)^n : , \psi)_0 ((\bar{X}^\varepsilon)^n : , \varphi)_0] =$$

$$\int_U \int_U G_{\alpha_0}^{\varepsilon, \varepsilon}(x, y)^n \varphi(x) \psi(y) dx dy \rightarrow \int_U \int_U G_{\alpha_0}(x, y)^n \varphi(x) \psi(y) dx dy =$$

$$= \mathbf{E}^\mu \langle : X^n : , \varphi \rangle_{\mathcal{E}', \mathcal{E}} \langle : X^n : , \psi \rangle_{\mathcal{E}', \mathcal{E}}, \text{ as } \varepsilon \rightarrow 0.$$

The part a) of the statement follows.

b) Let $\varepsilon, \varepsilon' \in (0, 1)$. Then by supercontractivity principle again and Lemma 5.4 in Appendix, for each $p \geq 2$ there is C such that

$$\begin{aligned} &(\mathbf{E}^\mu [(\int_U : (\bar{X}^\varepsilon)^n : dx - \int_U : (\bar{X}^{\varepsilon'})^n : dx)^p])^{2/p} \\ &\leq C \mathbf{E}^\mu [(\int_U : (\bar{X}^\varepsilon) : dx - \int_U : (\bar{X}^{\varepsilon'}) : dx)^2] \\ &= C \int_U \int_U (\bar{G}_{\alpha_0}^{\varepsilon, \varepsilon}(x, y)^n - \bar{G}_{\alpha_0}^{\varepsilon, \varepsilon'}(x, y)^n - \bar{G}_{\alpha_0}^{\varepsilon', \varepsilon}(x, y)^n + \bar{G}_{\alpha_0}^{\varepsilon', \varepsilon'}(x, y)^n) dx dy \\ &\leq C(\varepsilon \vee \varepsilon')^\kappa. \end{aligned}$$

Let φ_m be a sequence in $\mathcal{D}(U)$ such that $\varphi_m \rightarrow 1$ dx -a.e. and $\sup_m |\varphi_m|$ is bounded. It is easy to see that

$$\sup_{\varepsilon \in (0, 1)} \mathbf{E}^\mu [(\int_U : (\bar{X}^\varepsilon)^n : dx - \int_U : (\bar{X}^\varepsilon)^n : \varphi_m dx)^2] \rightarrow 0,$$

as $m \rightarrow \infty$. Direct computation shows that

$$\mathbf{E}^\mu [(\int_U : (\bar{X}^\varepsilon)^n : dx)^2] = \int_U \int_U (\bar{G}_{\alpha_0}^{\varepsilon, \varepsilon}(x, y)^n dx dy.$$

The statement is proved. ■

COROLLARY 4.1. *If $0 \leq \alpha < (\alpha_0 - 1)$, $\alpha \leq 2$, $p \geq 2$,*

$$\mathbf{E}^\mu (| : X^n : |_\alpha^p) \leq C \mathbf{E}^\mu (| : X^n : |_\alpha^2)^{p/2},$$

i.e. $: X^n :$ belongs to \mathcal{H}^α in this case.

Proof:

It follows from Lemma 4.1 that : $X^n \in H^\alpha(U)$ and : X^n : is zero on the boundary of U . Let $\rho(x) = \text{dist}(x, \partial U)$. By part b) of Lemma 4.1 $\mathbf{E}| : X(x)^n : \rho(x)^{-1/2}|$ is bounded if $\alpha_0 - 1 > 1/2$. So the statement follows by Lemma 5.1 in Appendix. ■

Now we discuss the relation between Wick products and multiple integrals.

4.1.1. *Wick products and multiple integrals.* Let $B = (I - \Delta)^{\alpha_0/2}X$. Then $X = (I - \Delta)^{-\alpha_0/2}B$, B is a Gaussian field and $\mathbf{E}B(\varphi) = 0$, $\mathbf{E}B(\varphi)B(\psi) = (\varphi, \psi)_0 = \int_U \varphi\psi dx$. Now we define a multiple integral with respect to B . Let (ζ_k) be an arbitrary CONS in $L_2(U)$. Then $(\zeta_{i_1}(x_1) \dots \zeta_{i_n}(x_n))$, $(i_1, \dots, i_n) \in \mathbb{N}_0^n$, form a CONS in $L_2(U^n)$ (here $\mathbb{N}_0 = \{0, 1, 2, \dots\}$). So, if $f \in L_2(U^n)$,

$$(4.5) \quad f = \sum_{i_1, \dots, i_n} f_{i_1, \dots, i_n} \zeta_{i_1}(x_1) \dots \zeta_{i_n}(x_n).$$

With any $(i_1, \dots, i_n) \in \mathbb{N}_0^n$ we associate an infinite multiindex $\alpha = (\alpha_k)_{k \geq 1}$, where α_k counts how many times k appears in (i_1, \dots, i_n) . Then firstly we define the multiple integral of $\zeta_{i_1} \dots \zeta_{i_n}$ with respect to B as

$$\begin{aligned} I_n(\zeta_{i_1} \dots \zeta_{i_n}) &= \int \zeta_{i_n}(x_1) \dots \zeta_{i_1}(x_n) B(dx_1) \dots B(dx_n) = \Pi_k H_{\alpha_k}(B(\zeta_k)) = \\ &=: \Pi_k B(\zeta_k)^{\alpha_k}, \end{aligned}$$

where $H_j(r) = (-1)^j e^{-r^2/2} \frac{d^j}{dr^j} e^{-r^2/2}$ is the Hermit polynomial of the j th order, $H_0 = 1$. So, if $(\alpha_k) \neq (\alpha'_k)$

$$\mathbf{E} \Pi_k H_{\alpha_k}(B(\zeta_k)) \Pi_k H_{\alpha'_k}(B(\zeta_k)) = 0.$$

Also,

$$\mathbf{E} I_n(\zeta_{i_1} \dots \zeta_{i_n})^2 = \alpha! = \Pi_k \alpha_k!.$$

We extend this definition by continuity and linearity to the whole $L_2(U^n)$. For $f \in L_2(U^n)$ given by (4.5) we define

$$\begin{aligned} I_n(f) &= \int f(x_1, \dots, x_n) B(dx_1) \dots B(dx_n) = \\ &= \sum_{i_1, \dots, i_n} f_{i_1 \dots i_n} I_n(\zeta_{i_1} \dots \zeta_{i_n}). \end{aligned}$$

We have the isometry

$$(4.6) \quad \mathbf{E} I_n(f)^2 = n! |f_{sym}|_0^2 \leq n! |f|_0^2,$$

where f_{sym} is a symmetrization of f .

REMARK 4.3. We notice here that

$$I_n(f) = I_n(f_{sym}),$$

and the definition does not depend on the choice of a particular basis. Also, for all $\varphi_1, \dots, \varphi_n \in \mathcal{E}$ we have

$$: B(\varphi_1) \dots B(\varphi_n) := \int \varphi(x_1) \dots \varphi(x_n) B(dx_1) \dots B(dx_n),$$

$$: X(\varphi_1) \dots X(\varphi_n) := : B(G_{\alpha_0/2}(\varphi_1)) \dots B(G_{\alpha_0/2}(\varphi_n)) :=$$

$$= \int G_{\alpha_0/2} \varphi(x_1) \dots G_{\alpha_0/2} \varphi_n(x_n) B(dx_1) \dots B(dx_n),$$

where $G_\alpha \varphi(x) = \int_U G_\alpha(x, y) \varphi(y) dy = (I - \Delta)^{-\alpha} \varphi(x)$.

Now we give an important multiplication formula of multiple integrals or Wick products. For $\varphi \in \mathcal{E}$ and $l \leq n$, we define the functions

$$F(\varphi) = F(\varphi, x_1, \dots, x_n) = \int_U G_{\alpha_0/2}(x_1, y) \dots G_{\alpha_0/2}(x_n, y) \varphi(y) dy,$$

$$\begin{aligned} F \smile_l F(\varphi) &= F \smile_l F(\varphi, x_1, \dots, x_{n-l}, y_1, \dots, y_{n-l}) = \\ &= \int F(\varphi, x_1, \dots, x_{n-l}, u_1, \dots, u_l) F(\varphi, u_1, \dots, u_l, y_1, \dots, y_{n-l}) du_1 \dots du_l. \end{aligned}$$

If $l = 0$, $F \smile_l F = F(\varphi, x_1, \dots, x_n) F(\varphi, y_1, \dots, y_n)$, and if $l = n$,

$$F \smile_l F = \int F(\varphi, u_1, \dots, u_n)^2 du_1 \dots du_n, I_0(F \smile_l F) = F \smile_l F.$$

Since our functions are symmetric, the order of indices does not really matter. Note that the functions $F \smile_l F$ are not symmetric. The following multiplication formula holds

THEOREM 4.2. *Let $\varphi \in \mathcal{D}(U)$. With the above notations*

$$\left(\int_U :X(x)^n : \varphi(x) dx \right)^2 = I_n(F(\varphi))^2 = \sum_{l=0}^n l! \binom{n}{l}^2 I_{2(n-l)}(F \smile_l F(\varphi)).$$

Proof:

The last equality is immediate consequence of Theorem 1.3 (p.76 in [1]). So, it remains to prove the first equality. Let ε does not exceed $1/4$ of the distance between the support of φ and the boundary of U . Define $X^\varepsilon(x) = X(w_\varepsilon(x - \cdot)) = B(G_{\alpha_0/2} w_\varepsilon(x - \cdot))$, where $w_\varepsilon(x) = \varepsilon^{-2} w(x/\varepsilon)$, $w \in C_0^\infty(\mathbf{R}^2)$, $w \geq 0$, $\int w dx = 1$, $x \in \text{supp } \varphi$. Let

$$G_{\alpha_0/2} w_\varepsilon(x, y) = \int G_{\alpha_0/2}(y, z) w_\varepsilon(x - z) dz.$$

Then

$$(4.7) \quad :X^\varepsilon(x)^n := \int G_{\alpha_0/2} w_\varepsilon(x, y_1) \dots G_{\alpha_0/2} w_\varepsilon(x, y_n) B(dy_1) \dots B(dy_n).$$

Let

$$F_\varepsilon(\varphi, y_1, \dots, y_n) = \int G_{\alpha_0/2} w_\varepsilon(x, y_1) \dots G_{\alpha_0/2} w_\varepsilon(x, y_n) \varphi(x) dx.$$

So, by (4.7) we have

$$\int_U :X^\varepsilon(x)^n : \varphi(x) dx = \int F_\varepsilon(\varphi, y_1, \dots, y_n) B(dy_1) \dots B(dy_n)$$

Now by Lemma 4.1 and (4.6) we obtain the first equality of the theorem passing to the limit as $\varepsilon \rightarrow 0$. ■

For the later considerations we need some estimates of the integrals involving Wick product. Fix a measurable subset $\tilde{U} \subset U$ such that $|\tilde{U}| > 0$ and $\text{dist}(\tilde{U}, \partial U) = \tau > 0$. Let $\varepsilon < 1/4\tau < 1$, $\alpha_0 = 1$. Define for $x \in \tilde{U}$, $n \geq 1$, $l \leq n$

$$F_\varepsilon(x, x_1, \dots, x_n) = F(w_\varepsilon(x - \cdot), x_1, \dots, x_n),$$

$$F_\varepsilon \smile_l F_\varepsilon(x) = F \smile_l F(w_\varepsilon(x - \cdot), x_1, \dots, x_{n-l}, y_1, \dots, y_{n-l}) =$$

$$= H_{\varepsilon,l}(x, x_1, \dots, x_{n-l}, y_1, \dots, y_{n-l}) = H_{\varepsilon,l}(x).$$

Then we have by Theorem 4.2 for $x \in \tilde{U}$

$$(4.8) \quad \begin{aligned} & (\int_U :X(z)^n : w_\varepsilon(x-z) dz)^2 = \\ &= \sum_{l=0}^n l! \binom{n}{l}^2 I_{2(n-l)}(F_\varepsilon \smile_l F_\varepsilon(x)) = \\ &= \sum_{l=0}^n l! \binom{n}{l}^2 I_{2(n-l)}(H_{\varepsilon,l}(x)), \end{aligned}$$

where $I_0(H_{\varepsilon,n}(x)) = \int_U \int_U G_1(z, z')^n w_\varepsilon(x-z) w_\varepsilon(x-z') dz dz'$.

LEMMA 4.2. *Let $\alpha_0 = 1$. There exist some constants $c_2 > c_1 > 0$ such that*

$$\int_{\tilde{U}} \mathbf{E}^\mu I_{2(n-l)}(H_{\varepsilon,l}(x))^2 dx \leq c_2 |\log \varepsilon|^{2l}, \text{ if } 0 \leq l < n,$$

$$c_1 |\log \varepsilon|^n \leq I_0(H_{\varepsilon,n}(x)) \leq c_2 |\log \varepsilon|^n, \text{ on } \tilde{U}.$$

Proof:

Let $0 \leq l < n$. Then there exists $C > 0$ such that for all $x \in \tilde{U}$

$$H_{\varepsilon,l}(x) = H_{\varepsilon,l}(x, x_1, \dots, x_{n-l}, y_1, \dots, y_{n-l}) = \int_U \int_U G_{1/2}(x_1, z) \dots$$

$$\dots G_{1/2}(x_{n-l}, z) G_1(z, z')^l G_{1/2}(y_1, z') \dots G_{1/2}(y_{n-l}, z') w_\varepsilon(x-z) w_\varepsilon(x-z') dz dz' \leq$$

$$\leq C |\log \varepsilon|^l \tilde{H}_{\varepsilon,l}(x, x_1, \dots, x_{n-l}) \tilde{H}_{\varepsilon,l}(x, y_1, \dots, y_{n-l}),$$

where $\tilde{H}_{\varepsilon,l}(x, x_1, \dots, x_{n-l}) = \int_U G_{1/2}(x_1, z) \dots G_{1/2}(x_{n-l}, z) w_\varepsilon(x-z) dz$. Thus

$$\int_{\tilde{U}} \int_{U^2} H_{\varepsilon,l}(x, x_1, \dots, x_{n-l}, y_1, \dots, y_{n-l})^2 dx_1 \dots dx_{n-l} dy_1 \dots dy_{n-l} dx \leq$$

$$\leq C^2 |\log \varepsilon|^{2l} \int_{\tilde{U}} (\int_U \tilde{H}_{\varepsilon,l}(x, x_1, \dots, x_{n-l})^2 dx_1 \dots dx_{n-l})^2 dx =$$

$$= C^2 |\log \varepsilon|^{2l} \int_{\tilde{U}} (\int_{U^2} G_1(z, z')^{n-l} w_\varepsilon(x-z) w_\varepsilon(x-z') dz dz')^2 dx \leq$$

$$\leq c' |\log \varepsilon|^{2l}.$$

Since $I_0(H_{\varepsilon,n}(x)) = \int_U \int_U G_1(z, z')^n w_\varepsilon(x-z) w_\varepsilon(x-z') dz dz'$, the last part of the statement follows by Lemma 5.2. ■

4.1.2. *Projection approximation of Wick products.* Define the projections of X by $\Pi^M X = X^M = \sum_1^M \langle X, e_k \rangle e_k$.

COROLLARY 4.2. *For each $n \in \mathbb{N}$, $p \geq 1$, $\alpha < \alpha_0 - 1$*

$$\sup_M \mathbf{E}^\mu [| : (X^M)^n : |_\alpha] < \infty, \text{ if } \alpha \leq 2,$$

$$\sup_M | : (X^M)^n : |_{\alpha,p} < \infty, \text{ if } \alpha \geq 0,$$

$$\sup_M \mathbf{E}^\mu | \int_U : (X^M)^n : dx |^p < \infty.$$

Proof:

It is an immediate consequence of Lemma 4.1, Corollary 4.1 and Theorems 1.12, 1.13 in [9]. ■

LEMMA 4.3. *For all $n \geq 1$, $p \geq 1$,*

$$\sup_M \mathbf{E}^\mu \exp \left\{ -p \int_U : (X^M)^{2n} : dx \right\} < \infty.$$

Proof:

Let $\bar{X}^M = \sum_1^M \langle X, e_k \rangle \bar{e}_k$, where \bar{e}_k is the corresponding extension of e_k . Let $\bar{X}^{M,\epsilon}$ be a convolution of \bar{X}^M and w_ϵ . By the definition of the Wick power and Theorem 1.13 in [9] there exists a constant b independent of M such that for each $\epsilon \in (0, 1)$

$$\begin{aligned} : \bar{X}^{M,\epsilon}(x)^{2n} : &\geq -b \left(\mathbf{E}^\mu \left[(\bar{X}^{M,\epsilon}(x)^2) \right] \right)^n \geq -b \left(\mathbf{E}^\mu \left[\bar{X}^\epsilon(x)^2 \right] \right)^n \geq \\ &\geq -b \left(\int \int w_\epsilon(x-z) \bar{G}_{\alpha_0}(z, z') w_\epsilon(x-z') dz dz' \right)^n \geq \\ &\geq -b \int \int w_\epsilon(x-z) |\bar{G}_{\alpha_0}(z, z')|^n w_\epsilon(x-z') dz dz'. \end{aligned}$$

Thus

$$(4.9) \quad \int_U : \bar{X}^{M,\epsilon}(x)^{2n} : dx \geq -c' \epsilon^{-2} \int_0^{2\epsilon} |\log u|^n u du \geq -c |\log \epsilon|^n$$

for some constants c', c independent of M .

On the other hand by Lemma 4.1 c) and Theorem 1.13 in [9]

$$(4.10) \quad \sup_M \mathbf{E}^\mu \left[\left(\int_U : (\bar{X}^{M,\epsilon})^{2n} : dx - \int_U : (\bar{X}^M)^{2n} : dx \right)^2 \right] \leq C \epsilon^\eta,$$

for some constant C and $\eta > 0$.

Using (4.9) and (4.10), we derive immediately the statement (see 5.2, Lemma 5.5 and the following remarks in [9]). ■

REMARK 4.4. Denote $G_{\alpha_0}^M(x, y) = \sum_1^M (1 + \lambda_k)^{-\alpha_0} e_k(x) e_k(y)$. Then

$$\int_U \int_U |G_{\alpha_0}^M(x, y) - G_{\alpha_0}(x, y)|^2 dx dy \rightarrow 0, \text{ as } M \rightarrow \infty,$$

i.e. $G_{\alpha_0}^M(x, y)$ converges in $L_2(U \times U, dx dy)$ to $G_{\alpha_0}(x, y) = \sum_1^\infty (1 + \lambda_k)^{-\alpha_0} e_k(x) e_k(y)$.

COROLLARY 4.3. For each $n \in \mathbf{N} = \{0, 1, 2, \dots\}$

$$a) \sup_M \int_U \int_U |G_{\alpha_0}^M(x, y)|^n dx dy < \infty;$$

$$b) \int_U \int_U |G_{\alpha_0}^M(x, y) - G_{\alpha_0}(x, y)|^n dx dy \rightarrow 0, \text{ as } M \rightarrow \infty;$$

Proof:

Indeed, for each n

$$\mathbf{E}^\mu \left[\left(\int_U : (X^M)^n : dx \right)^2 \right] = \int_U \int_U G_{\alpha_0}^M(x, y)^n dx dy.$$

Thus a) follows by Corollary 4.2, and we have b) by a) and Remark 4.4. ■

Now we will prove the convergence of $: (X^M)^n :$ to $: X^n :$.

LEMMA 4.4. Let $n \in \mathbf{N}, \alpha < \alpha_0 - 1$. Then

$$a) \mathbf{E}^\mu | : (X^M)^n : - : X^n : |_\alpha^2 \rightarrow 0, \text{ as } M \rightarrow \infty, \text{ if } \alpha \leq 2,$$

$$| : (X^M)^n : - : X^n : |_{\alpha, 2} \rightarrow 0, \text{ as } M \rightarrow \infty, \text{ if } \alpha \geq 0;$$

$$b) \mathbf{E}^\mu [(\int_U : (X^M)^n : dx - \int_U : X^n : dx)^2] \rightarrow 0, \text{ as } M \rightarrow \infty.$$

Proof:

Let $\alpha < 0$. Then

$$\begin{aligned} \mathbf{E}^\mu |:(X^M)^n:-:X^n:|_\alpha^2 &= \mathbf{E}^\mu \int_U \int_U G\alpha(x, y) (:(X^M(x)^n:- \\ &\quad -:X(x)^n:) (:(X^M(y)^n:-:X(y)^n:) dx dy = \\ &= \int_U \int_U G\alpha(x, y) (G_{\alpha_0}^M(x, y)^n - G_{\alpha_0}^M(x, y)^n - G_{\alpha_0}^M(x, y)^n + \\ &\quad + G_{\alpha_0}(x, y)^n) dx dy \rightarrow 0, \text{ as } M \rightarrow \infty, \end{aligned}$$

by Corollary 4.3 and Hölder inequality. Let $0 \leq \alpha < \alpha' < \alpha_0 - 1$. By Corollary 4.2,

$$\begin{aligned} (4.11) \quad \sup_M [|:(X^M)^n:-:X^n:|_{\alpha',2} + |:(X^M)^n:-:X^n:|_{0,2}] &< \infty, \\ \sup_M \mathbf{E}^\mu [|:(X^M)^n:-:X^n:|_{H^{\alpha'}}^2 + |:(X^M)^n:-:X^n:|_{H^{\alpha'}}^2] &< \infty. \end{aligned}$$

On the other hand,

$$\begin{aligned} (4.12) \quad \mathbf{E}^\mu |:(X^M)^n:-:X^n:|_0^2 &= |:(X^M)^n:-:X^n:|_{0,2} = \\ &= \int_U (G_{\alpha_0}(x, x) - G_{\alpha_0}^M(x, x)) dx \rightarrow 0, \end{aligned}$$

as $M \rightarrow \infty$. According to [5] by interpolation inequality for each $\varepsilon > 0$ there exists a constant C_ε such that

$$\begin{aligned} |:(X^M)^n:-:X^n:|_{\alpha,2} &\leq \varepsilon \sup_M |:(X^M)^n:-:X^n:|_{\alpha',2} + \\ &\quad + C_\varepsilon |:(X^M)^n:-:X^n:|_{0,2}, \end{aligned}$$

and a) follows by (4.11) and (4.12). Now

$$\begin{aligned} \mathbf{E}^\mu [(\int_U : (X^M)^n dx - \int_U : X^n : dx)^2] &= \\ &= \int_U \int_U (G_{\alpha_0}(x, y)^n - G_{\alpha_0}^M(x, y)^n) dx dy \rightarrow 0, \end{aligned}$$

as $M \rightarrow \infty$, by Corollary 4.3. ■

Owing to Lemma 4.3 and Lemma 4.4(b), we have

COROLLARY 4.4. (*cf. Lemma 5.5 in [9]*). *For every $\lambda > 0$ and $n \geq 1$,*

$$\exp\{-\lambda \int_U : X^{2n} : dx\} \in \cap_{p \geq 1} L_p(d\mu)$$

and

$$\mathbf{E}^\mu |\exp\{-\lambda \int_U : X^{2n} : dx\} - \exp\left\{-\lambda \int_U : (X^M)^{2n} : dx\right\}|^p \rightarrow 0,$$

as $M \rightarrow \infty$ for all $p \geq 1$.

4.2. Stationary Solutions of Certain SPDE's. We shall consider the following Cauchy problems in \mathcal{E}'

$$(4.13) \quad \begin{cases} dX_t = & -1/2(I - \Delta)^\varepsilon X_t + (I - \Delta)^\gamma dW_t, \\ X_0(\mathbf{P}) = & \mu = N(0, (I - \Delta)^{-\alpha_0}), \end{cases}$$

$$(4.14) \quad \begin{cases} dX_t = & -1/2(I - \Delta)^\varepsilon X_t dt - \lambda m(I - \Delta)^{2\gamma} : X^{2m-1} : dt + \\ & +(I - \Delta)^\gamma dW_t, \\ X_0(\mathbf{P}) = & \mu' = c_1 \exp\{-\lambda \int_U : X^{2m} : dx\} \mu(dX), \end{cases}$$

where $\alpha_0 = \varepsilon - 2\gamma, \varepsilon > 0, m \in \mathbb{N}$, W is a cylindrical Brownian motion in $\mathcal{H}^0 = L_2(U, dx)$.

It is easy to solve (4.13).

REMARK 4.5. The stationary solution to (4.13) can be given as $X_t = \sum_1^\infty X_t^k e_k$, where

$$\begin{cases} dX_t^k = & -1/2(1 + \lambda_k)^\varepsilon X_t^k dt + (1 + \lambda_k)^\gamma dW_t, \\ X_0^k(\mathbf{P}) = & N(0, (1 + \lambda_k)^{-\alpha_0}), \end{cases}$$

and W^k are independent standard Wiener processes. Thus

$$X_t^k = X_0^k e^{-1/2(1+\lambda_k)^\varepsilon t} + \int_0^t e^{-1/2(1+\lambda_k)^\varepsilon(t-s)} (1 + \lambda_k)^\gamma dW_s^k.$$

We want to construct a stationary weak solution to (4.14). Let Ω be the space of continuous \mathcal{E}' -valued trajectories on $[0, \infty)$, X_t be a canonical process in Ω , $\mathcal{F}_t = \sigma(X_s, s \leq t)$, $\mathcal{F} = \sigma(X_s, s \geq 0)$.

DEFINITION 4.1. a) A stationary measure \mathbf{P} on (Ω, \mathcal{F}) is a stationary solution of (4.14) if

$$\begin{aligned} \mathcal{L}_t &= X_t - X_0 + 1/2 \int_0^t (I - \Delta)^\varepsilon X_s ds + \\ &\quad \int_0^t \lambda m(I - \Delta)^{2\gamma} : X_s^{2m-1} : ds = (I - \Delta)^\gamma W_t, \end{aligned}$$

$$X_0(\mathbf{P}) = c_1 \exp\{-\lambda \int_U : X^{2m} : dx\} \mu(dX),$$

where W is a cylindrical Brownian motion in \mathcal{H}^0 ;

b) A stationary measure \mathbf{P}_0 on (Ω, \mathcal{F}) is a stationary solution of (4.13) if

$$\begin{aligned} \mathcal{L}_t &= X_t - X_0 + 1/2 \int_0^t (I - \Delta)^\varepsilon X_s ds = (I - \Delta)^\gamma W_t, \\ X_0(\mathbf{P}) &= \mu, \end{aligned}$$

where W is a cylindrical Brownian motion in \mathcal{H}^0 ;

It follows immediately from this definition the following characterization of a stationary solution.

REMARK 4.6. A stationary measure \mathbf{P} on (Ω, \mathcal{F}) is a stationary solution of (4.14) if and only if for each $s \leq t, \varphi \in \mathcal{E}$ the process $\langle \mathcal{L}_t, \varphi \rangle$ has independent increments and

$$\mathbf{P} \exp\{i \langle \mathcal{L}_t - \mathcal{L}_s, \varphi \rangle\} = \exp\{-1/2(t-s)|(I - \Delta)^\gamma \varphi|_0^2\},$$

$$X_0(\mathbf{P}) = c_1 \exp\{-\lambda \int_U : X^{2m} : dx\} \mu(dX).$$

For $M \in \mathbf{N}$ consider on \mathcal{E}' a measure

$$\mu'(dX) = c_M \exp\{-\lambda \int_U : (\sum_1^M \langle X, e_l \rangle e_l(x))^{2m} : dx\} \mu(dX).$$

Let $\xi_k = \langle X, e_k \rangle$, $k \geq 1$, and (W_t^k) be the sequence of independent standard Brownian motions such that (ξ_k) and (W_t^k) are also independent. For each $M \in \mathbf{N}$ we construct an approximation \mathbf{P}^M to the solution of (4.14) by solving the following infinite system of equations

$$(4.15) \quad \begin{aligned} dY_t^k &= -1/2(1 + \lambda_k)^\varepsilon Y_t^k dt + (1 + \lambda_k)^\gamma dW_t^k - \\ &\quad - \lambda m(1 + \lambda_k)^{2\gamma} \int_U : (\sum_1^M Y_t^l e_l(x))^{2m-1} : e_k(x) dx dt, \quad k \leq M, \\ dY_t^k &= -1/2(1 + \lambda_k)^\varepsilon X_t^k dt + (1 + \lambda_k)^\gamma dW_t^k, \quad k > M, \\ Y_0^k &= \xi_k, \quad k \geq 1. \end{aligned}$$

The problem is essentially a finite-dimensional one and it admits a unique stationary solution. The distribution of $Y_t = \sum_k Y_t^k e_k$ will define a stationary measure \mathbf{P}^M on Ω such that

$$(4.16) \quad \begin{aligned} dX_t &= -1/2(I - \Delta)^\varepsilon X_t dt + (I - \Delta)^\gamma dW_t - \\ &\quad - \lambda m(I - \Delta)^{2\gamma} \Pi_M : (\Pi_M X_t)^{2m-1} : dt, \\ X_0(\mathbf{P}^M) &= c_M \exp\{-\lambda \int_U : (\Pi_M X)^{2m} : dx\} \mu(dX). \end{aligned}$$

LEMMA 4.5. *There exists a stationary \mathbf{P}^M such that (4.16) holds.*

Proof:

By the definition of the Wick power

$$: (\sum_1^M \langle X, e_l \rangle e_l(x))^{2m} := \sum_{l=0}^m \frac{(2m)!}{l!(2m-2l)!} (\sum_{j=1}^M \langle X, e_j \rangle e_j(x))^{2m-2l} (-1/2 G_{\alpha_0}^M(x, x))^l.$$

Introduce the functions on \mathbf{R}^M

$$F(z) = -\lambda \int_U \sum_{l=0}^m \frac{(2m)!}{l!(2m-2l)!} (\sum_{j=1}^M z_j e_j(x))^{2m-2l} (-1/2 G_{\alpha_0}^M(x, x))^l dx,$$

$$K(z) = -1/2 \sum_1^M (1 + \lambda_k)^{\alpha_0} z_k^2.$$

There exist some constants b, b', K such that

$$(4.17) \quad -b' H(z) - K \leq F(z) \leq K - b H(z)$$

where $H(z) = \int_U (\sum_{j=1}^M z_j e_j(x))^{2m} dx$. Since (e_j) is a CONS in L_2 , there is $\eta > 0$ such that

$$(4.18) \quad \eta |z|^{2m} \leq H(z) \leq \eta^{-1} |z|^{2m}.$$

We notice here that the equation (4.15) can be rewritten in the following form for $k \leq M$

$$(4.19) \quad d\bar{Y}_t^k = 1/2(1 + \lambda_k)^{2\gamma} [\partial_k K(\bar{Y}_t) + \partial_k F(\bar{Y}_t)] dt + (1 + \lambda_k)^\gamma dW_t^k, \quad k \leq M,$$

where $\bar{Y}_t = (\bar{Y}_t^1, \dots, \bar{Y}_t^M)$. We will find a stationary solution of (4.19) by constructing an approximating sequence of stationary solutions. Let g be an increasing smooth function on \mathbf{R} such that $g(r) = r$, if $|r| \leq 1$, $g(r) = 2$, if $|r| \geq 2$. Let $g_\delta(r) = \delta^{-1}g(\delta r)$, $F^\delta(z) = g_\delta(F(z))$. By (4.17), (4.18) all the derivatives of F^δ are bounded. So for each $x = (x_1, \dots, x_M)$ there exists a unique solution (depending smoothly on x) $\bar{Y}_t^\delta = (\bar{Y}_t^{\delta,k})_{1 \leq k \leq M} = \bar{Y}_t^\delta(x)$ of the following equation

$$(4.20) \quad \begin{aligned} d\bar{Y}_t^{\delta,k} &= 1/2(1 + \lambda_k)^{2\gamma}[\partial_k K(\bar{Y}_t) + \partial_k F^\delta(\bar{Y}_t)]dt + (1 + \lambda_k)^\gamma dW_t^k, \\ \bar{Y}_0 &= x. \end{aligned}$$

Applying Ito formula we have

$$\begin{aligned} &\sum_{k=1}^M (1 + \lambda_k)^{-2\gamma} (\bar{Y}_t^{\delta,k})^2 \\ &= \sum_{k=1}^M (1 + \lambda_k)^{-2\gamma} x_k^2 - \int_0^t \sum_{k=1}^M (1 + \lambda_k)^{\varepsilon-2\gamma} \bar{Y}_s^{\delta,k} ds \\ &\quad - 2\lambda m \int_0^t \int_U :(\sum_1^M \bar{Y}_s^{\delta,k} e_k(y))^2 : ds \\ &\quad + Mt + 2 \int_0^t \sum_1^M (1 + \lambda_k)^{-\gamma} \bar{Y}_s^{\delta,k} dW_s^k. \end{aligned}$$

Thus for each $T > 0$, x

$$(4.21) \quad \sup_{\delta} \mathbf{E} \sup_{t \leq T} |\bar{Y}_t^\delta(x)|^2 < \infty.$$

For $\delta < \delta'$ we define stopping times $\tau_\delta = \inf(t : |F(\bar{Y}_t^\delta(x))| \geq \delta'^{-1})$, $\tau_{\delta'} = \inf(t : |F(\bar{Y}_t^{\delta'}(x))| \geq \delta'^{-1})$, $\tau = \tau_\delta \wedge \tau_{\delta'}$. Then $\bar{Y}_{\wedge \tau}^\delta = \bar{Y}_{\wedge \tau}^{\delta'} \mathbf{P}\text{-a.s.}$ So $\tau_\delta = \tau_{\delta'} = \tau \mathbf{P}\text{-a.s.}$ By (4.21), (4.17), (4.18)

$$\mathbf{P}(\tau_{\delta'} < T) \rightarrow 0, \text{ as } \delta' \rightarrow 0.$$

Thus there exists a measurable family $\bar{Y}_t = \bar{Y}_t(x)$, $x \in \mathbf{R}^M$, of stochastic processes such that $\bar{Y}_{\wedge \tau_\delta}(x) = \bar{Y}_{\wedge \tau_{\delta'}}(x)$ for each $\delta' > 0$. We can write also (because of the estimate (4.21) for each x)

$$(4.22) \quad \mathbf{E} \sup_{t \leq T} |\bar{Y}_t^{\delta'}(x) - \bar{Y}_t(x)| \rightarrow 0, \text{ as } \delta' \rightarrow 0,$$

and $\bar{Y}_t = \bar{Y}_t(x)$ is a solution of (4.19) starting at x , i.e. $\bar{Y}_0 = x$. On the other hand for each δ \bar{Y}^δ is a diffusion associated to the operator

$$Af(x) = 1/2 \sum_{k=1}^M (1 + \lambda_k)^{2\gamma} \partial_{kk}^2 f(x) + 1/2 \sum_{k=1}^M (1 + \lambda_k)^{2\gamma} [\partial_k K(x) + \partial_k F^\delta(x)] \partial_k f(x).$$

Thus its adjoint operator is

$$\begin{aligned} A^* f(x) &= 1/2 \sum_{k=1}^M (1 + \lambda_k)^{2\gamma} \partial_{kk}^2 f(x) \\ &\quad - 1/2 \sum_{k=1}^M (1 + \lambda_k)^{2\gamma} [\partial_k K(x) + \partial_k F^\delta(x)] \partial_k f(x) - \\ &\quad - 1/2 \sum_{k=1}^M (1 + \lambda_k)^{2\gamma} (\partial_{kk}^2 K(x) + \partial_{kk}^2 F^\delta(x)) f(x) . \end{aligned}$$

By [23] for each $\varphi \in C_0^\infty(\mathbf{R}^M)$ there exists a classical solution of the Cauchy problem

$$(4.23) \quad \begin{aligned} \partial_t u &= Au, \text{ in } \mathbf{R}^M, \\ u(0, \cdot) &= \varphi. \end{aligned}$$

By Ito formula, $u(t, x) = \mathbf{E}\varphi(\bar{Y}_t^\delta(x))$. Let $\Phi^\delta(x) = \exp\{K(x) + F^\delta(x)\}$. Since $A^*\Phi^\delta(x) = 0$, we have each $t > 0$

$$\begin{aligned} \frac{d}{dt} \int \mathbf{E}\varphi(\bar{Y}_t^\delta(x)) \Phi^\delta(x) dx &= \int \partial_t u(t, x) \Phi^\delta(x) dx = \\ &= \int Au(t, x) \Phi^\delta(x) dx = \int u(t, x) A^* \Phi^\delta(x) dx = 0. \end{aligned}$$

Thus $c_{M,\delta} \Phi^\delta(x) dx$ is an invariant measure for Y^δ . Let $\Phi(x) = \exp\{K(x) + F(x)\}$, $\varphi \in C_0^\infty(\mathbf{R}^M)$. Then for each $t \geq 0$ by (4.22)

$$\int \mathbf{E}\varphi(\bar{Y}_t^\delta(x)) \Phi^\delta(x) dx \rightarrow \int \mathbf{E}\varphi(\bar{Y}_t(x)) \Phi(x) dx, \text{ as } \delta \rightarrow 0,$$

i.e. for each $\varphi \in C_0^\infty(\mathbf{R}^M)$, $\int \mathbf{E}\varphi(\bar{Y}_t(x)) \Phi(x) dx$ does not depend on t . Let $Y_t^k = \bar{Y}_t^k(\xi^M)$, $k \leq M$, $\xi^M = (\xi_1, \dots, \xi_M)$. Then $(Y_t^k)_{k \leq M}$ is a stationary solution of (4.19), $Y_0^k = \xi_k$, $k \leq M$. For each $k > M$, we solve the one-dimensional equation

$$dY_t^k = -1/2(1 + \lambda_k)^\epsilon Y_t^k dt + (1 + \lambda_k)^\gamma dW_t^k,$$

with initial condition $Y_0^k = \xi_k$. So we get a stationary solution $(Y_t^k)_{k \geq 1}$ of (4.15) and the distribution of $Y_t = \sum_k Y_t^k e_k$ defines the needed measure \mathbf{P}^M on the space of trajectories Ω . ■

Now we prove the main statement regarding the existence of the stationary solutions of (4.14).

THEOREM 4.3. *Let $0 \leq \alpha_0 - 1 \leq 2$ and $m \in \mathbf{N}$. Then there exists a stationary solution \mathbf{P} of (4.14). Moreover, if $2\gamma < \alpha_0 - 1$, then it is unique and $\mathbf{P}|_{\mathcal{F}_t} \sim \mathbf{P}_0|_{\mathcal{F}_t}$ for all $t > 0$, where \mathbf{P}_0 is the distribution of (4.13)*

Proof:

Let $\phi^M(X) = c_M \exp\{-\lambda \int_U :(\Pi_M X)^{2m} : dx\}$, $\phi(X) = c_1 \exp\{-\lambda \int_U :X^{2m} : dx\}$. For $M \geq 1$ we have by Lemma 4.5 a stationary approximation \mathbf{P}^M such that (4.16) holds. The sequence $\{\mathbf{P}^M, M \geq 1\}$ is relatively compact if, for each $\varphi \in \mathcal{E}$ the sequence of one-dimensional distributions $\{X_\cdot(\varphi)(\mathbf{P}^M), M \geq 1\}$ is relatively compact (see [22]). Let $\beta \in \mathbf{R}$ be such that $\beta + 4\gamma < \alpha_0 - 1$. Then by Hölder

inequality and Corollary 4.2 for each $t \geq 0$

$$\begin{aligned}
& \sup_M \mathbf{P}^M |((I - \Delta)^{2\gamma} : (\Pi_M X_t)^{2m-1})|_{\beta}^4 = \\
& = \sup_M \mathbf{E}^{\mu} \phi^M(X) | : (\Pi_M X)^{2m-1} : |_{\beta+4\gamma}^4 \leq \\
(4.24) \quad & \leq C \sup_M (\mathbf{E}^{\mu} | : (\Pi_M X)^{2m-1} : |_{\beta+4\gamma}^8)^{1/2} \leq \\
& = C \sup_M (\mathbf{E}^{\mu} | : (\Pi_M X)^{2m-1} : |_{\beta+4\gamma}^2)^2 < \infty.
\end{aligned}$$

Thus for each $\varphi \in \mathcal{E}$ there exists a constant C independent of M such that for each $0 < s \leq t$

$$\mathbf{P}^M |X_t(\varphi) - X_s(\varphi)|^4 \leq C|t - s|^2.$$

This means (see [6]) that the sequence $\{X_0(\varphi)(\mathbf{P}^M), M \geq 1\}$ is relatively compact. Let

$$\begin{aligned}
\mathcal{L}_t^M &= X_t - X_0 + \int_0^t 1/2(I - \Delta)^{\varepsilon} X_s ds + \\
& + \int_0^t \lambda m(I - \Delta)^{2\gamma} \Pi_M : (\Pi_M X_s)^{2m-1} : ds = \eta_t + b_t^M,
\end{aligned}$$

where $\eta_t = X_t + \int_0^t 1/2(I - \Delta)^{\varepsilon} X_s ds$, $b_t^M = \int_0^t \lambda m(I - \Delta)^{2\gamma} \Pi_M : (\Pi_M X_s)^{2m-1} : ds$. Then for each $s \leq t$, $\varphi \in \mathcal{E}$

$$\begin{aligned}
(4.25) \quad & \mathbf{P}^M \exp\{i(\mathcal{L}_t^M - \mathcal{L}_s^M, \varphi)\} = \exp\{-1/2|(I - \Delta)^{\gamma} \varphi|_0^2(t - s)\}, \text{ and} \\
& X_0(\mathbf{P}^M) = \phi^M(X) \mu(dX).
\end{aligned}$$

By Corollaries 4.2 and 4.4,

$$(4.26) \quad \phi^M(X) \mu(dX) \rightarrow \phi(X) \mu(dX) \text{ by variation,}$$

as $M \rightarrow \infty$. Now we estimate the 'bad' part b_t^M of the drift coefficient. Let $b_t = \int_0^t \lambda m(I - \Delta)^{2\gamma} : X_s^{2m-1} : ds$. Then for $\beta + 4\gamma < \alpha_0 - 1$, $M \geq M'$

$$\begin{aligned}
& \mathbf{P}^M |b_t^M - b_t|^2 \leq C \int_0^t \mathbf{P}^M |(I - \Delta)^{2\gamma} \Pi_{M'} : (\Pi_{M'} X_s)^{2m-1} : - \\
& - (I - \Delta)^{2\gamma} \Pi_{M'} : X_s^{2m-1} :|_{\beta}^2 ds + \\
& + C \int_0^t \mathbf{P}^M |(I - \Delta)^{2\gamma} \Pi_{M'} : X_s^{2m-1} : - (I - \Delta)^{2\gamma} : X_s^{2m-1} :|_{\beta}^2 ds = \\
& = A_1 + A_2.
\end{aligned}$$

By stationarity of \mathbf{P}^M and Hölder inequality

$$\begin{aligned}
A_1 &\leq C \mathbf{E}^{\mu} \phi^M(X) | : (\Pi_{M'} X)^{2m-1} : - : X^{2m-1} : |_{\beta+4\gamma}^2 \leq \\
&\leq C \mathbf{E}^{\mu} | : (\Pi_{M'} X)^{2m-1} : - : X^{2m-1} : |_{\beta+4\gamma}^4, \\
A_2 &\leq C \mathbf{E}^{\mu} \phi^M(X) |\Pi_{M'} : X^{2m-1} : - : X^{2m-1} :|_{\beta+4\gamma}^2 \leq \\
&\leq C \mathbf{E}^{\mu} |\Pi_{M'} : X^{2m-1} : - : X^{2m-1} :|_{\beta+4\gamma}^4.
\end{aligned}$$

Thus by Lemma 4.4

$$(4.27) \quad \lim_{M'} \sup_{M \geq M'} \mathbf{P}^M |b_t^{M'} - b_t|_\beta^2 = 0 .$$

Since $\{\mathbf{P}^M, M \geq 1\}$ is relatively compact, there exist a subsequence (M_k) and a measure \mathbf{P} on (Ω, \mathcal{F}) such that (\mathbf{P}^{M_k}) converges weakly to \mathbf{P} . Because of (4.26) \mathbf{P} is a stationary measure and $X_0(\mathbf{P}) = \mu'(dX) = \phi(X)\mu(dX)$. Similarly, we have

$$(4.28) \quad \mathbf{P}|b_t^{M'} - b_t|_\beta^2 \rightarrow 0, \text{ as } M' \rightarrow \infty .$$

Then for each M' , $s \leq t$, $\varphi \in \mathcal{E}$

$$\begin{aligned} & \exp\{-1/2(t-s)|(I-\Delta)^\gamma \varphi|_0^2\} = \mathbf{P}^{M_k} \exp\{i(\mathcal{L}_t^{M_k} - \mathcal{L}_s^{M_k}, \varphi)\} = \\ & = \mathbf{P}^{M_k} [\exp\{i\langle \eta_t - \eta_s, \varphi \rangle\} \exp\{i\langle b_t^{M_k} - b_s^{M_k}, \varphi \rangle\} - \\ & - \exp\{i\langle \eta_t - \eta_s, \varphi \rangle\} \exp\{i\langle b_t^{M'} - b_s^{M'}, \varphi \rangle\}] + \\ & + \mathbf{P}^{M_k} [\exp\{i\langle \eta_t - \eta_s, \varphi \rangle\} \exp\{i\langle b_t^{M'} - b_s^{M'}, \varphi \rangle\}] = A_1 + A_2 . \end{aligned}$$

For each M' $A_2 \rightarrow \mathbf{P}[\exp\{i\langle \eta_t - \eta_s, \varphi \rangle\} \exp\{i\langle b_t^{M'} - b_s^{M'}, \varphi \rangle\}]$, as $k \rightarrow \infty$, by weak convergence of measures. On the other hand, by (4.27)

$$\limsup_k |A_1| \leq C \sup_{M_k \geq M'} (\mathbf{P}^{M_k} |b_t^{M'} - b_t|_\beta + \mathbf{P}^{M_k} |b_s^{M'} - b_s|_\beta) .$$

Now by (4.27) and (4.28) it follows that

$$\exp\{-1/2(t-s)|(I-\Delta)^\gamma \varphi|_0^2\} = \mathbf{P} \exp\{i(\mathcal{L}_t - \mathcal{L}_s, \varphi)\} .$$

According to Theorem 3.2 and Corollary 3.3 in order to have the local equivalence of measures and uniqueness of \mathbf{P} it suffices that for all $t > 0$

$$(\mathbf{P} + \mathbf{P}_0) \int_0^t \int_U ((I-\Delta)^\gamma : X_t^{2m-1} :)^2 dx dt < \infty .$$

By Hölder inequality and supercontractivity again

$$\begin{aligned} & \int_0^1 \int_U \mathbf{P}((I-\Delta)^\gamma : X_t^{2m-1} :)^2 dx dt = \\ & = \int_U \mathbf{E}^\mu \phi(X)((I-\Delta)^\gamma : X^{2m-1} :)^2 dx \leq \\ & \leq C \mathbf{E}^\mu |X_t^{2m-1}|_{2\gamma}^4 \leq C' (\mathbf{E}^\mu |X^{2m-1}|_{2\gamma}^2)^2 . \end{aligned}$$

The last term is finite by Lemma 4.1, if $2\gamma < \alpha_0 - 1$. ■

We shall prove now that for $2\gamma < \alpha_0 - 1$ the unique stationary solution to (4.14) in this case is a Markovian one.

Let $\mathbb{F} = (\mathcal{F}_t)$ be a canonical filtration of σ -algebras on Ω , i.e. $\mathcal{F}_t = \sigma(X_s, s \leq t)$, $t \geq 0$. For a given probability measure ν on \mathcal{E}' we introduce the set $\mathcal{S}(\nu)$ of all probability measures \mathbf{P} on (Ω, \mathcal{F}) such that $X_0(\mathbf{P}) = \nu$ and the process $\mathcal{L}_t = X_t + 1/2 \int_0^t (I-\Delta)^\epsilon X_r dr$ is a Wiener process starting at the moments s from zero with a corresponding covariance operator $(I-\Delta)^{2\gamma} t$. Similarly, we introduce the set $\mathcal{S}'(\nu)$ of all probability measures \mathbf{P} on (Ω, \mathcal{F}) such that $X_0(\mathbf{P}) = \nu$, $\int_0^t |X_r^{2m-1}|_{2\gamma}^2 < +\infty$ \mathbf{P} -a.s. for all $t > 0$ and the process $\mathcal{L}_t = X_t + 1/2 \int_0^t (I-\Delta)^\epsilon X_r dr + \lambda m \int_0^t (I-$

$\Delta)^{2\gamma} : X_r^{2m-1} : dr$ is a continuous \mathcal{E}' -valued (\mathcal{F}_t) -martingale with covariance operator $(I - \Delta)^{2\gamma} t$. One can say \mathbf{P} is a weak solution to (4.13) (respectively (4.14)) starting at the time moment 0 with a given initial distribution ν . If $\nu = \delta_\xi, \xi \in \mathcal{E}'$, then we denote the corresponding set $\mathcal{S}(\xi)$ (respectively, $\mathcal{S}'(\xi)$) and say that \mathbf{P} is a weak solution of (4.13) (respectively, of (4.14)) starting at the time moment 0 from ξ .

REMARK 4.7. Obviously for each $\xi \in \mathcal{E}'$ there exists a unique solution $\mathbf{P}_\xi^0 \in \mathcal{S}(\xi)$. On the other hand, if $\bar{\mathbf{P}}_\xi \in \mathcal{S}'(\xi)$, then by definition for each $t > 0$ $\int_0^t | : X_r^{2m-1} : |_{2\gamma}^2 dr < +\infty$ $\bar{\mathbf{P}}_\xi$ -a.s. and it is unique by Theorem 3.2. Moreover, $\bar{\mathbf{P}}_\xi|_{\mathcal{F}_t} = \rho_t \mathbf{P}_\xi^0|_{\mathcal{F}_t}$, where

$$\begin{aligned}\rho_t &= \exp\left\{\int_0^t | : X_r^{2m-1} : d\mathcal{L}_r - 1/2 \int_0^t | : X_r^{2m-1} : |_{2\gamma}^2 dr\right\}, \\ \mathcal{L}_t &= X_t + 1/2 \int_0^t (I - \Delta)^\epsilon X_r dr\end{aligned}$$

(see Theorem 3.2).

Let $\mathfrak{A} = \{\xi \in \mathcal{E}' : \int_0^t | : X_r^{2m-1} : |_{2\gamma}^2 dr < \infty \text{ } \mathbf{P}_\xi^0\text{-a.s. for all } t > 0 \text{ and } \mathbf{P}_\xi^0 \rho_t = 1 \text{ for all } t > 0\}$. It is a measurable subset of \mathcal{E}' . Let G_ξ be a probability measure on Ω such that $\rho_t \mathbf{P}_\xi^0|_{\mathcal{F}_t} = G_\xi|_{\mathcal{F}_t}$ for all $t > 0$, if $\xi \in \mathfrak{A}$, and $G_\xi = \delta_\xi$, if $\xi \notin \mathfrak{A}$. We notice here that if $\xi \in \mathfrak{A}$, then by Theorem 3.2 $G_\xi \in \mathcal{S}'(\xi)$, it is unique and $G_\xi|_{\mathcal{F}_t} \sim \mathbf{P}_\xi^0|_{\mathcal{F}_t}$ for all $t > 0$.

Let $\mathbf{P} \in \mathcal{S}(\nu)$ (respectively, $\bar{\mathbf{P}} \in \mathcal{S}'(\nu)$) and τ be a \mathbb{F} -stopping time. There exist the regular conditional probabilities $\mathbf{P}_\omega^\tau(\cdot) = \mathbf{P}(\cdot|\mathcal{F}_\tau)$, $\bar{\mathbf{P}}_\omega^\tau(\cdot) = \bar{\mathbf{P}}(\cdot|\mathcal{F}_\tau)$. Obviously

$$\mathbf{P}_\omega^\tau(X_t = X_t(\omega), t \leq \tau(\omega)) = 1, \quad \mathbf{P}\text{-a.s.}$$

and a similar equality holds for $\bar{\mathbf{P}}$. Let us introduce the mappings $\mathcal{J}_\tau : \Omega \times \Omega \rightarrow \Omega$, by formula

$$\mathcal{J}_\tau(\omega, \omega')_t = \omega'_{\tau(\omega)+t}, \quad t \geq 0.$$

The following statement concerning the solutions of martingale problems is a standard one.

LEMMA 4.6. Let $\bar{\mathbf{P}} \in \mathcal{S}'(\nu)$ (respectively, $\mathbf{P} \in \mathcal{S}(\nu)$) and τ be a bounded \mathbb{F} -stopping time. Then $\mathcal{J}_\tau(\omega, \cdot)(\bar{\mathbf{P}}_\omega^\tau) \in \mathcal{S}'(X_{\tau(\omega)})$ $\bar{\mathbf{P}}$ -a.s. (respectively $\mathcal{J}_\tau(\omega, \cdot)(\mathbf{P}_\omega^\tau) \in \mathcal{S}(X_{\tau(\omega)})$ \mathbf{P} -a.s.).

Proof:

Let $\varphi \in \mathcal{E}$, $\Gamma \in \mathcal{F}_\tau$, f be an arbitrary bounded measurable function on $(\mathcal{E}')^n$, $r_1 < \dots < r_n \leq s < t$. Then

$$\begin{aligned}0 &= \int_\Gamma \int [\mathcal{L}_{t+\tau}(\varphi) - \mathcal{L}_{s+\tau}(\varphi)] f(X_{\tau+r_1}, \dots, X_{\tau+r_n}) \bar{\mathbf{P}}_\omega^\tau(d\omega') \bar{\mathbf{P}}(d\omega) = \\ &= \int_\Gamma \int [\mathcal{L}_t(\varphi) - \mathcal{L}_s(\varphi)] f(X_{r_1}, \dots, X_{r_n}) \mathcal{J}_\tau(\omega, \cdot)(\bar{\mathbf{P}}_\omega^\tau)(d\omega') \bar{\mathbf{P}}(d\omega).\end{aligned}$$

according to the definition of conditional expectation. Also,

$$\int_\Gamma \bar{\mathbf{P}}_\omega^\tau \left(\int_{\tau(\omega)}^{\tau(\omega)+t} | : X_r^{2m-1} : |_{2\gamma}^2 dr < \infty \right) \bar{\mathbf{P}}(d\omega) =$$

$$\int_\Gamma \mathcal{J}_\tau(\omega, \cdot)(\bar{\mathbf{P}}_\omega^\tau) \left(\int_0^t | : X_r^{2m-1} : |_{2\gamma}^2 dr < \infty \right) \bar{\mathbf{P}}(d\omega) = \bar{\mathbf{P}}(\Gamma).$$

Since φ, s, t, f, Γ are arbitrary we derive the statement in a standard way (see [26]) for $\bar{\mathbf{P}}$ and similarly for \mathbf{P} . ■

PROPOSITION 4.1. *Let $2\gamma < \alpha_0 - 1$ and $\mathfrak{A} \subset \mathcal{E}'$ be the set defined in Remark 4.7. Then*

a) $\mu(\mathfrak{A}) = 1$, i.e. for μ -almost all $\xi \in \mathcal{E}'$ there exists a unique $\mathbf{P}_\xi \in \mathcal{S}'(\xi)$ and for each $t > 0$ $\mathbf{P}_\xi|_{\mathcal{F}_t} \sim \mathbf{P}_\xi^0|_{\mathcal{F}_t}$ (in fact, $G_\xi = \mathbf{P}_\xi$ μ -a.s.)

b) for each probability measure $\tilde{\mu}$ on \mathfrak{A} (in particular for all probability measures satisfying $\tilde{\mu} \prec \mu$) there exists a unique $\mathbf{P}_{\tilde{\mu}} \in \mathcal{S}'(\tilde{\mu})$ and $\mathbf{P}_{\tilde{\mu}} = \int G_\xi \mu(d\xi)$ (in particular, $\int G_\xi \mu'(d\xi) = \mathbf{P} \in \mathcal{S}'(\mu')$).

Proof:

By Theorem 4.3 there exists a unique $\mathbf{P} \in \mathcal{S}'(\mu')$ and for each $t > 0$ $\mathbf{P} \int_0^t | : X_r^{2m-1} : |_{2\gamma}^2 dr < \infty$. Also there exists a unique $\mathbf{P}^0 \in \mathcal{S}(\mu)$ and $\mathbf{P}^0 \int_0^t | : X_r^{2m-1} : |_{2\gamma}^2 dr < \infty$ for each t . Applying Lemma 4.6 for \mathbf{P} and \mathbf{P}^0 (the stopping time $\tau = 0$) we obtain by Remark 4.7 that $\mathbf{P}(\cdot | \mathcal{F}_0) = G_{X_0}$, $X_0 \in \mathfrak{A}$ \mathbf{P} -a.s. Since $\mu' \sim \mu$, we have $\mu'(\mathfrak{A}) = \mu(\mathfrak{A}) = 1$. Then, obviously $\mathbf{P}_{\tilde{\mu}} = \int G_\xi \tilde{\mu}(d\xi) \in \mathcal{S}'(\tilde{\mu})$. Also, $\mathbf{P}_{\tilde{\mu}}$ a.s. $\int_0^t | : X_r^{2m-1} : |_{2\gamma}^2 dr < +\infty$ for each $t > 0$. Let $\mathbf{P}_\xi^0 \in \mathcal{S}(\xi)$, $\xi \in \mathcal{E}'$, $\mathbf{P}_\xi^0 = \int \mathbf{P}_\xi^0 \tilde{\mu}(d\xi)$. Then \mathbf{P}_ξ^0 is a unique solution of (4.13) with the initial distribution $\tilde{\mu}$. The second statement now follows by our results. ■

Now we discuss the Markov property of $\mathbf{P} \in \mathcal{S}'(\tilde{\mu})$ for a probability measure $\tilde{\mu}$ on \mathfrak{A} (in particular for any $\tilde{\mu} \prec \mu$). On the set of trajectories Ω we define the shift operators $\theta_t : w \mapsto w_{+t}$, $w \in \Omega$, $t \geq 0$.

PROPOSITION 4.2. *Let $2\gamma < \alpha_0 - 1$, $\tilde{\mu}$ be a probability measure on $\mathfrak{A} \subset \mathcal{E}'$ and $\bar{\mathbf{P}} \in \mathcal{S}'(\tilde{\mu})$, where \mathfrak{A} is defined in Remark 4.7. Then for each bounded \mathbb{F} -stopping times τ , and all $B \subset \Omega$*

$$\bar{\mathbf{P}}(\theta_\tau^{-1}(B) | \mathcal{F}_\tau) = G_{X_\tau}(B) \quad \bar{\mathbf{P}}\text{-a.s.},$$

i.e. it is a strong Markov process with respect to the natural filtration \mathbb{F} .

Proof:

Let τ be a bounded \mathbb{F} -stopping time and f be a bounded measurable function on \mathcal{E}' . Then

$$\bar{\mathbf{P}}(f(X_{\tau+s}) | \mathcal{F}_\tau) = \bar{\mathbf{P}}_\omega^\tau f(X_{\tau(\omega)+s}) = G_{X_{\tau(\omega)}} f(X_s) \quad \bar{\mathbf{P}}\text{-a.s.},$$

where G is a function defined in Remark 4.7. The first equality here is a definition and the second follows by Remark 4.7 and Lemma 4.6. ■

4.2.1. Ergodicity. In our consideration of ergodicity we follow [3]. Let $2\gamma < \alpha_0 - 1$. Thus by Proposition 4.2 and Theorem 4.3 there exists a unique stationary solution $\mathbf{P} \in \mathcal{S}'(\mu')$ and for all bounded \mathbb{F} -stopping times τ and measurable $B \subset \Omega$

$$\mathbf{P}(\theta_\tau^{-1}(B) | \mathcal{F}_\tau) = G_{X_\tau}(B) \quad \mathbf{P}\text{-a.s.}$$

We investigate now the ergodicity of it. Let \mathcal{F}' be a \mathbf{P} -completion of \mathcal{F} . A set $B \in \mathcal{F}'$ is called (θ_t) -invariant if $\theta_t^{-1}B = B$ \mathbf{P} -a.s. for all $t > 0$. Then the class \mathcal{I} of all (θ_t) -invariant sets is a sub σ -algebra of \mathcal{F}' . The measure \mathbf{P} is ergodic if the invariant σ -algebra is trivial a.s., i. e. for any $B \in \mathcal{I}$, $G_\xi(B) = 0$ μ -a.e. or $G_\xi(B) = 1$ μ -a.e. It was noticed that $\mathbf{P} = \int G_\xi \mu'(d\xi)$. Let \mathcal{B}' be a μ -completion of the Borel σ -algebra on \mathcal{E}' . A set $C \in \mathcal{B}'$ is called (T_t) -invariant if $G_\xi(X_t \in C) = 1_C(\xi)$ μ -a.e. for all $t > 0$. The set of all (T_t) -invariant sets \mathcal{A} is a σ -sub-algebra of \mathcal{B}' . We say \mathcal{A} is trivial a.s. if for any $C \in \mathcal{A}$ $1_C(\xi) = 1$ μ -a.e. or $1_C(\xi) = 0$ μ -a.e.

REMARK 4.8. We notice here that $\mathbf{P} \in \mathcal{S}'(\mu')$ is ergodic if and only if \mathcal{A} is trivial a.s.

Proof:

a) Assume that \mathbf{P} is ergodic and let $C \in \mathcal{A}$. Then $1_C(X_t)$ is a G_ξ -martingale for μ -almost all ξ . Indeed, for each measurable $H \subset \mathcal{E}'$

$$\begin{aligned}\mathbf{P}1_H(X_0)\mathbf{P}(1_C(X_{t+s})|\mathcal{F}_t) &= \int 1_H(\xi)G_\xi(1_C(X_{t+s})|\mathcal{F}_t)\mu' d\xi = \\ &= \mathbf{P}1_H(X_0)G_{X_t}(1_C(X_s)) = \int 1_H(\xi)G_\xi(G_{X_t}(1_C(X_s)))\mu' d\xi = \\ &= \mathbf{P}1_H(X_0)1_C(X_t) = \int 1_H(\xi)G_\xi(1_C(X_t))\mu' d\xi.\end{aligned}$$

Since H is an arbitrary measurable set, it follows that μ -a.s. $1_C(X_t)$ is a G_ξ -martingale. Set $Z = \limsup_{t \rightarrow \infty} 1_C(X_t)$. It is obviously (θ_t) -invariant. By the convergence theorem of martingales μ -a.s. $1_C(\xi) = G_\xi(Z)$. On the other hand, $G_\xi(Z) = 0$ μ -a.s. or $G_\xi(Z) = 1$ μ -a.s. by ergodicity of \mathbf{P} .

b) Assume that \mathcal{A} is trivial a.s. Let $B \in \mathcal{I}$, $f(\xi) = G_\xi(B)$. Then f is (T_t) -invariant. Indeed, for all measurable $H \subset \mathcal{E}'$

$$\begin{aligned}\mathbf{P}(1_H(X_0)B) &= \mathbf{P}(1_H(X_0)\theta_t^{-1}B) = \mathbf{P}(1_H(X_0)\mathbf{P}(\theta_t^{-1}B|\mathcal{F}_t)) = \\ &= \mathbf{P}(1_H(X_0)G_{X_t}(B))\end{aligned}$$

Thus μ -a.s. $G_\xi(B) = G_\xi(G_{X_t}(B))$ for all $t > 0$. So f is μ -a.s. a constant function. On the other hand μ -a.s. $f(X_t)$ is a G_ξ -martingale, because by the Markov property for every $s, t \geq 0$

$$\begin{aligned}\mathbf{P}(1_H(X_0)f(X_{t+s})1_F) &= \mathbf{P}(1_H(X_0)1_F\mathbf{P}[f(X_{t+s})|\mathcal{F}_t]) = \\ &= \mathbf{P}(1_H(X_0)G_{X_t}f(X_s)1_F) = \mathbf{P}(1_H(X_0)f(X_t)1_F)\end{aligned}$$

By martingale convergence theorem $\lim_{t \rightarrow \infty} f(X_t) = 1_B$ G_ξ -a.s. Therefore we have $f = 1$ μ -a.s. or $f = 0$ μ -a.s., proving that \mathbf{P} is ergodic. ■

Now we prove the ergodicity of \mathbf{P} .

PROPOSITION 4.3. *Let $2\gamma < \alpha_0 - 1$. Then $\mathbf{P} \in \mathcal{S}'(\mu')$ is ergodic.*

Proof:

Let $A \in \mathcal{A}$ be an (T_t) -invariant set

$$G_\xi(X_t \in A) = 1_A(\xi), \mu\text{-a.s. for all } t > 0,$$

such that $\mu(A) > 0$. According to Remark 4.8 it is enough to show that $\mu(A) = 1$. Let \mathfrak{A} be the set defined in Remark 4.8. By Proposition 4.1 $\mu(\mathfrak{A}) = 1$. Thus for all $t > 0$

$$G_\xi(X_t \in A) = 1, \text{ for } \mu\text{-almost all } \xi \in A \cap \mathfrak{A}$$

and $\mu(A \cap \mathfrak{A}) = \mu(A) > 0$. So there exists $\xi_0 \in A \cap \mathfrak{A}$ and $t > 0$ such that $G_{\xi_0}(X_t \in A) = 1$. By Remark 4.8 and our results we have the equivalence of measures $G_\xi(X_t \in \cdot) \sim \mathbf{P}_\xi^0(X_t \in \cdot)$, where $\mathbf{P}_\xi^0 \in \mathcal{S}(\xi)$, $\xi \in \mathfrak{A}$. On the other hand by the absolute continuity of Gaussian measures (see [6]) we have the equivalence of $\mathbf{P}_\xi^0(X_t \in \cdot) \sim \mu$, for all $\xi \in \mathfrak{A}$. Thus for all $\xi \in \mathfrak{A}$, $G_\xi(X_t \in \cdot) \sim G_{\xi_0}(X_t \in \cdot)$. Since $G_{\xi_0}(X_t \in A) = 1$ we have $G_\xi(X_t \in A) = 1$ for all $\xi \in \mathfrak{A}$, i.e. $G_\xi(X_t \in A) = 1$ μ -a.s. So the measure \mathbf{P} is ergodic by Remark 4.8. ■

We proved in Proposition 4.7 that if $\mathbf{P} \in \mathcal{S}'(\mu')$ then $\mathbf{P} = \int G_\xi \mu'(d\xi)$. Let \mathfrak{A} be the set defined in Remark 4.8. So By Proposition 4.1 $\mu(\mathfrak{A}) = \mu'(\mathfrak{A}) = 1$. From the ergodicity of \mathbf{P} we derive the following property.

PROPOSITION 4.4. *For every pair of probability measures π_1, π_2 on \mathfrak{A} , we have*

$$\lim_{t \rightarrow \infty} \|(\pi_1 - \pi_2)T_t\| = 0,$$

where $\pi_1 T_t$ is a measure $\int G_\xi(X_t \in \cdot) \pi_1(d\xi)$ and $\|\cdot\|$ is the norm of the total variation. In particular, if $\pi_2 = \mu'$

$$\lim_{t \rightarrow \infty} \|\pi_1 T_t - \mu'\| = 0$$

Proof:

Set $\mathcal{G}_t = \theta_t^{-1}\mathcal{F}$. Then for any probability ν on \mathfrak{A} ,

$$(4.29) \quad \lim_{t \rightarrow \infty} \sup_{A \in \mathcal{T}_t} |\mathbf{P}_\nu(A \cap B) - \mathbf{P}_\nu(A)\mathbf{P}_\nu(B)| = 0,$$

where $\mathbf{P}_\nu = \int G_\xi \nu(d\xi)$. Indeed,

$$\begin{aligned} \sup_{A \in \mathcal{T}_t} |\mathbf{P}_\nu(A \cap B) - \mathbf{P}_\nu(A)\mathbf{P}_\nu(B)| &= \sup_{A \in \mathcal{T}_t} \left| \int_A (\mathbf{P}_\nu(B | \mathcal{T}_t) - \mathbf{P}_\nu(B)) d\mathbf{P}_\nu \right| \leq \\ &\leq \int_\Omega |\mathbf{P}_\nu(B | \mathcal{T}_t) - \mathbf{P}_\nu(B)| d\mathbf{P}_\nu. \end{aligned}$$

Since $\cap_t \mathcal{G}_t = \mathcal{I}$, $\mathbf{P}_\nu(B | \mathcal{G}_t)$ converges to $\mathbf{P}_\nu(B | \mathcal{I}) = \mathbf{P}_\nu(B)$. Therefore the last member in the above inequality converges to zero. Let $\pi_1 - \pi_2 = \pi^+ - \pi^-$ be the Jordan-Hahn decomposition of the measure $\pi_1 - \pi_2$ and let S^+, S^- be the associated sets of the decomposition. Define $\pi = \pi^+ + \pi^-$. By (4.29)

$$\lim_{t \rightarrow \infty} \sup_{\Gamma} |\mathbf{P}_{\pi^+}(X_t \in \Gamma) - \mathbf{P}_\pi(X_0 \in S^+) \mathbf{P}_\pi(X_t \in \Gamma)| = 0,$$

$$\lim_{t \rightarrow \infty} \sup_{\Gamma} |\mathbf{P}_{\pi^-}(X_t \in \Gamma) - \mathbf{P}_\pi(X_0 \in S^-) \mathbf{P}_\pi(X_t \in \Gamma)| = 0.$$

Since $(\pi_1 - \pi_2)(S^+ \cup S^-) = 0$,

$$\lim_{t \rightarrow \infty} \sup_{\Gamma} |\mathbf{P}_{\pi^+}(X_t \in \Gamma) - \mathbf{P}_{\pi^-}(X_t \in \Gamma)| = 0,$$

and the statement follows. ■

4.2.2. Singularity of measures. We prove here that a stationary solution \mathbf{P} of (4.14) and a stationary solution \mathbf{P}_0 of (4.13) are singular in the case $\alpha_0 = \varepsilon - 2\gamma = 1$.

PROPOSITION 4.5. *Let $\varepsilon - 2\gamma = 1$. Then \mathbf{P} and \mathbf{P}_0 are singular on \mathcal{F}_t for each $t > 0$.*

Proof:

a) First of all we prove that for each $t > 0$

$$\int_0^t \int_U (X_s(y)^{2m-1})^2 dy ds = +\infty \text{ } \mathbf{P}\text{-a.s.}$$

Let $B_t = (I - \Delta)^{1/2} X_t$. Fix a measurable subset $\tilde{U} \subset U$ such that $|\tilde{U}| > 0$ and $\text{dist}(\tilde{U}, \partial U) = \tau > 0$. Let $\varepsilon < 1/4\tau < 1$. Define $\bar{V}_t^\varepsilon = \int_0^t \int_{\tilde{U}} (f : X_s(y)^{2m-1} :$

$w_\varepsilon(x - y) dy)^2 dx ds$. Then according to (4.8) and Lemma 4.2

$$(4.30) \quad \bar{V}_t^\varepsilon = \int_0^t \int_{\tilde{U}} \sum_{l=0}^{n-1} l! \binom{n}{l}^2 \int H_{\varepsilon,l}(x, x_1, \dots, x_{n-l}, y_1, \dots, y_{n-l}) \\ B_s(dx_1) \dots B_s(dy_{n-l}) dx ds \\ + t \int_{\tilde{U}} \left(\int_{U^2} G_1(z, z')^n w_\varepsilon(x - z) w_\varepsilon(x - z') dz dz' dx \right)$$

where $n = 2m - 1$. Let:

$$A_s^{\varepsilon,l}(x) = \int H_{\varepsilon,l}(x, x_1, \dots, x_{n-l}, y_1, \dots, y_{n-l}) B_s(dx_1) \dots B_s(dy_{n-l})$$

and fix $p_0 \in (1, (2m-1)/(2m-2))$. Then by Young inequality

$$\mathbf{P}|A_s^{\varepsilon,l}(x)| = \mathbf{P}_0 \phi(X_s) |A_s^{\varepsilon,l}(x)| \leq \mathbf{P}_0 [(1/p'_0) \phi(X_s)]^{p'_0} + \\ +(1/p_0) |A_s^{\varepsilon,l}(x)|^{p_0},$$

where $\phi(X) = d\mu'/d\mu, 1/p_0 + 1/p'_0 = 1$. So, by Lemma 4.4,

$$\mathbf{P}|A_s^{\varepsilon,l}(x)| \leq C_1 + C_2 \mathbf{P}_0(|A_s^{\varepsilon,l}(x)|^{p_0}) \leq C_1 + C_2 (\mathbf{P}_0 |A_s^{\varepsilon,l}(x)|^2)^{1/\tilde{p}},$$

where $\tilde{p} p_0 = 2$. Using Hölder inequality we obtain

$$\int_{\tilde{U}} \mathbf{P}|A_s^{\varepsilon,l}(x)| dx \leq C'_1 + C'_2 (\int_{\tilde{D}} \mathbf{P}_0(|A_s^{\varepsilon,l}(x)|^2) dx)^{1/\tilde{p}}$$

From here and Lemma 4.2 it follows

$$\int_{\tilde{U}} \mathbf{P}|A_s^{\varepsilon,l}(x)| dx \leq \\ \leq C'_1 + C_2 |\log \varepsilon|^{2l/\tilde{p}} = C'_1 + C_2 |\log \varepsilon|^{lp_0},$$

and $lp_0 < 2m - 1$. We have by (4.2) that for some $c'_1 > 0$

$$\bar{V}_t^\varepsilon \geq c'_1 t |\log \varepsilon|^{2m-1} - \\ - \int_0^t \int_{\tilde{U}} \sum_{l=0}^{n-1} l! \binom{n}{l}^2 |A_s^{\varepsilon,l}(x)| dx ds,$$

i.e.

$$\bar{V}_t^\varepsilon |\log \varepsilon|^{-(2m-1)} \geq c'_1 t - K_t^\varepsilon,$$

where $K_t^\varepsilon = |\log \varepsilon|^{-(2m-1)} \int_0^t \int_{\tilde{U}} \sum_{l=0}^{n-1} l! \binom{n}{l}^2 |A_s^{\varepsilon,l}(x)| dx ds, n = 2m - 1$. By the above estimates

$$\mathbf{P} K_t^\varepsilon \rightarrow 0, \text{ as } \varepsilon \rightarrow 0.$$

Thus for some sequence $\varepsilon_n \rightarrow 0$

$$\lim_n \bar{V}_t^{\varepsilon_n} = +\infty \text{ P-a.s.,}$$

So, $\int_0^t \int_U (X_s(y)^{2m-1})^2 dy ds = +\infty$ P-a.s.

b) Consider the projections $X_t^N = \Pi^N X_t = \sum_{k=1}^N \langle X_t, e_k \rangle e_k$ of the canonical process X_t . Let $\mathcal{F}_t^N = \sigma(X_s^N, s \leq t), \mathbb{F}^N = (\mathcal{F}_{t+})_{t \geq 0}$. Define

$$h_t^k = -\lambda m \langle X_t^{2m-1}, e_k \rangle, K_t^N = \sum_{k=1}^N h_t^k e_k,$$

$$h_t^{N,k} = \mathbf{P}[h_t^k \mid \mathcal{F}_t^N], \bar{K}_t^N = \sum_{k=1}^N h_t^{N,k} e_k.$$

Fix $T > 0$. Let $\mathbf{P}_0^N = \mathbf{P}_0|_{\mathcal{F}_T^N}$, $\mathbf{P}^N = \mathbf{P}|_{\mathcal{F}_T^N}$. Then \mathbf{P}_0^N is uniquely determined by the fact that \mathbf{P}_0 -a.s. for all $t \leq T$

$$(4.31) \quad \begin{cases} dX_t^N = -1/2(I - \Delta)X_t^N dt + dW_t^N, \\ X_0^N(\mathbf{P}_0) = \mu^N = N(0, \Pi^N(I - \Delta)^{-1}\Pi^N), \end{cases}$$

where $W_t^N = \sum_{k=1}^N B_t^k e_k$, and B_t^k are independent standard Wiener processes. Similarly, \mathbf{P}^N -a.s. for all $t \leq T$

$$(4.32) \quad \begin{cases} dX_t^N = -1/2(I - \Delta)X_t^N dt - \tilde{K}_t^N dt + d\tilde{W}_t^N, \\ X_0^N(\mathbf{P}) = \mu^N = \mathbf{P}[\phi(X) | \mathcal{F}_0^N] \mu. \end{cases}$$

where $\tilde{W}_t^N = \sum_{k=1}^N \tilde{B}_t^k e_k$, and \tilde{B}_t^k are independent standard Wiener processes. Also,

$$\mathbf{P}^N\left(\int_0^T \sum_{k=1}^N (h_s^{N,k})^2 ds < \infty\right) = 1,$$

and by Theorem 3.2 $\mathbf{P}^N \prec \mathbf{P}_0^N$. Moreover, $\mathbf{P}^N = \rho_T \mathbf{P}_0^N$ on \mathcal{F}_T^N , where

$$\rho_T = \rho_0 \exp\left\{-\int_0^T \sum_{k=1}^N h_s^{N,k} dB_s^k - 1/2 \int_0^T \sum_{k=1}^N (h_s^{N,k})^2 ds\right\}.$$

and $\rho_0 = (\mathbf{P}(\phi(X)|\mathcal{F}_0^N))^{-1}$. According to Remark 3.2, $\mathbf{P}_0|_{\mathcal{F}_T}$ and $\mathbf{P}|_{\mathcal{F}_T}$ are singular if (\mathbf{P}_0^N) and (\mathbf{P}^N) are entirely separated. For each $M > 0$, we have

$$\liminf_N \int_0^T \sum_{k=1}^N (h_s^{N,k})^2 ds \geq \int_0^T \sum_{k=1}^M \lim_N (h_s^{N,k})^2 ds.$$

But

$$\lim_N h_s^{N,k} = \lim_N \mathbf{P}[h_s^k | \mathcal{F}_s^N] = h_s^k.$$

Since M is arbitrary

$$\liminf_N \int_0^T \sum_{k=1}^N (h_s^{N,k})^2 ds \geq c \int_0^T \left| : X_s^{2m-1} : \right|_0^2 ds = \infty$$

by the first part of the proof. Thus (\mathbf{P}_0^N) and (\mathbf{P}^N) are entirely separated by Theorem 3.6, and we are done. ■

5. Appendix

5.1. Dirichlet Laplacian and its Green Functions. In this section we define the Hilbert scale generated by Laplacian with zero boundary conditions and study some properties of the associated Green function.

Let $U \subset \mathbf{R}^d$ be a bounded smooth domain. We remind some facts about function spaces $H^s(U)$. By definition it is the space of generalized functions $f \in \mathcal{D}'(U)$ which are restrictions of $g \in H^s(\mathbf{R}^d)$ (we write $f = Rg$ and call R a restriction operator). The corresponding norm is defined by

$$|f|_{H^s} = |f|_{H^s(U)} = \inf_{Rg=f} |g|_{H^s(\mathbf{R}^d)},$$

where $|g|_{H^s(\mathbf{R}^2)} = |\mathcal{F}^{-1}(1 + |\xi|^s)\mathcal{F}g|_{L_2(\mathbf{R}^2)}$, \mathcal{F} is a Fourier transform and \mathcal{F}^{-1} its inverse. For $s \geq 0$ we denote $H_0^s = H_0^s(U)$ the completion of $C_0^\infty(U)$ with respect to the $H^s(U)$ -norm. Let $H_{00}^{1/2}(U) = \{u \in H_0^{1/2}(U) : |u|_{H_{00}^{1/2}}^2 = |\rho^{-1/2}u|_0^2 + |u|_{H^{1/2}}^2 < \infty\}$, where $\rho(x) = \text{dist}(x, \partial U)$. Consider the Dirichlet problem

$$(5.1) \quad \begin{cases} (-\Delta)u = f & \text{in } U \\ u|_{\partial U} = 0. \end{cases}$$

Let (e_k) be the basis in $L_2(U)$ of the eigenfunctions of the Dirichlet Laplacian (5.1), i.e. $(-\Delta)e_k = \lambda_k e_k$, $e_k|_{\partial U} = 0$, $k = 1, 2, \dots$. It is a standard fact that $e_k \in H_0^1(U) \cap C^\infty(\bar{U})$ and

$$\lambda_k \sim k^{2/d}$$

(see e.g. [11]). Introduce the spaces $\mathcal{H}^s(U) = \{f \in L_2(U) : f = \sum_k f_k e_k\}$, and $|f|_s = (\sum_k (1 + \lambda_k)^s f_k^2)^{1/2} < \infty$, $s \geq 0$. According to [27] $\mathcal{E} = \cap_s \mathcal{H}^s$ (with the sets $V_{s,\varepsilon} = \{f : |f|_s \leq \varepsilon\}$, $s \geq 0$, $\varepsilon > 0$, as zero neighbourhood basis) is a countably Hilbert nuclear space, its dual $\mathcal{E}' = \cup_{s \geq 0} (\mathcal{H}^s)'$. We have here Gelfand's triple $\mathcal{E} \subset L^2(U) \subset \mathcal{E}'$.

Let introduce also an intermediate interpolation space $\tilde{H}^{3/2}(U) = [H^2(U) \cap H_0^1(U), H_0^1(U)]_{1/2}$. Obviously $\tilde{H}^{3/2}(U) \subset H^{3/2}(U)$ and the embedding is continuous, i.e. the norm of $\tilde{H}^{3/2}(U)$ is stronger. We point out the following relations between $H^s(U)$ and $\mathcal{H}^s(U)$.

LEMMA 5.1. *Let $s = 2k + \gamma$, $k \in \mathbf{N} = \{0, 1, 2, \dots\}$, $\gamma \in [0, 2]$. Then*

$\mathcal{H}^s(U)$ coincides with the set of all $u \in H^s(D)$ such that $(I - \Delta)^l u \in H_0^1(U)$, if $l = 0, 1, \dots, k - 1$, and

$$(I - \Delta)^k u \in \begin{cases} H^\gamma(U) \cap H_0^1(U) & , \text{if } 1 \leq \gamma \leq 2, \gamma \neq 3/2, \\ \tilde{H}^{3/2}(U) & , \text{if } \gamma = 3/2, \\ H_0^\gamma(U) & , \text{if } 0 \leq \gamma \leq 1, \gamma \neq 1/2, \\ H_{00}^{1/2}(U) & , \text{if } \gamma = 1/2, \end{cases}$$

Moreover, $\mathcal{H}^s(U)$ and $H^s(U)$ norms are equivalent if $s \neq 2k + 1/2, 2k + 3/2$. If $s = 2k + 1/2$, the $|u|_s$ norm is equivalent to $|u|_{H^s(U)} + |(I - \Delta)^k u|_{H_{00}^{1/2}(U)}$ norm. If $s = 2k + 3/2$, the $|u|_s$ norm is equivalent to $|u|_{H^s(U)} + |(I - \Delta)^k u|_{\tilde{H}^{3/2}(U)}$.

Proof:

a) Let $k = 0, \gamma \in [0, 1]$. By the choice of (e_k) $H_0^1(U) = \mathcal{H}^1(U)$, $H_0^0(U) = L_2(U) = \mathcal{H}^0(U)$ and the corresponding norms are equivalent. Then by interpolation and Theorem 11.7 of Chapter 1 in [4] $H_0^\gamma(U) = \mathcal{H}^\gamma(U)$, if $0 \leq \gamma \leq 1, \gamma \neq 1/2$, $H_{00}^{1/2}(U) = \mathcal{H}^{1/2}(U)$ and the corresponding norms are equivalent. Let $\gamma \in [1, 2], \gamma \neq 3/2$, $u \in H^\gamma(U) \cap H_0^1(U)$. Then by Theorem 8.3 of Chapter 2 in [4] the norms $|u|_{H^\gamma}$ and $|(I - \Delta)u|_{H^{\gamma-2}}$ are equivalent. If $\gamma \neq 3/2$, $H^{\gamma-2}(U) = (H_0^{\gamma-1}(U))' = (\mathcal{H}^{\gamma-1})' = \mathcal{H}^{\gamma-2}$ and the norms are equivalent, i.e. $|(I - \Delta)u|_{H^{\gamma-2}} \sim |(I - \Delta)u|_{\gamma-2}$. On the other hand, let $(I - \Delta)u = \sum_k f_k e_k$, $f^N = \sum_{k=1}^N f_k e_k$, $u^N = \sum_{k=1}^N (1 + \lambda_k)^{-1} f_k e_k$. Then $(I - \Delta)u^N = f^N$ and by the same theorem 8.3 in [4] $|u^N|_{H^\gamma} \sim |(I - \Delta)u^N|_{H^{\gamma-2}} \sim |(I - \Delta)u^N|_{\gamma-2} = |f^N|_{\gamma-2} = |u^N|_\gamma$. Also, $|f^N - f|_{\gamma-2} \rightarrow 0$, as $N \rightarrow \infty$. So, by the theorem 8.3 of Chapter 2 in [4] $u^N \rightarrow u$ in $H^\gamma(U)$ and equivalently in $\mathcal{H}^\gamma(U)$. Thus $H^s(U) \cap H_0^1(U) = \mathcal{H}^s(U)$ and

the norms are equivalent if $1 \leq s \leq 2$, $s \neq 3/2$. If $\gamma = 3/2$, the statement follows by interpolation between the pairs $(H^2(U) \cap H_0^1(U), H_0^1(U))$ and $(\mathcal{H}^2(U), \mathcal{H}^1(U))$.

b) Assume $k \geq 1$, $s = 2k + \gamma$. By Theorem 5.1 of Chapter 2 in [4] we have the equivalence of the following norms

$$|u|_{H^s(U)} \sim |(I - \Delta)u|_{H^{s-2k}(U)} \sim \dots \sim |(I - \Delta)^k u|_{H^{s-2k}(U)}.$$

Because of our assumptions ant the part a) of the proof the last norm is equivalent to $|(I - \Delta)^k u|_{s-2k}$, if $s \neq 2k + 1/2, 2k + 3/2$. If $s = 2k + 1/2$, then $|u|_{H^s(U)} + |(I - \Delta)^k u|_{H_0^{1/2}}$ is equivalent to $|(I - \Delta)^k u|_{s-2k} + |(I - \Delta)^k u|_{H_0^{1/2}} \sim |(I - \Delta)^k u|_{1/2}$ by part a). Similarly, if $s = 2k + 3/2$, then $|u|_{H^s(U)} + |(I - \Delta)^k u|_{\tilde{H}^{3/2}(U)}$ is equivalent to $|(I - \Delta)^k u|_{H^{3/2}(U)} + |(I - \Delta)^k u|_{\tilde{H}^{3/2}} \sim |(I - \Delta)^k u|_{3/2}$ by part a). Let $(I - \Delta)^k u = \sum_j b_j e_j$, $u^N = \sum_{j=1}^N (1 + \lambda_j)^{-k} b_j e_j$. Then $\sum_j (1 + \lambda_j)^{s-2k} b_j^2 < \infty$, $|(I - \Delta)^k u^N|_{s-2k} = |u^N|_s$. Passing to the limit as $N \rightarrow \infty$ we obtain our statement by Theorem 5.1 of Chapter 2 in [4]. ■

We have the following obvious consequence of the previous statement.

REMARK 5.1. For each $s \neq 2k+1/2, 2k+3/2$ $H_0^s(U) \subset \mathcal{H}^s(U)$. Also, $C_0^\infty(U) \subset \cap_s \mathcal{H}^s(U) = \mathcal{E}$, and all $e_k \in \mathcal{E}$.

For $s < 0$ we define $\mathcal{H}^s(U)$ as a completion of $L_2(U)$ or \mathcal{E} in the norm

$$|f|_s = \left(\sum_k (1 + \lambda_k)^s f_k^2 \right)^{1/2},$$

where $f = \sum_k f_k e_k$.

REMARK 5.2. One can notice here that for $s \geq 0$ $\mathcal{H}^s(U)' = \mathcal{H}^{-s}(U) \subset \mathcal{E}' \subset \mathcal{D}'(U)$,

$$\langle f, \varphi \rangle_{\mathcal{E}', \mathcal{E}} = \sum_k f_k \varphi_k,$$

where $f = \sum_k f_k e_k \in \mathcal{E}' = \cup_s \mathcal{H}^{-s}(U)$, $\varphi = \sum_k \varphi_k e_k \in \mathcal{E}$. Also $\mathcal{H}^{-1}(U) = H^{-1}(U) = H_0^1(U)'$.

Now we define fractional powers of the $I - \Delta$ by

$$(I - \Delta)^s f = \sum_k (1 + \lambda_k)^s f_k e_k,$$

where $f = \sum_k f_k e_k$. Consider the following equation in U

$$(5.2) \quad (I - \Delta)^\alpha u = f.$$

If $f = \sum_k f_k e_k \in \mathcal{H}^s(U)$, then $u = (I - \Delta)^{-\alpha} f = \sum_k (1 + \lambda_k)^{-\alpha} f_k e_k \in \mathcal{H}^{s+2\alpha}(U)$ is a solution of (5.2).

REMARK 5.3. We notice that for $\alpha > 0$

$$(5.3) \quad (I - \Delta)^{-\alpha} f = c(\alpha) \int_0^\infty t^\alpha T_t f e^{-t} \frac{dt}{t},$$

where $v = T_t f$ solves the equation

$$(5.4) \quad \begin{cases} \partial_t v = \Delta v, & \text{in } U, \\ v(0, \cdot) = f, \\ v(t, \cdot) |_{\partial U} = 0, \end{cases}$$

and $c(\alpha) = (\int_0^\infty t^\alpha e^{-t} \frac{dt}{t})^{-1}$. If $p(t, x, y)$ is the fundamental solution of (5.4), then

$$(I - \Delta)^{-\alpha} f = c(\alpha) \int_0^\infty t^\alpha T_t f e^{-t} \frac{dt}{t} = \int_U G_\alpha(x, y) f(y) dy,$$

where

$$(5.5) \quad G_\alpha(x, y) = c(\alpha) \int_0^\infty t^\alpha e^{-t} p(t, x, y) \frac{dt}{t}$$

is called a Green function of the problem (5.2).

REMARK 5.4. We have by definitions that for each $\alpha > 0, \beta > 0$

$$(5.6) \quad G_{\alpha+\beta}(x, y) = \int_U G_\alpha(x, z) G_\beta(z, y) dz,$$

Now we discuss some regularity properties of the Green function.

LEMMA 5.2. Let $\alpha > 0$. Then

1) for $s = 0, 1, 2, \dots$ there is a constant $c > 0$ such that

$$|\partial_x^s G_\alpha(x, y)| \leq \begin{cases} c|x - y|^{2\alpha-d-s}, & \text{if } s > 2\alpha - d, \\ c_1 + c_2 |\log|x - y||, & \text{if } s = 2\alpha - d, \\ c, & \text{if } s < 2\alpha - d; \end{cases}$$

2) for $s = 0, 1, 2, \dots$ and $\beta \in (0, 1)$ there exist a constant c such that

$$|\partial_x^s G_\alpha(x, y) - \partial_x^s G_\alpha(x', y)| |x - x'|^{-\beta} \leq \begin{cases} c|x'' - y|^{2\alpha-d-s-\beta}, & \text{if } s + \beta > 2\alpha - d, \\ c_1 + c_2 |\log|x'' - y||, & \text{if } s + \beta = 2\alpha - d, \\ c, & \text{if } s + \beta < 2\alpha - d; \end{cases}$$

where $|x'' - y| = \min\{|x - y|, |x' - y|\}$.

Proof:

Fix $\alpha > 0$. Then by Remark 5.3

$$G_\alpha(x, y) = c(\alpha) \int_0^\infty t^\alpha e^{-t} p(t, x, y) \frac{dt}{t},$$

where $p(t, x, y)$ is a fundamental solution of (2.1). According to [15] for each $s = 1, 2, \dots$

$$(5.7) \quad |\partial_x^s p(t, x, y)| \leq c \exp\{-c'|x - y|^2/t\} / t^{(d+s)/2},$$

$$|\partial_x^s p(t, x, y) - \partial_x^s p(t, x', y)| \leq C|x - x'|^\beta t^{-(d+s+\beta)/2} \exp\{-c(|x - y| \wedge |x' - y|)^2/t\}.$$

Thus for each $s = 1, 2, \dots, \beta \in (0, 1)$ we have by (5.7)

$$(5.8) \quad \begin{aligned} |\partial_x^s G_\alpha(x, y)| &= \left| \int_0^\infty t^\alpha e^{-t} \partial_x^s p(t, x, y) \frac{dt}{t} \right| \leq \\ &\leq C \int_0^\infty t^\alpha e^{-t} t^{-(d+s)/2} \exp\{-|x - y|^2/t\} \frac{dt}{t}, \\ |\partial_x^s G_\alpha(x, y) - \partial_x^s G_\alpha(x', y)| |x - x'|^{-\beta} &\leq \\ Ct^{-(d+s+\beta)/2} \exp\{-c(|x - y| \wedge |x' - y|)^2/t\}. \end{aligned}$$

Now the statement follows. ■

The following estimates for the mixed derivatives hold.

COROLLARY 5.1. Let $\alpha > 0, k \in \mathbf{N}, \beta \in (0, 1)$. Then there is a constant C such that

$$\begin{aligned} |\partial_x^k \partial_y^l G_\alpha(x, y)| &\leq C \text{ if } k, l < \alpha - 1, \\ \sup_{x, x', y, y'} |(\partial_x^k \partial_y^l G_\alpha(x, y) - \partial_x^k \partial_y^l G_\alpha(x', y)) - \\ &-(\partial_x^k \partial_y^l G_\alpha(x, y') - \partial_x^k \partial_y^l G_\alpha(x', y'))| |x - x'|^{-\beta} |y - y'|^{-\beta} &\leq C, \\ \text{if } k + \beta < \alpha - 1, l + \beta < \alpha - 1. \end{aligned}$$

Proof:

By Remark 5.4, if $k < \alpha - 1$,

$$\partial_x^k \partial_y^l G_\alpha(x, y) = \int_U \partial_x^k G_{\alpha/2}(x, z) \partial_y^l G_{\alpha/2}(z, y) dz.$$

Since $\partial_x^k G_\alpha(x, y) = \partial_y^k G_\alpha(y, x)$ the statement follows by Lemma 5.2 and Hölder inequality. ■

Denote $C^{j+\beta}(U \times U)$ the set of all j -times continuously differentiable functions $g(x, y)$ on $U \times U$ with finite norm

$$\begin{aligned} \|g\|_{j, \beta} &= \sup_{x, y} \sum_{k, l \leq j} |\partial_x^k \partial_y^l g(x, y)| + \sum_{k+l=j} [\partial_x^k \partial_y^l g]_\beta, \text{ where } [v]_\beta = \\ &= \sup_{x, x', y, y'} |v(x', y) - v(x, y) - v(x', y') + v(x, y')| |x - x'|^{-\beta} |y - y'|^{-\beta}. \end{aligned}$$

So G_α belongs to $C^{j, \beta}(U \times U)$, where $j + \beta < \alpha - 1, \beta \in (0, 1), j = [\alpha - 1]$, if $\alpha - 1 \notin \mathbf{N}$, $j = \alpha - 2$, if $\alpha - 1 \in \mathbf{N}$. As the following statement shows the space $C^{j+\beta}(U \times U)$ is an algebra.

LEMMA 5.3. Let $f, g \in C^{j+\beta}(U \times U)$. Then $h = fg \in C^{j+\beta}(U \times U)$.

Proof:

Let $v_1 = \partial_x^k \partial_y^{k'} f(x, y), v_2 = \partial_x^l \partial_y^{l'} g(x, y), k + k' + l + l' = j$. It is enough to prove that $[v_1 v_2]_\beta < \infty$. In fact, $[v_1 v_2]_\beta$ is no greater than:

$$\begin{aligned} &\sup_{x, y, y'} |v_1(x, y) - v_1(x, y')| |y - y'|^{-\beta} \sup_{y, x, x'} |v_2(x, y) - v_2(x', y)| |x - x'|^{-\beta} + \\ &+[v_1]_\beta \sup_{x, y} |v_2(x, y)| + [v_2]_\beta \sup_{x, y} |v_1(x, y)| + \sup_{x, x', y'} \\ &|v_1(x, y') - v_1(x', y')| |x - x'|^{-\beta} \sup_{x, y, y'} |v_2(x, y) - v_2(x, y')| |y - y'|^{-\beta}, \end{aligned}$$

and the statement follows. ■

We shall consider some smooth approximations of the Green functions G_α in the case $d = 2$. By Theorem 3.3.4 in [5] for each natural number $N = 0, 1, 2, \dots$ there exists an extension operator $S = S^N$ mapping continuously $H^s(U)$ to $H^s(\mathbf{R}^2)$ for all $|s| < N$ and such that $RS = I$, where R is a restriction operator (see formula 3.3.4/6 in [5]). Moreover, there exists $c > 0$ such that $\text{supp } Sf \subset \{x : |x| < c\}$ for each $f \in H^s(U), |s| < N$. Fix $N > \max\{2(\alpha_0 - 1), 2\}$ and define for $\alpha \geq 1$

$$\bar{G}_\alpha(x, y) = \bar{G}_{\alpha, N}(x, y) = S_x^N S_y^N G_\alpha(x, y).$$

Let $w \in C_0^\infty(\mathbf{R}^2), w \geq 0, \int w dx = 1$. Define $w_\varepsilon(x) = \varepsilon^{-2} w(x/\varepsilon)$ and

$$\bar{G}_\alpha^{\varepsilon, \varepsilon'}(x, y) = \bar{G}_{\alpha, N}^{\varepsilon, \varepsilon'}(x, y) = \int \int \bar{G}_{\alpha, N}(z, z') w_\varepsilon(x - z) w_{\varepsilon'}(y - z') dz dz'.$$

Now we evaluate the rate of convergence of $\bar{G}^{\varepsilon, \varepsilon'}$.

LEMMA 5.4. *Let $\alpha \geq 1$. Then for each $n \geq 1$ there exist $c, \kappa > 0$ such that*

$$\int \int |\bar{G}_{\alpha, N}^{\varepsilon, \varepsilon'}(x, y) - \bar{G}_{\alpha, N}(x, y)|^n dx dy \leq C(\varepsilon \vee \varepsilon')^\kappa.$$

Moreover,

$$\int |\bar{G}_{\alpha, N}^{\varepsilon, \varepsilon'}(x, x) - \bar{G}_{\alpha, N}(x, x)|^n dx \leq C(\varepsilon \vee \varepsilon')^\kappa, \text{ if } \alpha > 1.$$

Proof:

Let $\alpha > 1$. By Lemma 5.2 and by the properties of the extension operator S (see [5], Theorem 3.3.4) for each $0 < \beta < 2(\alpha - 1) \wedge 1$

$$|\bar{G}_\alpha(x, y) - \bar{G}_\alpha(x', y)| \leq c|x - x'|^\beta.$$

Because of the symmetry a similar inequality holds for the difference with respect to y and the statement follows immediately with any $\kappa < n[2(\alpha - 1) \wedge 1]$.

Let $\alpha = 1$. By the definition of $S = S^N$ we have for $x, y \in U$

$$|\bar{G}_1(x, y)| \leq c'_1 + c'_2 |\log|x - y||,$$

$$|\partial_x \bar{G}_1(x, y)| \leq c'|x - y|^{-1}.$$

Denote $B = \int_U \int_U |\bar{G}_1^{\varepsilon, \varepsilon'}(x, y) - \bar{G}_1(x, y)| dx dy$. Let $\delta = \varepsilon \vee \varepsilon' \in (0, 1)$ and fix $\eta \in (0, 1)$. Then we split the integral

$$B = \int_{\{|x-y| \leq \delta^\eta\} \cap U_\delta^2} + \int_{\{|x-y| > \delta^\eta\} \cap U_\delta^2} + \int_{U^2 \setminus U_\delta^2} = B_1 + B_2 + B_3,$$

where $U^2 = U \times U$, $U_\delta^2 = \{(x, y) \in U^2 : d(x, \partial U) > \delta^\eta, d(y, \partial U) > \delta^\eta\}$. We notice that by Holder's inequality

$$B_1 \leq C \int_{|x-y| \leq 3\delta^\eta} |\bar{G}_1(x, y)|^n dx dy \leq C\delta^{\kappa_1}$$

with any $\kappa_1 < \lambda\eta, \lambda < 2$. If $|x - y| > \delta^\eta$, then $|z - z'| > c'\delta^\eta$. Thus

$$\begin{aligned} B_2 &\leq \int_{|x-y| > \delta^\eta} |\bar{G}_1(x, y) - \bar{G}_1(z, z')|^n w_\varepsilon(x - z) w_{\varepsilon'}(y - z') dz dz' dx dy \leq \\ &\leq C\delta^n / \delta^{n\eta} = C\delta^{n(1-\eta)}. \end{aligned}$$

By Hölder inequality again

$$B_3 \leq C\delta^{\eta(1-\eta)}.$$

Now the statement follows. ■

COROLLARY 5.2. *Let the assumptions of Lemma 5.4 be satisfied. Then for each $n \geq 1, \theta > 0$ there exist $\kappa > 0, c > 0$ such that*

$$\int_U \int_U G_\theta(x, y) |\bar{G}_{\alpha, N}^{\varepsilon, \varepsilon'}(x, y) - \bar{G}_{\alpha, N}(x, y)|^n dx dy \leq c(\varepsilon \vee \varepsilon')^\kappa,$$

$$\int_U \int_U G_\theta(x, y) G_\alpha(x, y)^n dx dy < \infty.$$

Proof:

The statement follows obviously by Hölder inequality and Lemma 5.4. ■

For $\varphi \in C_0^\infty(\mathbf{R}^2)$, $\gamma \in (0, 1)$ introduce $|\partial|^\gamma \varphi = |\partial_x|^\gamma \varphi(x) = \mathcal{F}^{-1}[|\xi|^\gamma \mathcal{F}\varphi](x) = N(\gamma) \int [\varphi(x+y) - \varphi(x)] \frac{dy}{|y|^{2+\gamma}}$,

where \mathcal{F} denotes the Fourier transform and \mathcal{F}^{-1} its inverse.

COROLLARY 5.3. Let $\alpha > 1$, $0 < k+\gamma < \alpha-1$, $\gamma \in [0, 1)$, $k \in \mathbf{N} = \{0, 1, 2, \dots\}$. Let $K_k^{\varepsilon, \varepsilon'}(x, y) = \partial_y^k \partial_x^k (\bar{G}_\alpha^{\varepsilon, \varepsilon'}(x, y))^n$. Then

$$\sup_{\varepsilon'} \sup_x | |\partial_x|^\gamma |\partial_y|^\gamma (K_k^{\varepsilon, \varepsilon'}(x, y) - K_k^{0, \varepsilon'}(x, y))_{y=x} | \rightarrow 0, \text{ as } \varepsilon \rightarrow 0,$$

$$\sup_x | |\partial_x|^\gamma |\partial_y|^\gamma \partial_y^k \partial_x^k (\bar{G}_\alpha(x, y)^n)_{y=x} | < \infty.$$

Proof:

Fix $\kappa' > \kappa$ such that $k+\gamma < k+\kappa' < \alpha-1$. Then by Corollary 5.1 and Lemma 5.3 for each $\eta > 0$

$$\begin{aligned} & \sup_{\varepsilon'} \sup_x | |\partial_x|^\gamma |\partial_y|^\gamma (K_k^{\varepsilon, \varepsilon'}(x, y) - K_k^{0, \varepsilon'}(x, y))_{y=x} | \\ & \leq C \left\{ \int_{|z| \leq \eta} \frac{dz}{|z|^{2-(\kappa'-\kappa)}} + \sup_{x, \varepsilon'} \left| \int_{|z| > \eta} \int_{|z'| > \eta} [(K_k^{\varepsilon, \varepsilon'}(x+z, y+z') \right. \right. \\ & \quad \left. \left. - K_k^{0, \varepsilon'}(x, y+z')) - (K_k^{\varepsilon, \varepsilon'}(x+z, y) - K_k^{0, \varepsilon'}(x, y))] \frac{dz}{|z|^{2+\gamma}} \frac{dz'}{|z'|^{2+\gamma}} \right| \right\}. \end{aligned}$$

Now the statement follows by Corollary 5.1 and Lemma 5.3. ■

Bibliography

- [1] P.-A. Meyer (1993): Quantum Probability for Probabilists. Lecture Notes in Math. **1538**, Springer Verlag, New York.
- [2] P.-A. Meyer (1977): Notes sur les intégrales stochastiques I. Intégrales Hilbertiennes. *Séminaire de Probabilités XI*, Lecture Notes in Math. **581**, 446-463.
- [3] H. Kunita (1972): Stochastic flows and stochastic differential equations. Cambridge University Press, Cambridge.
- [4] J. L. Lions and E. Magenes (1972): Non-homogeneous boundary value problems. Springer Verlag, New York, N. Y.
- [5] H. Triebel (1983): Theory of functions spaces. Birkhäuser, Basel.
- [6] I. I. Gihman and A. V. Skorohod (1979): The theory of stochastic processes I. Springer Verlag, New York.
- [7] J. Glimm and A. Jaffe (1987): Quantum Physics. A functional integral point of view. Springer Verlag, New York.
- [8] M. Metivier and J. Pellaumail (1980): Stochastic Integration. Academic Press, New York.
- [9] B. Simon (1974): The $P(\phi)_2$ Euclidian (Quantum) Field Theory. Princeton Univ. Press, Princeton.
- [10] L. Schwartz (1964): Sous-espaces hilbertiens d'espaces vectoriel topologiques et noyeux associés (noyeux reproduisants). *Journal d'Analyse Math.*, **13**, 115-256.
- [11] M.A. Shubin (1987): Pseudodifferential operators and spectral theory. Springer Verlag, New York, N.Y.
- [12] J. Jacod and A. N. Shiryaev (1987): Limit theorems for stochastic processes. Springer Verlag, New York, N.Y.
- [13] J. Jacod (1979): Calcul stochastique et problèmes de martingales. Lecture Notes in Math., **714**, Springer Verlag, New York, N.Y.
- [14] S. N. Ethier and T. G. Kurtz (1985): Markov Processes. Characterization and Convergence. John Wiley & Sons, New York, N.Y.
- [15] O. A. Ladyzhenskaya, V. A. Solonnikov and Ural'tseva (1968): Linear and quasilinear equations of parabolic type. Translations of mathematical monographs, **23**, American Mathematical Society.

- [16] R. Sh. Liptser and A. N. Shirayev (1982): Statistics of Random Processes. Springer-Verlag, New York, N.Y.
- [17] G. Jona-Lasinio and P. K. Mitter (1985): On the stochastic quantization of field theory. *Comm. Math. Phys.*, **101**, 409-436.
- [18] V. S. Borkar, R. T. Chari and S. K. Mitter (1988): Stochastic quantization of field theory in finite volume and infinite volume. *J. Funct. Anal.*, **81**, 184-206.
- [19] Gatarek, D., and Goldys, B., Existence, uniqueness and ergodicity for stochastic quantization equation, Existence,uniqueness and ergodicity for the stochastic quantization equation, *Studia Math.* 119 (1996), no2, 179-193.
- [20] S. Albeverio and M. Röckner (1989): Classical Dirichlet forms on topological vector spaces, construction of an associated diffusion process. *Prob. Theory Rel. Fields*, **83**, 405-435.
- [21] Y. Hu and G. Kallianpur (1977): Exponential integrability and application to stochastic quantization. (preprint), 1-61.
- [22] I. Mitoma (1983): Tightness of probabilities on $C([0, 1]; \mathcal{S}')$ and $D([0, 1]; \mathcal{S}')$. *Annals of Probab.*, **11**, 989-999.
- [23] A. Lunardi and V. Vespi (1995): Optimal L^∞ and Schauder estimates for elliptic and parabolic operators with unbounded coefficients. (preprint), 1-30.
- [24] N. Dunford and J. T. Schwartz (1958): Linear Operators, Part I. Interscience, New York, N. Y.
- [25] M. Reed and B. Simon (1972): Methods of modern mathematical physics, I:Functional Analysis. Academic Press, New York, N. Y.
- [26] D. W. Stroock and S. R. S. Varadhan (1979): Multidimensional diffusion processes. Springer Verlag, New York, N.Y.
- [27] I. M. Gel'fand and N. Y. Vilenkin (1964): Generalized Functions, vol.4. Academic Press, New York, N.Y.
- [28] J. B. Walsh (1986): An introduction to stochastic partial differential equations. in *Ecole d'Eté de Probabilités de Saint-Flour XIV 1984*, 265-439. Lecture Notes in Math., **1180**, Springer-Verlag, New York, N. Y.
- [29] R. Mikulevicius and B. L. Rozovskii (1994): Uniqueness and absolute continuity of weak solutions for parabolic SPDE's. *Acta Applic. Math.*, **35**, 179-192.
- [30] R. Mikulevicius and B. L. Rozovskii (1998): Normalized Stochastic Integrals in Topological Vector Spaces. In: *Séminaire de Probabilités XXXIII*. Lecture Notes in Mathematics, Springer Verlag, New York, N.Y. (to appear)
- [31] S. Albeverio and M. Röckner (1989): Dirichlet forms, quantum fields and stochastic quantization. in *Stochastic analysis, path integration and dynamics*. Research Notes in Mathematics, vol. 200, 1 - 21. Editors: K.D. Elworthy, J.C. Zambrini. Harlow. Longmann.
- [32] S. Albeverio and M. Röckner (1991): Stochastic differential equations in infinite dimensions: solutions via Dirichlet forms. *Probab. Th. Rel. Fields*, **89**, 347 - 386.
- [33] M. Röckner and T. S. Zhang (1992): Uniqueness of generalized Schrödinger operators and applications. *J. Funct. Anal.*, **105**, 187 - 231.
- [34] S. Albeverio, M. Röckner and T. S. Zhang (1993): Girsanov transform for symmetric diffusions with infinite dimensional state space. *Annals of Probability*, **21**, 961 - 978.
- [35] T. S. Zhang and M. Röckner (1996): Finite dimensional approximation of diffusion processes on infinite dimensional state spaces. *Stochastics and Stochastic Reports*, **57**, 37 - 55.
- [36] S. Albeverio, Y. G. Kondratiev and M. Röckner (1996): Ergodicity for the stochastic dynamics of quasi-invariant measures with applications to Gibbs states. (SFB-343-Preprint 1996)
- [37] V. Liskevich and M. Röckner (1997): Strong uniqueness for a class of infinite dimensional Dirichlet operators and applications to stochastic quantization. (preprint)
- [38] G. Parisi and Y. S. Wu (1981): Perturbation theory without gauge fixing. *Scientia Sinica*, **24**, 383 - 496.
- [39] G. DaPrato and L. Tubaro (1996): Introduction to stochastic quantization. (preprint)
- [40] G. Jona-Lasinio and R. Seneor (1991): On a class of stochastic reaction-diffusion equations in two space dimensions. *J. Phys.*, **A 24**, 4123 - 4128.
- [41] P. K. Mitter (1986): Stochastic approach to Euclidean field theory (Stochastic quantization). in *New perspectives in Quantum Field Theory*. Editors: J. Abad, M. Asorey, A. Cruz. World Scientific, Singapore.

- [42] G. Jona-Lasinio and P. K. Mitter (1990): Large deviation estimates in the stochastic quantization of Φ_2^4 . *Commun. Math. Phys.*, **130**, 111-121.
- [43] B. Grigelionis and R. Mikulevicius (1985): Weak solutions of stochastic evolution equations. Lecture Notes in Control and Inf. Sc., **69**, 43-54. Springer Verlag, New York, N.Y.
- [44] B. Grigelionis and R. Mikulevicius, R. (1983): Stochastic evolution equations and densities of conditional distributions. Lecture Notes in Control and Inf. Sc., **49**, 49-86. Springer Verlag, New York, N. Y.
- [45] M. Metivier (1988): Stochastic Partial Differential Equations in Infinite Dimensional Spaces. Scuola Normale Superiore, Pisa.
- [46] M. Viot (1976): Solutions faibles d'équations aux dérivées partielles non linéaires. Thèse, Université Pierre et Marie Curie, Paris.
- [47] E. Pardoux (1975): Equations aux dérivées partielles stochastiques non linéaires monotones. Etude des solutions fortes de type Ito. Thèse, Université de Paris Sud, Orsay.
- [48] B. Grigelionis and R. Mikulevicius (1984): On the functional limit theorems. in *Stochastic Space - Time Models and Limit Theorems*, Editors: L. Arnold, P. Kotelenetz, 191-216. Springer Verlag New York, N.Y.
- [49] G. Kallianpur, I. Mitoma and R. Wolpert (1988): Diffusion equations in duals of nuclear spaces. (technical report No 234)
- [50] A. V. Skorohod (1965): Studies in the Theory of Random Processes. Addison-Wesley, Reading, MS.
- [51] N. V. Krylov and B. L. Rozovskii (1981): Stochastic evolution systems. *J. Soviet Math.* **16**, 1233-1276.
- [52] S. M. Kozlov (1978): Some problems in stochastic partial differential equations. Proc. of Petrovskii Seminar. **4**, 148-172.
- [53] B. L. Rozovskii (1990): Stochastic Evolution Systems. Kluwer Academic Publ., Dordrecht.
- [54] J. Jacod and J. Memin(1981): Sur un type de convergence intermediaire entre la convergence en loi et la convergence in probabilité. Lecture Notes in Math., **850**. Springer-Verlag, New York, N.Y.
- [55] A. K. Zvonkin and N. V. Krylov (1975): Strong solutions of stochastic differential equations. in *Proceedings of the School-Seminar on the Theory of Random Processes, Druskininkai, November 25-30, 1974*. Vilnius.
- [56] J. Doob (1953): Stochastic Processes. John Wiley, Ney York, N.Y.
- [57] N. Ikeda and S. Watanabe (1981): Stochastic Differential Equations and Diffusion Processes. North Holland, Amsterdam.
- [58] D. A. Dawson (1993): Measure-valued Markov Processes. in *Ecole d'Eté de Probabilités de Saint-Flour XXI 1991*, 1-262. Lect. Notes in Math., **1541**, Springer-Verlag, New York, N.Y.
- [59] H. Allouba (1997): Different Types of SPDE's in the Eyes of Girsanov's Theorem. (preprint)
- [60] I. Gyongy, I. Existence and uniqueness results for semilinear stochastic partial differential equations. *Stochastic Processes and their Applications* (to appear)
- [61] S. Roelly-Coppoletta (1986): A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics*, **17**, 43-65.
- [62] N. El Karoui and S. Méléard (1990): Martingale measures and stochastic calculus. *Prob. Theor. Relat. Fields*, **84**, 83-101.

This page intentionally left blank

Indexes

This page intentionally left blank

Notation Index

$(1 + \beta)$ -branching superprocess,	56	$\mathcal{F}_p^n(\tau)$,	191
$(1 - \Delta)^{n/2}$,	187	\mathcal{F}_t ,	190
(L, Λ) ,	190, 215	$\mathcal{H}_p^2(T)$,	189
(Ω, \mathcal{F}, P) ,	190	$\mathcal{H}_{p,0}^n(\tau)$,	191
(ϕ, u) ,	191	\mathcal{H}_p^n ,	191
(ξ, Φ, K) -superprocess,	63	$\mathcal{H}_p^n(\tau)$,	191
(u, ϕ) ,	191	\mathcal{R}_1 ,	188
$B(\mathbb{R}^d)$,	207	\mathcal{R}_{λ_0} ,	189, 219
$B^{ n +\gamma}$,	207	ϵ ,	157
$B^{ n +\gamma}(l_2)$,	207	γ ,	207
$C(\epsilon)$,	174	γ_t ,	172
C^α ,	220	\hat{h}_m ,	212
$C^{ n +\gamma}(\mathbb{R}^d)$,	207	\mathcal{P} ,	190
$C^{ n -1,1}(\mathbb{R}^d)$,	207	\mathfrak{D} ,	201
C_0^∞ ,	187	$\mathfrak{D}(l_2)$,	201
$E_\phi(\omega, \mathbf{k})$,	160	$\nu_{j,\ell}(d\omega, dk)$,	156
$E_{j,\ell}(\omega, k)$,	156	$(0, \tau]$,	191
$H_p^{1,2}(T)$,	187	$\phi(t, \mathbf{x})$,	159
H_p^n ,	187	$\phi_K(t, \mathbf{x})$,	165
$H_p^n(\mathbb{R}^d)$,	186	$\sigma_{\mathbf{k}}$,	171
$L(t, \cdot)$,	214	$\bar{\mathbf{v}}(t, \mathbf{x})$,	154, 155
L_p ,	187	ξ -superprocess,	53
Lu ,	186	ξ_t ,	165
T_t ,	187	n_ν ,	210
$W_2^n(\mathbb{R}^d)$,	185	n_b ,	210
W_p^n ,	191	n_c ,	210
$\Gamma_{j,\ell}(t, \mathbf{x})$,	156	p -tempered measure,	56
$\Lambda(t, \cdot)$,	214	u_t^k ,	186, 190
Λu ,	186	z ,	157
α^{ij} ,	196, 208	$ g _2^2$,	191
$\alpha_{\mathbf{k}}$,	171	$\ \cdot\ _{1,2,p}$,	187
\mathbb{D} ,	192	$\ \cdot\ _{\mathcal{H}_p^n(\tau)}$,	191
\mathbb{H}_p^n ,	191	$\ \cdot\ _{n,p}$,	187
$\mathbb{H}_p^n(\tau)$,	191	$\ \cdot\ _p$,	187
$\mathbb{H}_p^n(\tau, l_2)$,	191	$\ g\ _{n,p}$,	191
$\mathbb{L}_{\dots\dots}$,	191		
\mathbb{R}^d ,	186	ASEP,	113
\mathbb{S} ,	192	CS,	91
\bar{a} ,	209	HSL,	109, 112, 119
$\beta(r)$,	156	IPP,	110, 113
\mathcal{C}^α ,	220	RM,	25, 30
\mathcal{D} ,	186	RNG,	27–30, 33, 39
$\mathcal{E}(r)$,	156	RT,	27
$\mathcal{E}^{(reg)}(r)$,	157		
$\mathcal{E}_\phi(r)$,	160		
\mathcal{F}_p^n ,	191		

This page intentionally left blank

Subject Index

- Kelvin-Helmholtz instability, 21
- adjoint covariance space, 252
- auto-regressive, 166
- averaged equation, 30, 31
- Bakry-Emery conditions, 49
- branching
- catalytic, 95
 - measure-valued process
 - critical, 62
 - particle system, 45
 - interacting, 45
- Brownian
- excursion density, 57
 - rescaling, 90
- Brownian motion
- historical, 58
- Buckley-Leverett, 11, 34, 37
- Burgers equation, 110, 111, 148
- inviscid, 110, 111, 116
 - stochastic, 111, 123–125
 - with viscosity, 110, 111, 119, 120
- Burgers vector, 8
- Cahn–Allen equation, 113, 144
- stochastic, 130, 135, 138
- Cahn–Hilliard equation, 111, 146
- stochastic, 111, 135, 140
- Campbell measure, 87
- normalized, 77
- canonical measure, 52
- Poisson, 52
 - super-Brownian motion, 52
- catalytic branching, 95
- clan
- recurrent, 81
 - transience, 81
- closure, 6, 8, 11, 18, 30, 32, 36
- cluster representation, 52
- coarse graining, 110, 114, 116
- collision
- local time, 73
 - system, 91
- constitutive laws, 8, 30
- continuous random tree, 58
- Corrsin’s hypothesis, 36
- covariance
- operator function, 246
 - space, 252
- critical phenomena, 110
- cylindrical
- Brownian motion, 246
 - martingale, 246
 - standard Brownian motion, 246
- Darcy’s law, 11, 12, 34, 35, 37
- dead mass process, 69
- diffusive clustering, 75, 83
- diffusivity constant, 177
- direct numerical simulation, 18
- Dirichlet Laplacian, 318
- discrete Laplacian generator, 53
- discretization
- Euler scheme, 166, 170
 - Fourier domain, 165
 - stream function, 161
 - time, 161
- dislocation, 8
- dispersion, 34, 36–38
- coefficient, 35
- dissipation scale, 157
- effective equation, 6, 16, 30
- empirical measure process, 47
- entrance law, 83
- equation of state, 7, 8
- equilibrium clan, 84
- Euler scheme, 166
- exchange particle system, 46
- exchangeable particle system
- non-interacting, 46
- excursions, 53
- exponential martingale problem, 50
- extinction law, 52
- Feller branching
- fixed point property, 82
 - process, 51
- financial decision making, 3
- fluctuating hydrodynamic, 108, 111, 120, 134
- fluctuations, 108, 109
- critical, 129

- macroscopic effects of, 141, 147
- nonlinear, 111, 122, 125, 126, 130, 135, 141
 - of the shock, 116, 120
 - small(normal), 121, 122, 134, 135
- fractal, 35, 37
- fractional Brownian process, 163
- fracture, 9
- front tracking, 18
- full martingale problem, 64
- Gibbs measure, 129–131, 133, 138, 148
- Glauber dynamics, 109, 128
 - nonlinear fluctuations, 136
 - pattern formation, 141
 - with long range interactions, 131
- Glauber+Kawasaki dynamics, 146, 148
- historical
 - Brownian motion, 58
 - cluster representation, 59
 - martingale problem, 59
 - modulus of continuity, 60
- homogenization, 13, 17, 155, 176, 177, 179
 - simulations, 178
- Hurst exponent, 35
- hydrodynamic
 - behavior, 110
 - equations, 109, 112
 - laws, 107
 - scaling limit, 109, 112, 119, 120, 133
- hydrodynamic limit
 - central limit theorem, 110
 - large deviations, 148
 - law of large numbers, 110
- IC, 55
- incompressible limit, 112, 122
- independent particle process, 110, 113–115
- integral
 - martingale problem, 50
 - scale, 39, 157
 - scales, 38
- integrated super-excision, 57
- interacting particle system (IPS), 109
- interaction chain, 83
- interface
 - fluctuations, 127, 128
 - growth model, 126, 127
 - motion, 113, 147
- intermittency, 10, 15, 38
- Ising, 34
 - model, 109, 128, 130
 - Ising model, 131
- Kac potentials, 128
- Kardar–Parisi–Zhang equation, 127
- Kawasaki dynamics, 109, 128, 146, 148
 - nonlinear fluctuations, 140
 - with long range interactions, 131
- Kelvin–Helmholtz instability, 18
- Kolmogorov, 10, 13
 - law, 39
 - point, 39
 - spectrum, 38, 154
 - classical, 156
 - cut-off, 157
 - parameters, 157
 - theory of turbulence, 156
 - integral and dissipation scales, 157
- Lévy catalytic medium, 96
- Laplace functional, 51
- level set, 20
- lifting, 48
 - additive functional, 63
- local time
 - collision, 73
 - super-Brownian, 69
- log Laplace equation, 51
- Lorenz equations, 5
- low dimensional clumping, 84
- LSI, 246
- Lyapunov exponent, 155, 172
 - approximation, 170
 - positivity, *see also* mixing
- martingale measure, 249
- martingale problem, 49, 263
 - Historical, 59
 - MP_θ , 89
 - absolute continuity of measures, 261
 - collision system, 91
 - equivalent formulation, 49
 - exponential, 50
 - full, 64
 - integral, 50
 - $\text{MP}_{I.C.}$, 92
 - partial, 63
- measure-valued
 - additive functional, 62
 - cylindrical martingale, 249
- mesh refinement, 19
- metastable behavior, 129, 148
- microscopic models, 108
- mixing, 22, 31, 176

- layer, 25
- model
 - stepping stone, 100
- moment, 31
 - expansion, 15
- multi-phase flow, 9, 10, 21, 32
- multiplicative property
 - strong, 52
 - weak, 52
- Navier-Stokes
 - corrections, 110, 112, 115, 119
 - equation, 107, 110, 112, 122
- non local evolution equations, 132, 134
- non-interacting exchangeable particle system, 46
- normalized Campbell measure, 77
- occupation
 - density field, 69
 - time process, 57
- occupation measure
 - weighted, 68
- Ornstein-Uhlenbeck process, 111, 120, 121
- Palm measure, 78
- parametrix method, 100
- partial martingale problem, 63
- particle models
 - non-reversible, 110, 111, 148
 - reversible, 128, 131, 148
- particle system
 - exchangeable, 46
- path-valued process, 58
- permeability, 12, 34
- persistence, 74
- petroleum reservoir, 3
- phase
 - segregation, 144, 147
 - separation, 144, 147
 - transition, 110, 128–130, 132
- Poisson
 - cluster representation, 52
 - point processes, 155, 178
- porous
 - media, 11, 34, 37
 - medium, 12
- random
 - cluster, 47
 - field, 5, 7, 12, 13, 20, 34
- Rayleigh-Taylor
 - instability, 18, 19
 - mixing, 37
- reaction-diffusion equations, 108
 - stochastic, 111, 135, 140
- renormalization, 8
 - group, 16
- Reynolds stress, 32
- Richtmyer-Meshkov instability, 19
- Riemann problem, 19, 23
- RM
 - instability, 22
 - mixing, 22
- RNG, 38
 - dynamics, 27
 - equation, 17
 - operation, 16
- RT
 - instability, 21, 26
 - mixing, 22, 27, 33
 - problem, 24
- sampling rate, 92
- scale
 - macroscopic, 109, 110
 - microscopic, 109, 110
- scaling
 - diffusive, 109, 112, 113, 116, 120, 148
 - Euler, 109, 110, 112, 116
- Schwartz pair, 246
- second class particle, 118
- shear
 - band, 9
 - flow, 163, 164
- shocks
 - absence of, 110, 115
 - fluctuations of, 116, 119
 - formation of, 114
 - location of, 118
 - models with, 111
 - propagation of, 111
 - speed of the, 117
- simple exclusion process
 - asymmetric, 113
 - symmetric, 129
 - *weakly asymmetric, 119
- smooth function, 50
- statistical mechanics
 - equilibrium, 108, 110, 129, 140
 - non equilibrium, 109
- stepping stone model, 100
- stochastic
 - heat equation, 125, 128
 - stochastic integral, 254
 - normalized, 257
 - stochastic quantization equation, 296

- stationary solutions, 306
- absolute continuity, 309
- ergodicity, 314
- Markov property, 313
- singularity of measures, 315
- stream function, 159
- strong solution, 262
- structure function, 163
- super Brownian motion, 45, 275
- Taylor-Saffman instability, 18
- tempered distributions, 54
- total mass process, 51
- turbulence, 6, 10, 14, 15, 32
- two-phase flow, 31
- velocity field
 - homogeneous in space, 155
 - incompressibility, 156
 - isotropy, 156
 - Markovian, 156
 - stationarity, 155
 - vorticity, 168, 179
- wavelets, 167, 169, 178
- weak solution, 262
- weather, 5
 - forecasts, 3
- weighted occupation measure, 68
- white noise, 108, 119, 121, 122, 124, 125,
 - 134, 145, 250
- Wick products, 298
 - approximations, 303
 - multiple integrals, 301

Selected Titles in This Series

(Continued from the front of this publication)

- 33 **Vladimir G. Berkovich**, Spectral theory and analytic geometry over non-Archimedean fields, 1990
- 32 **Howard Jacobowitz**, An introduction to CR structures, 1990
- 31 **Paul J. Sally, Jr. and David A. Vogan, Jr., Editors**, Representation theory and harmonic analysis on semisimple Lie groups, 1989
- 30 **Thomas W. Cusick and Mary E. Flahive**, The Markoff and Lagrange spectra, 1989
- 29 **Alan L. T. Paterson**, Amenability, 1988
- 28 **Richard Beals, Percy Deift, and Carlos Tomei**, Direct and inverse scattering on the line, 1988
- 27 **Nathan J. Fine**, Basic hypergeometric series and applications, 1988
- 26 **Hari Bercovici**, Operator theory and arithmetic in H^∞ , 1988
- 25 **Jack K. Hale**, Asymptotic behavior of dissipative systems, 1988
- 24 **Lance W. Small, Editor**, Noetherian rings and their applications, 1987
- 23 **E. H. Rothe**, Introduction to various aspects of degree theory in Banach spaces, 1986
- 22 **Michael E. Taylor**, Noncommutative harmonic analysis, 1986
- 21 **Albert Baernstein, David Drasin, Peter Duren, and Albert Marden, Editors**, The Bieberbach conjecture: Proceedings of the symposium on the occasion of the proof, 1986
- 20 **Kenneth R. Goodearl**, Partially ordered abelian groups with interpolation, 1986
- 19 **Gregory V. Chudnovsky**, Contributions to the theory of transcendental numbers, 1984
- 18 **Frank B. Knight**, Essentials of Brownian motion and diffusion, 1981
- 17 **Le Baron O. Ferguson**, Approximation by polynomials with integral coefficients, 1980
- 16 **O. Timothy O'Meara**, Symplectic groups, 1978
- 15 **J. Diestel and J. J. Uhl, Jr.**, Vector measures, 1977
- 14 **V. Guillemin and S. Sternberg**, Geometric asymptotics, 1977
- 13 **C. Pearcy, Editor**, Topics in operator theory, 1974
- 12 **J. R. Isbell**, Uniform spaces, 1964
- 11 **J. Cronin**, Fixed points and topological degree in nonlinear analysis, 1964
- 10 **R. Ayoub**, An introduction to the analytic theory of numbers, 1963
- 9 **Arthur Sard**, Linear approximation, 1963
- 8 **J. Lehner**, Discontinuous groups and automorphic functions, 1964
- 7.2 **A. H. Clifford and G. B. Preston**, The algebraic theory of semigroups, Volume II, 1961
- 7.1 **A. H. Clifford and G. B. Preston**, The algebraic theory of semigroups, Volume I, 1961
- 6 **C. C. Chevalley**, Introduction to the theory of algebraic functions of one variable, 1951
- 5 **S. Bergman**, The kernel function and conformal mapping, 1950
- 4 **O. F. G. Schilling**, The theory of valuations, 1950
- 3 **M. Marden**, Geometry of polynomials, 1949
- 2 **N. Jacobson**, The theory of rings, 1943
- 1 **J. A. Shohat and J. D. Tamarkin**, The problem of moments, 1943

ISBN 0-8218-2100-8



9 780821 821008

SURV/64.S

AMS on the Web
www.ams.org