

INAUGURAL - DISSERTATION  
zur  
Erlangung der Doktorwürde  
der  
Naturwissenschaftlich-Mathematischen Gesamtfakultät  
der  
RUPRECHT - KARLS - UNIVERSITÄT  
HEIDELBERG

vorgelegt von  
**Diplom-Physiker Christian Reichert**  
aus Germersheim

Tag der mündlichen Prüfung: 04.10.2006



# Deterministic and Stochastic Modelling of a Catalytic Surface Reaction

Gutachter: **Prof. Dr. Dr. h. c. mult. Willi Jäger**  
**Priv.-Doz. Dr. Karl Oelschläger**



*Für Sév und Léonie*



## Abstract

Catalytic surface reactions are of great importance both for chemical industry and as model systems for the study of pattern formation far from thermodynamic equilibrium. A reaction that has been investigated extensively in experiments is the oxidation of carbon monoxide on platinum. In the present work we first develop a mathematical model for CO oxidation on Pt which is valid over a wide pressure range. This requires the use of different model types. While at low pressures in the gas phase the system can be described by a deterministic model in the form of ordinary or partial differential equations, a stochastic particle model is needed at higher pressures due to rising fluctuations. A numerical bifurcation analysis for the deterministic model is performed, which yields good agreement with experimental findings. Subsequently, we investigate the consistency of deterministic differential equations models and stochastic particle models for reaction-diffusion systems in a more general setting. We rigorously derive partial differential equations as limit dynamics of certain linear and nonlinear ‘mesoscopic’ stochastic particle models in the limit of large particle numbers. The convergence proofs combine techniques from numerical analysis and the theory of Markov processes. Finally, we use the stochastic particle model for CO oxidation on Pt to simulate the spontaneous nucleation and subsequent dying out of pulses (‘raindrop patterns’) that has been observed experimentally.

## Zusammenfassung

Katalytische Oberflächenreaktionen sind von großer Bedeutung sowohl für die chemische Industrie als auch als Modellsysteme für die Untersuchung von Strukturbildung weit weg vom thermodynamischen Gleichgewicht. Ein experimentell intensiv untersuchtes Beispiel ist die Oxidation von Kohlenmonoxid an Platinoberflächen. In der vorliegenden Arbeit entwickeln wir zunächst ein mathematisches Modell für die CO-Oxidation an Pt, das über einen weiten Druckbereich Gültigkeit besitzt. Hierzu ist es erforderlich, verschiedene Modelltypen zu verwenden. Während bei niedrigen Drücken in der Gasphase deterministische Modelle in der Form von gewöhnlichen oder partiellen Differentialgleichungen eine gute Beschreibung des Systems bieten, ist es bei höheren Drücken aufgrund der auftretenden Fluktuationen erforderlich, ein stochastisches Vielteilchenmodell zu verwenden. Eine numerische Bifurkationsanalyse des deterministischen Modells ergibt eine gute Übereinstimmung mit experimentellen Resultaten. Anschließend untersuchen wir in einem allgemeineren Rahmen die Konsistenz von Differentialgleichungsmodellen und ‘mesoskopischen’ stochastischen Vielteilchenmodellen für Reaktions-Diffusions-Systeme. Partielle Differentialgleichungen ergeben sich als Approximation der stochastischen Dynamik im Limes großer Teilchenzahlen. Für die Konvergenzbeweise benutzen wir Techniken aus der numerischen Analysis und der Theorie der Markov-Prozesse. Schließlich verwenden wir das stochastische Modell für die CO-Oxidation an Pt, um ‘Regentropfenmuster’, d.h., die spontane Nukleation von Pulsen verbunden mit anschließendem Aussterben, zu simulieren.



# Contents

<b>List of mathematical notation</b>	<b>iv</b>
<b>Introduction</b>	<b>1</b>
Complex systems and mathematical modelling . . . . .	1
Contents and main results . . . . .	7
Remarks on notation and style . . . . .	11
<b>1 Experimental background and mathematical modelling</b>	<b>13</b>
Overview . . . . .	13
1.1 Experimental background . . . . .	13
1.1.1 Experimental methods . . . . .	14
1.1.2 Reaction steps of CO oxidation on Pt . . . . .	15
1.1.3 Spatio-temporal pattern formation . . . . .	18
1.2 Macroscopic deterministic modelling . . . . .	21
1.2.1 Chemical reactions and surface diffusion . . . . .	21
1.2.2 Mass balance . . . . .	26
1.2.3 Numerical bifurcation analysis . . . . .	27
1.2.4 Heat production and transfer . . . . .	28
1.2.5 Energy balance . . . . .	31
1.2.6 Summary . . . . .	32
1.3 Mesoscopic stochastic modelling . . . . .	33
1.3.1 Jumps caused by reaction and diffusion events . . . . .	34
1.3.2 Temperature drift . . . . .	35
1.3.3 Summary . . . . .	36
Discussion . . . . .	37

<b>2</b>	<b>Law of large numbers for linear models</b>	<b>39</b>
	Overview . . . . .	39
2.1	Macroscopic PDE models and mesoscopic stochastic particle models . . . . .	41
	2.1.1 The macroscopic PDE model . . . . .	42
	2.1.2 The mesoscopic stochastic particle model . . . . .	44
2.2	The macroscopic PDE and a semi-discrete approximation . . . . .	52
	2.2.1 Weak formulation of the PDE . . . . .	52
	2.2.2 A semi-discrete finite-difference approximation . . . . .	56
	2.2.3 External approximation schemes . . . . .	59
2.3	Law of large numbers for the example model . . . . .	63
	2.3.1 Convergence of the approximation . . . . .	63
	2.3.2 Convergence of the particle density . . . . .	67
2.4	Law of large numbers for the general linear model . . . . .	73
	2.4.1 Convergence of the approximation . . . . .	73
	2.4.2 Convergence of the particle density . . . . .	75
	Discussion . . . . .	79
<b>3</b>	<b>Nonlinearities and a refined law of large numbers</b>	<b>81</b>
	Overview . . . . .	81
3.1	Nonlinear reaction kinetics . . . . .	82
	3.1.1 Lipschitz conditions . . . . .	82
	3.1.2 Invariant regions . . . . .	87
3.2	Nonlinear diffusion . . . . .	97
	3.2.1 Crowding effects . . . . .	97
	3.2.2 Gradient-activated diffusion . . . . .	108
3.3	A refined law of large numbers . . . . .	115
	Discussion . . . . .	117
<b>4</b>	<b>Stochastic simulations</b>	<b>119</b>
	Overview . . . . .	119
4.1	The simulation algorithm . . . . .	119
4.2	Simulation of raindrop patterns . . . . .	123
	Discussion . . . . .	126
	<b>References</b>	<b>127</b>
	<b>List of figures</b>	<b>132</b>

List of tables	133
Index	135

## List of mathematical notation

$n_s, n_r$	number of species and number of reactions
$\mathbb{R}_0^+$	the nonnegative reals
$G$	bounded domain in $\mathbb{R}^m$
$Q_T$	space-time cylinder $G \times (0, T)$
$\mathcal{Z}_h(z_0)$	grid in $\mathbb{R}^m$ with mesh $h$ and origin $h z_0$
$\mathcal{G}_h, \mathcal{G}_h^1, \bar{\mathcal{G}}_h, \bar{\mathcal{G}}_h^1$	sets of grid points related to the domain $G$
$L^2(G)$	square-integrable real-valued functions on $G$
$H^1(G), H_0^1(G)$	Sobolev spaces of functions on $G$
$H^{-1}(G)$	dual space of $H_0^1(G)$
$\mathcal{L}^2(\mathcal{G}_h), \mathcal{L}^2(\bar{\mathcal{G}}_h)$	discrete analogues of $L^2(G)$
$\mathcal{H}_0^1(\mathcal{G}_h)$	discrete analogue of $H_0^1(G)$
$\mathcal{H}^1(\bar{\mathcal{G}}_h)$	discrete analogue of $H^1(G)$
$L^2(G)$	cartesian product of $n_s$ copies of $L^2(G)$
$H_0^1(G)$	cartesian product of $n_s$ copies of $H_0^1(G)$
$H^{-1}(G)$	dual space of $H_0^1(G)$
$\mathcal{L}^2(\mathcal{G}_h)$	cartesian product of $n_s$ copies of $\mathcal{L}^2(\mathcal{G}_h)$
$\mathcal{H}_0^1(\mathcal{G}_h)$	cartesian product of $n_s$ copies of $\mathcal{H}_0^1(\mathcal{G}_h)$
$\mathcal{H}^{-1}(\mathcal{G}_h)$	dual space of $\mathcal{H}_0^1(\mathcal{G}_h)$
$L^2(0, T; X)$	square-integrable functions from $[0, T]$ to the Banach space $X$
$H^1(0, T; H_0^1(G), L^2(G))$	functions in $L^2(0, T; H_0^1(G))$ having a generalized time derivative in $L^2(0, T; H^{-1}(G))$
$C([0, T], X)$	continuous functions from $[0, T]$ to the Banach space $X$
$\partial_i = \partial_{x_i} = \partial/\partial x_i$	partial derivative with respect to $x_i$
$\partial_i^+, \partial_i^-$	forward and backward discrete partial derivatives
$\langle \cdot, \cdot \rangle_X$	dual pairing between $X$ and $X^*$
$(\cdot, \cdot)_X$	scalar product in $X$
$\rightarrow, \rightharpoonup$	strong and weak convergence
$\Omega$	sample space
$P$	probability measure
$\mathcal{A}, \mathcal{F}_t$	sigma-fields of subsets of $\Omega$
$2^A$	power set of $A$
$E[\cdot], E \int \dots$	expected value
$a \wedge b$	minimum of $a$ and $b$
$a \vee b$	maximum of $a$ and $b$

# Introduction

‘Large composite systems are variegated and full of surprises. Perhaps the most wonderful is that despite their complexity on the small scale, sometimes they crystallize into large-scale patterns that can be conceptualized rather simply, just as crazy swirls of colors crystallize into a meaningful picture when we step back from the wall and take a broader view of a mural.’

Auyang (1998)

This thesis is concerned with the mathematical modelling of a complex physico-chemical system: the oxidation reaction of carbon monoxide on a platinum surface. Although the terms ‘complex system’ and ‘mathematical modelling’ are by now widely used, we shall try in the first part of this introduction to sketch their meaning and our personal point of view. In the second part we give an overview of the contents of the thesis and the main results. This part assumes a certain familiarity of the reader with the subject matter. We conclude this introduction by some remarks on notation and style.

## Complex systems and mathematical modelling

### Complex systems

A system is understood in various branches of science as a part of reality that is composed of many interacting constituents. It is separated, at least conceptually, from the rest of the world which takes the role of the system’s environment: solids, liquids and gases as well as biological macromolecules are composed of single atoms or simple molecules; organisms are made up of a large number of cells; in the brains of behaving animals information is processed by a network of neurons; biological populations are composed of individual organisms; an economy is a system of consumers and producers. ‘Complex systems’ is by now the most widely accepted label for an interdisciplinary field of research that started to emerge a few decades ago. (Some classical references are, e.g., von Bertalanffy (1968); Nicolis & Prigogine (1977); Haken (1983).)

What makes a particular system a ‘complex’ system? In a broad sense a system may be called complex if it is complicated by some subjective judgement. A still fuzzy but

more restrictive working definition is to call a system complex if its constituents, by their cooperative activities, are able to form and maintain spatio-temporal patterns or functional structures on the macroscopic (system) level.

This is best explained by an example. A classical magician's trick is 'turning water into wine': a glass of slightly alkaline water is poured into a wine glass on the bottom of which are hidden a few drops of phenolphthalein solution. This organic compound is colourless in neutral solution. However, it has an intense red colour in alkaline solution, which yields the desired effect (as long as no one wants to taste the 'wine'). At second sight this change of colour is not surprising. It can be readily explained by the fact that in alkaline solution the phenolphthalein molecule loses two hydrogen atoms. This shifts absorption from the ultra-violet to the blue-green part of the spectrum and makes the solution appear red. The periodic switching of the colour of a solution from blue to red during the celebrated Belousov-Zhabotinsky (BZ) reaction, on the other hand, caused protest among chemists who saw in it a violation of the laws of thermodynamics (Zhabotinsky, 1991; Murray, 2004a). These colour changes are due to oscillations of the concentration of a redox indicator. They are an example of temporal pattern formation and cannot be explained by a simple extrapolation of the features of a single molecule to the whole system. Their appearance constitutes a 'new' or 'emergent' property of the system. If the BZ reaction is prepared in a Petri dish one can observe, in addition, spatio-temporal concentration patterns such as spirals and target patterns (the spreading of concentric rings).

The BZ reaction is a homogeneously catalysed reaction because the catalyst and the reactants are in the same phase. Surface reactions such as the oxidation of carbon monoxide on platinum belong to the class of heterogeneously catalysed reactions. Oscillations in the reaction rate of CO oxidation on Pt had first been observed at atmospheric pressures with a polycrystalline catalyst. In later studies on well-defined single-crystal surfaces a large variety of patterns has been found that are often similar to those observed with the BZ reaction (Ertl, 1991; Eiswirth & Ertl, 1995; Imbihl & Ertl, 1995). CO oxidation on Pt has since become a model system for the study of pattern formation in reaction-diffusion systems. Other prominent surface reactions are, e.g., the NO + CO reaction on platinum or the NO + H<sub>2</sub> reaction on rhodium which exhibits spirals with 'corners' (Imbihl & Ertl, 1995).

Pattern formation can be observed also in biological systems (oscillations in populations; cardiac rhythm; circadian rhythms, i.e., periodic changes between sleep and awakesness; mammalian coat patterns, e.g., the famous question 'how the leopard got its spots'; cf. Murray (2004a,b)). Other classical examples of pattern formation are convection patterns in fluids (e.g., Rayleigh-Bénard convection or Taylor-Couette flow; cf. Cross & Hohenberg (1993)).

Functional structures do not appear in chemical systems nor hydrodynamic systems but are omnipresent in biology. We mention two important examples: proteins and the brain. Proteins are biological macromolecules. They are chains composed of 20 types of amino-

acids which are assembled according to the genetic code in the genome (Alberts et al., 2004). These chains fold in a complicated three-dimensional structure of minimal free energy in which up to four levels can be distinguished. Because of their characteristic shape and surface topography, proteins generally show great specificity in their ability to bind to other molecules. Their particular structure thus enables proteins to perform specialised tasks, e.g., to catalyse biochemical reactions or to serve as receptor molecules.

Information is processed in the brain by a large assembly of interconnected neurons. On the macroscopic level the brain has sometimes been divided on anatomical grounds in different areas that were thought to specialise in different kinds of information. It seems, however, that this point of view is only partially correct. Even relatively simple tasks, e.g., the tactual discrimination of the shape of two objects, are performed in collaboration by different macro-circuits of a size of the order of  $10^7$  neurons located in different areas so that there is no localised neuronal population responsible for one particular task. Moreover, the populations that are involved in different tasks may overlap (Roland, 1993).

### Mathematical modelling of complex systems

An interesting account of the meaning of ‘mathematical modelling’ is given in (Aris, 1978), where a mathematical model is characterised as a ‘complete and consistent set of mathematical equations which is thought to correspond to some other entity, its prototype’. We agree with this characterisation, although ‘set of mathematical equations’ should perhaps be replaced by some more general concept. Virtually all mathematical models of complex systems come in the guise of either a dynamical system (in the mathematical sense) or a stochastic process, two concepts that shall be introduced informally in the sequel.

A *dynamical system* is a triple consisting of a time set  $\mathcal{T}$ , a *state space*  $S$ , and a family of operators  $\phi_t : S \rightarrow S$ ,  $t \in \mathcal{T}$ , that represent the time evolution of the system. That is, an initial state  $u_0 \in S$  is mapped by  $\phi_t$  to the state  $u(t) = \phi_t(u_0) \in S$  at time  $t$ . A dynamical system is the mathematical formalisation of the concept of a deterministic evolution. If the system is not controlled from the outside, i.e., if it is *autonomous*, the evolution operators usually satisfy the condition  $\phi_{t+s}(u_0) = \phi_t(\phi_s(u_0))$ , which implies  $u(t+s) = \phi_t(u(s))$ .

The mathematical concept that formalises a random evolution is a *stochastic process*. Randomness typically comes into play in two different ways: it can be caused by either a fluctuating environment or by the randomness of certain intrinsic processes, e.g., chemical reactions. Randomness is represented mathematically by a sample space  $\Omega$  on which a probability measure  $P$  is defined that assigns a probability to certain ‘reasonable’ subsets of  $\Omega$  called *events*. Every ‘reasonable’ mapping  $u$  from  $\Omega$  to a space  $S$  can then be thought of as a *random element* in  $S$  or a *random variable* taking values in  $S$ . A stochastic process is a triple consisting of a time set  $\mathcal{T}$ , a state space  $S$  and a family of random variables  $u(t)$ ,  $t \in \mathcal{T}$ , that take values in  $S$ . That is, for each  $t \in \mathcal{T}$ ,  $u(t)$  is a random element in  $S$ .

Let us discuss how some of the examples for complex systems introduced above can be modelled by dynamical systems or stochastic processes. Generally, for complex systems are composed of a large number of constituents, mathematical models can be set up on the level of the constituents (microscopic models, individual-based models) or on the system level (macroscopic models). In addition, there are mesoscopic models that ‘interpolate’ between micro- and macro-models. Which level of detail is appropriate depends on the specific problem at hand. The question whether there is a supremacy of microscopic over macroscopic models, in the sense that the latter can be ‘derived’ from the former, incites philosophical debates (Auyang, 1998).

Pattern formation in reaction-diffusion systems in homogeneous solution, as e.g. the BZ reaction, are typically modelled on the macroscopic level by systems of ordinary or partial differential equations (PDEs) for the concentrations of the chemical species. Here the time set consists of the nonnegative reals, the state space is  $\mathbb{R}^n$  or an appropriate function space (e.g.,  $L^2$ , the space of square-integrable functions) and the evolution operators are usually obtained implicitly by solving the differential equations. The same approach can be applied to model surface reactions (Krischer et al., 1992; Bär, 1993) and biological populations (Murray, 2004a,b).

On the microscopic level reaction-diffusion systems are often modelled by ‘microscopic’ stochastic particle models (interacting particle systems, lattice-gas models, stochastic spatial models), where mutually interacting particles randomly move on a lattice, say  $\mathbb{Z}^m$  ( $m = 1, 2, 3$ ), representing certain ‘sites’ (e.g., adsorption sites in the case of surface reactions). Here the state space is the space of all possible particle configurations on the lattice, i.e.,  $S = C^{\mathbb{Z}^m}$ , where the set  $C$  contains all possible configurations of a single site. In the simplest case  $C$  might be the set  $\{0, 1\}$ , where 0 stands for empty, and 1 for occupied. For CO oxidation on Pt the first model of this type was introduced by Ziff et al. (1986). A more elaborate model is described in Rosé et al. (1994). In a coarser ‘mesoscopic’ approach the particles move randomly on a lattice now representing certain cells or compartments of mesoscopic size. Particles are randomly created or destroyed with intensities that depend on the local particle density. The state space is given by the collection of all possible particle numbers in the cells, i.e.,  $S = (\mathbb{N}_0^{n_s})^{\mathbb{Z}^m}$ , where  $n_s$  denotes the number of species. In computer simulations the use of microscopic particle models is at present limited to rather small systems, e.g., CO oxidation on a platinum field emitter tip where only about  $500 \times 500$  adsorption sites are involved. For the description of a fluctuation phenomenon in CO oxidation on Pt which occurs on a larger scale ( $\approx 1 - 100 \mu\text{m}$ ) and involves about  $10^9$  adsorption sites, so-called raindrop patterns, a mesoscopic approach is more appropriate (cf. Chapters 1 and 4). The mesoscopic models find application also in biology, e.g. in the modelling of fluctuations in the rates of gene expression, which can lead to the appearance of different phenotypic outcomes in initially uniform populations (Arkin et al., 1998).

We briefly comment on mathematical models for the complex systems exhibiting functional structures mentioned above. Comparing to reaction-diffusion systems, for instance, it is probably fair to say that the mathematical modelling of such systems has so far been of limited success. Proteins and neural networks are at present modelled almost exclusively on the microscopic level (Snow et al., 2005; Dayan & Abbott, 2001), although in the context of neural networks there do exist models working with population densities (Ermentrout, 1998). Since proteins are macromolecules, their state can be specified on the microscopic level by giving the position and momentum of each of their atoms. The atoms are assumed to interact with each other via phenomenological forces. In addition one must introduce a model for the protein's environment (the solvent). In order to predict the final shape and functionality of the protein, one would like to use such a microscopic model to simulate the folding process. However, molecular dynamics simulations are currently limited to the nanosecond to microsecond regime, while even small proteins fold on the microsecond to second timescale (Snow et al., 2005). This shows the limitations of a 'brute force' microscopic approach. Similarly, neural networks are often modelled by large systems of coupled differential equations for the dynamics of the firing rates or the membrane potentials of all neurons. Certain aspects of brain activity and learning can successfully be reproduced by such models. However, the connection to actual neural circuits in the brain and the emergence of more complicated behaviour remain unclear.

We have pointed out that complex systems at various levels of detail are modelled either by dynamical systems or stochastic processes. These are two vast sub-areas of contemporary applied mathematics. The problems addressed there range from the fundamental problem of establishing the existence of certain dynamical systems or stochastic processes to the analysis of their qualitative features. Which mathematical questions are particularly relevant for the study of complex systems? We give a few examples.

- **Bifurcations.** In the study of pattern formation one is interested in a pattern that is established after a certain period of transient behaviour. This corresponds to the fact that the trajectories of autonomous 'dissipative' dynamical systems asymptotically approach a subset of the state space  $S$  called the 'global attractor'. It may contain or consist of stable steady states (states  $u \in S$  for which  $\phi_t(u) = u$  for all  $t \in \mathcal{T}$ ), stable periodic orbits (subsets  $B$  of  $S$  for which exists a period  $T$  such that  $\phi_{t+T}(u) = u$  for all  $u \in B$ ) or even a 'strange attractor'. Roughly speaking, qualitative changes in the structure of the global attractor (e.g., changes in the number of stable steady states or stable periodic orbits) upon variation of parameters are called bifurcations. Spatio-temporal pattern formation is often explained by bifurcation: the pattern is associated with the appearance of a nontrivial solution bifurcating from a trivial or homogeneous one. The physical background is that pattern formation typically occurs in open systems (systems that exchange energy and/or matter with their environment)

far away from thermodynamic equilibrium. Typically, one can experimentally tune a control parameter, e.g., a temperature gradient, that gradually drives the system from thermodynamic equilibrium, where it is uniform, to the non-equilibrium situation where patterns are formed. In order to understand a dynamical system depending on parameters one tries to figure out its bifurcation diagram or tableau. That is, one tries to determine all regions in parameter space with qualitatively different dynamics, either analytically or numerically (Guckenheimer, 1986).

- ▶ **Model reduction.** Often one is interested in finding out the ‘effective’ number of degrees of freedom in a dynamical system. In the simplest case certain systems of ordinary differential equations (ODEs) can be reduced, at least heuristically, by setting the time derivative of some ‘fast-relaxing’ variables to zero. In a more complicated situation one might want to prove that the behaviour of an infinite-dimensional dynamical system defined by a PDE can approximately be determined by the projection of the dynamics on a finite-dimensional subspace of the infinite-dimensional state space.
- ▶ **Coarse graining.** We have already seen that complex systems may be modelled on at least two levels, the microscopic (constituent) level and the macroscopic (system) level. The introduction of models at several levels of detail naturally raises the question whether these models are consistent and how they can be linked together. For reaction-diffusion systems macroscopic PDE models and microscopic or mesoscopic stochastic particle models can sometimes be linked by a so-called law of large numbers. That is, in the limit of large particle numbers, a particle density associated to the stochastic model approaches the solution of the macroscopic PDE.
- ▶ **Stochastics.** Fluctuations, either due to environmental influences or internally created, offers new possibilities for the dynamical behaviour. If the dynamics is purely deterministic, then the system will always stay close to a stable steady state, for instance, once it has reach its vicinity. In the presence of fluctuations the system might leave this vicinity and then ‘deterministically’ approach a second stable steady state. The interplay between stochastic and deterministic dynamics may lead to a sort of ‘oscillation’ with a random period. Often a quantity of interest for this type of problem is an exit time  $\tau_B$ , the time to reach the boundary of a neighbourhood  $B$  of a certain state  $u$  when starting in  $u$  (or, in other words, the hitting time of  $S \setminus B$ ). Exit times and hitting times are random variables whose distribution or expected value can sometimes be calculated.

The short selection above is by no means complete. Some of these typical problems are important for the understanding of CO oxidation on Pt and will be tackled later, either with mathematical rigour, by a computational approach, or in a purely heuristic way.

## Contents and main results

The main motivation for the present work is the experimental observation of temporal, spatio-temporal and noise-induced pattern formation in the oxidation of carbon monoxide on platinum single crystal surfaces. Therefore we present the necessary experimental background to CO oxidation on Pt in the first section of the next chapter and give an overview of the observed patterns at low ( $\lesssim 10^{-4}$  mbar) to intermediate ( $\approx 10^{-2}$  mbar) pressures in the gas phase. We shall focus mainly on the Pt(110) surface and mention only briefly the surfaces with Miller indices (100) and (111). In order to capture both deterministic and random phenomena, we then introduce a macroscopic PDE model and a mesoscopic stochastic particle model, both including a temperature variable. It follows a discussion of the bifurcation structure of the kinetic part of the macroscopic model with fixed temperature. Heat effects are considered in more detail in Chapter 4. Our model extends and improves the well-known (isothermal) model for Pt(110) by Krischer et al. (1992) in several respects. As compared to Krischer et al. (1992), the term for oxygen adsorption is changed, which yields a better correspondence with experimental results at higher pressures. Moreover, we use a different ansatz for the kinetics of the surface structural phase transition so that the model can easily be adapted to other platinum surfaces by changing certain parameters. The present work is a continuation of Reichert (2000) and Reichert et al. (2001), where we focused on the isothermal and homogeneous case.

The introduction of two different models, a macroscopic deterministic model and a mesoscopic stochastic particle model, naturally raises the question whether these models are consistent. This issue is dealt with in a quite general setting in Chapters 2 and 3 where we investigate how certain classes of PDE models and stochastic particle models are related. We prove rigorously that a particle density associated to the stochastic particle model converges to the solution of the PDE in the limit of large particle numbers.

The stochastic particle models we call mesoscopic can be thought of as a combination of a continuous-time version of the classical urn model by P. and T. Ehrenfest for diffusion through a membrane (see, e.g., Karlin & Taylor (1975)) and the standard stochastic model for chemical reactions (see, e.g., van Kampen (1992)). That is, we think of a chemical reactor as being composed of cells or compartments of length  $l$ . Each cell may contain up to about  $n$  particles of each species. These particles jump randomly from a cell to an adjacent one with rate  $d$ . Moreover, if we denote the vector of particle concentrations (the particle numbers divided by  $n$ ) by  $\mathbf{u}_l = (u_{l,1}, \dots, u_{l,n_s})$ ,  $n_s$  being the number of species, the particle numbers in a cell change according to  $n_r$  possible reactions that occur randomly with rates  $n K_i(\mathbf{u}_l)$ . Stochastic particle models of this type have been described since the early seventies by many authors in physics (Nicolis & Prigogine, 1977; Gillespie, 1977; Haken, 1983; van Kampen, 1992; Gardiner, 2004) and mathematics (Kurtz, 1981; Arnold & Theodosopulu, 1980; Kotelenez, 1986, 1988; Blount, 1991, 1993, 1994; Guias, 2002). In the physical literature

the model is often simply called ‘the’ stochastic model for chemical reactions; in mathematics it also goes under the name of ‘density-dependent population process’.

The methods for proving laws of large numbers for such particle models are closely related to techniques for solving the limit PDEs. In order to distinguish our approach from the one based on Kotelenetz’s work (Kotelenetz, 1986, 1988) which has been used in the literature cited above, we start with a short, informal discussion of the simplest nonlinear limit PDE. That is, we consider a scalar reaction-diffusion equation

$$\partial_t u - \Delta u = f(u) \tag{1}$$

on a bounded domain  $G \subset \mathbb{R}^m$  with sufficiently smooth boundary together with (homogeneous) Dirichlet boundary conditions. Here  $f$  is a linear combination of the reaction rates  $K_i$  involving the stoichiometric coefficients.

As is often the case in PDE theory, we look for generalised solutions. In the *semigroup approach* one first defines  $S(t) = e^{t\Delta}$ , the semigroup of bounded linear operators on  $L^2(G)$  generated by the Laplacian (with Dirichlet boundary conditions). By variation of constants, a classical solution of (1) also satisfies

$$u(t) = S(t)u(0) + \int_0^t S(t-s)f(u(s)) ds. \tag{2}$$

Any function  $u : [0, T] \rightarrow L^2(G)$  that satisfies Eq. (2), on the other hand, is called a generalised *mild* solution of Eq. (1).

In the *variational approach* one multiplies Eq. (1) with a test function  $v$  and integrates over the domain  $G$ . After an integration by parts one observes that a classical solution also satisfies

$$\frac{d}{dt}(u(t), v)_{L^2(G)} + a(u(t), v) = (f(u(t)), v)_{L^2(G)} \quad \text{for all } v \in H_0^1(G), \tag{3}$$

where the bilinear form  $a(\cdot, \cdot)$  on  $H_0^1(G) \times H_0^1(G)$  is given by

$$a(u, v) = (\nabla u, \nabla v)_{(L^2(G))^m}.$$

A function  $u \in H^1(0, T; H_0^1(G), L^2(G))$  that solves (3) is called a *weak* solution of Eq. (1).

While for semilinear equations as the one discussed above both the semigroup approach and the variational approach are successful, it is probably fair to say that the variational approach is more powerful when tackling quasilinear equations, e.g., equations involving a nonlinear diffusion operator.

In order to connect the stochastic process modelling the dynamics of the absolute particle numbers that has been described above to the solution of the PDE (1), we rescale it and define an appropriate step function-valued particle density  $u_l$  on a set of grid points  $\mathcal{G}_l$  that corresponds to the continuous domain  $G$ . (Recall that  $l$  is the edge length of a cell.) With the

aid of Dynkin's formula it can be seen that this particle density solves a system of stochastic differential equations

$$du_l(z, t) = ((d/2m)l^2 \Delta_l u_l + f(u_l)) dt + dM_l(z, t), \quad z \in \mathcal{G}_l^1. \quad (4)$$

Here  $\Delta_l$  is a discretized Laplacian, the processes  $M_l(z, \cdot)$  are certain martingales and  $\mathcal{G}_l^1$  is the 'interior' of  $\mathcal{G}_l$ . This observation can be used to guess the limit PDE: if one shrinks the cell size  $l$  and increases the number of particles per cell  $n$  and the hopping rate  $d$  in an appropriate way, then, one hopes, the fluctuating martingale terms will become small, and the equations defined on the grid  $\mathcal{G}_l^1$  will approximate the equation defined on the continuous domain  $G$ .

There are (at least) two possibilities to make these considerations rigorous, which are closely related to the semigroup approach and the variational approach to solving the limit PDE. In the first approach (Kotelenez, 1986, 1988) one observes that by variation of constants the stochastic particle density  $u_l$  satisfies the equation

$$u_l(t) = S_l(t)u_l(0) + \int_0^t S_l(t-s)f(u_l(s)) ds + \int_0^t S_l(t-s) dM_l(s). \quad (5)$$

Here  $S_l(t)$  is the semigroup generated by the discrete Laplacian, and the second integral term is a so-called stochastic convolution integral. Then one subtracts Eq. (5) from Eq. (2) and estimates  $\|u(t) - u_l(t)\|_{L^2(G)}$ . This yields a law of large numbers of the form

$$P \left[ \sup_{t \leq T} \|u - u_l\|_{L^2(G)} \geq \varepsilon \right] \rightarrow 0 \quad (l \rightarrow 0) \quad (6)$$

for arbitrary  $\varepsilon > 0$ .

Our technique, which will be introduced in Chapters 2 and 3, is related to the variational approach to solving Eq. (1). We also start with the observation that Eq. (4) looks like a spatially semi-discretized finite difference approximation of Eq. (1) perturbed by a martingale noise term. In a first step we then consider the (deterministic) step function-valued solutions  $v_l$  of

$$\frac{d}{dt}v_l(z, t) - (d/2m)l^2 \Delta_l v_l(z, t) = f(v_l(z, t)), \quad z \in \mathcal{G}_l^1, \quad (7)$$

and show that they converge to the weak solution of Eq. (1) in  $L^2(0, T; L^2(G))$ . Subsequently, in the second step, we show that

$$E \int_0^T \|u_l - v_l\|_{L^2(G)}^2 dt \rightarrow 0 \quad (l \rightarrow 0). \quad (8)$$

In order to obtain (8), we first use Dynkin's formula to show that the process

$$\begin{aligned} & \|u_l(t) - v_l(t)\|_{\mathcal{L}^2(\mathcal{G}_l)}^2 - \|u_l(0) - v_l(0)\|_{\mathcal{L}^2(\mathcal{G}_l)}^2 \\ & + 2 \int_0^t a_l(u_l(s) - v_l(s), u_l(s) - v_l(s)) ds \\ & - 2 \int_0^t (f(u_l(s)) - f(v_l(s)), u_l(s) - v_l(s))_{\mathcal{L}^2(\mathcal{G}_l)} ds - R_l(t), \quad t \leq T, \end{aligned} \quad (9)$$

is a (local) martingale. Here  $a_l(\cdot, \cdot)$  is the discrete analogue of the bilinear form  $a(\cdot, \cdot)$ ,  $\mathcal{L}^2(\mathcal{G}_l)$  is a finite-dimensional subspace of step-functions in  $L^2(G)$ , and  $R_l$  is a remainder term that vanishes in the limit. Formula (9) clearly reveals the relationship between the stochastic particle model and the variational structure underlying Eq. (1). Sufficient conditions for convergence are  $(d/2m)l^2 \rightarrow 1$  and  $d/n \rightarrow 0$  as  $n \rightarrow \infty, l \rightarrow 0$ . These are the same as the conditions needed in Kotelenz (1988) to show (6).

The ideas outlined above are carried out in Chapter 2 first for linear models, in order to introduce our method of proof with a minimum of technical and notational difficulties. In Chapter 3 we then study the convergence of certain nonlinear stochastic particle models. We first treat models with Lipschitz continuous reaction rates that have semilinear systems of the type of Eq. (1) as limit dynamics. It turns out that this constitutes only a slight extension of the linear case. Furthermore, we discuss the more realistic case of reaction functions  $\mathbf{f} = (f_1, \dots, f_{n_s})$  admitting an *invariant region*. More precisely, we assume that the vector field  $\mathbf{f}$  induced by the reaction rates  $K_i$  points inwards on the boundary of the cube  $[0, 1]^{n_s}$ . In addition, we allow for non-diffusing species. This model generalises the single-species models with polynomial reaction kinetics treated in Kotelenz (1988) and Blount (1994), and the two-species model discussed in Guiaş (2002).

In the second part of Chapter 3 we treat two models with a nonlinear diffusion mechanism. For the sake of simplicity, we restrict the discussion to single-species models in one space dimension without chemical reactions. We first investigate what happens when the intensity for a jump of a particle to a neighbouring cell depends on the local concentration, i.e.,  $d = d(u_l(z))$ , where  $d(\cdot)$  is monotonously increasing on  $\mathbb{R}_0^+$ . It turns out that, if  $D(\cdot)$  denotes the limit of  $\frac{1}{2}l^2d(\cdot)$ , the diffusive mass flux on the macroscopic level is given by

$$J = -(D'(u)u + D(u)) \partial_x u. \quad (10)$$

Thereafter, we have a look at the case when the intensity for a jump to a neighbouring cell depends on the absolute value of the discrete concentration gradient, i.e.,  $d = d(|\partial^+ u_l(z)|)$  for a jump to the right and  $d = d(|-\partial^- u_l(z)|)$  for a jump to the left, respectively. The resulting diffusive mass flux is

$$J = -D(|\partial_x u|) \partial_x u, \quad (11)$$

where  $D(\cdot)$  is the limit of  $l^2d(\cdot)$ .

The limit behaviour of the particle models with nonlinear diffusion mechanism has (to the best of our knowledge) not been investigated yet in the context of mesoscopic stochastic particle models. The first model ( $d = d(u_l(z))$ ) resembles the so-called zero-range process that is studied extensively in the literature on interacting particle systems where a macroscopic limit is obtained by rescaling the space and time variables (Kipnis & Landim, 1999). Finally, in the last section of Chapter 3, we demonstrate for a linear example model from Chapter

2 how the law of large numbers obtained there can be refined. Possible improvements and extensions of the results are discussed at the end of Chapter 3.

In Chapter 4 we turn to the more practical problem of simulating paths of the particle density process for the models considered in the previous chapters. The discussion applies, in particular, to the particle density process associated to the model for CO oxidation on Pt(110) that is introduced in Chapter 1. We propose a simulation algorithm which is then used for the simulation of the raindrop patterns described in Chapter 1.

More details about the contents are given in an overview paragraph at the beginning of each chapter. At the end of each chapter the results are discussed and related to other work.

## Remarks on notation and style

The language and level of mathematical rigour is adapted to the specific purposes of each chapter and thus varies slightly throughout the thesis. The ‘hard’ mathematics is contained in Chapters 2 and 3, while Chapters 1 and 4 are concerned with modelling and simulation. Overall, we have tried to use only standard notation, although this is practically impossible, since this is interdisciplinary work and the same letters often have different meanings in different fields. For example, it is customary in probability theory to reserve the letter  $\Omega$  for a sample space, while in analysis it often denotes a domain. We point out that we do not keep to the convention of denoting random variables and stochastic processes by capital letters. Vector-valued functions are always printed in bold face, whereas a point  $x = (x_1, \dots, x_m)$  in  $\mathbb{R}^m$  is printed in normal face. The letter  $C$  is reserved for a positive constant that can vary from line to line. Abstract vector spaces are generally assumed to be real. If  $X$  is a Hilbert space, then the scalar product in  $X$  is denoted by  $(\cdot, \cdot)_X$ . The dual pairing between a Banach space  $X$  and its dual  $X^*$  is denoted by  $\langle x^*, x \rangle_{X^*, X} = \langle x^*, x \rangle_X$ ,  $x^* \in X^*$ ,  $x \in X$ . A list of mathematical notation can be found on page iv.



# Chapter 1

## Experimental background and mathematical modelling

### Overview

In the first section of the present chapter we describe the experimental background to CO oxidation on Pt single crystal surfaces at low ( $\lesssim 10^{-4}$  mbar) to intermediate ( $\approx 10^{-2}$  mbar) pressures in the gas phase. We shall focus mainly on the Pt(110) surface and mention only briefly the surfaces with Miller indices (100) and (111). In addition, we indicate how some of the experimental results can be included in a mathematical model on the mesoscopic or macroscopic level. In Section 1.2 we develop a macroscopic deterministic PDE model for CO oxidation on Pt surfaces starting from a basic isothermal model that involves three species. The results of a numerical bifurcation analysis of the kinetic part for Pt(110) are discussed in Section 1.2.3. Subsequently the model is augmented by an equation for the (rescaled) surface temperature to capture also thermal effects. Finally, in Section 1.3, we introduce a corresponding mesoscopic stochastic particle model that accounts for fluctuations in the reaction processes.

### 1.1 Experimental background

On all platinum surfaces the CO oxidation reaction proceeds via a Langmuir-Hinshelwood mechanism: both CO and oxygen have to be adsorbed from the gas phase at adsorption sites on the platinum surface before reaction can occur. The adsorption of O<sub>2</sub> is dissociative, i.e., the oxygen molecule is split in two oxygen atoms, which constitutes the catalytic action of the surface. Adsorbed CO molecules are mobile: they perform a diffusive hopping from site to site. When a diffusing CO molecule encounters an oxygen atom both react to carbon dioxide CO<sub>2</sub> that desorbs immediately. The reaction is exothermal, the heat of reaction being about

20–30 kJ/mol. A basic reaction scheme reads



where  $\star$  denotes a vacant adsorption site.

### 1.1.1 Experimental methods

For a detailed understanding of the mechanism of any surface reaction one has to use experimental methods from surface physics and chemistry ('surface science'). In spite of the ambient gas phase, these methods allow to retrieve information about the processes taking place directly on the surface ('in situ'). In the following we shall briefly discuss some techniques that are important for the experimental study of CO oxidation on Pt. For more information see Zangwill (1988); Imbihl & Ertl (1995); Thomas & Thomas (1996); Rotermund (1997a); Oura et al. (2003). An essential prerequisite for many experimental techniques is vacuum technology (see, e.g., Oura et al. (2003)), at least for the preparation and purification of the catalytic surface. Therefore experiments are performed in an ultra-high vacuum (UHV) chamber in which a platinum probe of a typical size of 1 cm<sup>2</sup> is installed. The temperature of the probe can be regulated by a feedback-controlled heating mechanism. A gas inlet system allows to operate the UHV chamber as a *constant flow reactor*; the concentrations of CO and O<sub>2</sub> in the gas phase are held fixed to a good approximation, and the reaction product CO<sub>2</sub> is continuously removed.

A widely used method for determining surface structures and their changes is *low-energy electron diffraction* (LEED), where incoming electrons are scattered from the outer surface layers. The diffraction pattern provides information about the surface structure as well as evidence for superlattices. Because of the short mean free path of the electrons in the gas phase, the method can only be applied at low pressures ( $\lesssim 10^{-4}$  mbar). To get an idea about the global adsorbate coverages one can measure the work function of the probe and compare it to the work function of an empty surface. This method has been applied, e.g., for the study of global concentration oscillations on Pt(110) (Eiswirth, 1987). The development of the *photoemission electron microscope* (PEEM) permitted the visualisation of spatio-temporal concentration patterns. The spatial resolution of the PEEM is about 0.2  $\mu\text{m}$  and its temporal resolution is about 20 ms. In the PEEM images oxygen-covered areas appear dark due to the high work function of the O-covered surface; a CO-covered area is characterised by a light grey, and the empty parts of the surface appear bright. Measurements with atomic resolution can be performed with the *scanning tunnelling microscope* (STM), but because of its yet limited temporal resolution it cannot be applied to CO oxidation on Pt (at temperatures where pattern formation occurs). Experimental results from the atomic to the mesoscopic

scale are available for the catalytic oxidation of hydrogen on Pt(111) (Sachs et al., 2001). Another possibility to visualise surface reactions with nearly atomic resolution is to follow the reaction on a field emitter tip. This method has been applied to study the effects of internal noise in CO oxidation on Pt(110) (Suchorski et al., 1999, 2001; Imbihl, 2005).

All methods mentioned so far make use of electrons to characterise the structure or coverage of the surface. Therefore they cannot be employed at higher pressures in the gas phase, since the mean free path of the electrons is then too short. More recently, several methods using electromagnetic radiation (either in the infrared or in the visible part of the spectrum) have been developed which allow to observe the surface even at atmospheric pressures (Rotermund et al., 1995; Rotermund, 1997a). One of them is *ellipsomicroscopy for surface imaging* (EMSI) which is based on the change of polarisation of an outgoing with respect to an incoming light beam.

Finally, it should be mentioned that all known experimental methods on the mesoscopic or macroscopic scale provide only qualitative information about the surface coverages. Exact quantitative information about the coverages is usually not available.

### 1.1.2 Reaction steps of CO oxidation on Pt

Here we discuss in more detail the reaction steps of CO oxidation on Pt outlined at the beginning of this section. At the same time we introduce some generalities about the mathematical modelling of surface reactions on the mesoscopic and the macroscopic level. Reasonable values for the parameters introduced below are specified in Section 1.2 for the Pt(110) surface at low to intermediate pressures in the gas phase. For further discussions and references we refer to Eiswirth (1987) and Krischer et al. (1992). For further information about the kinetics of surface reactions see, e.g., Zangwill (1988) or Oura et al. (2003). In the following we denote the CO coverage by  $u$ , the oxygen coverage by  $v$ , and the fraction of surface in the  $1 \times 1$  structural phase (see below) by  $w$ .

#### Surface reconstruction

Conceptually cutting through a platinum crystal under different angles yields different single crystal surfaces. The surface of polycrystalline platinum is composed of domains which are orientated like crystal planes with Miller indices (100), (110), or (111). It is a well-known fact that many solids reconstruct in the vicinity of their surface in order to reach an energetically more favourable structure. Investigations with LEED and other methods proved that this is also the case for platinum; Pt(100) reconstructs in a quasi-hexagonal (hex) structure, while reconstructed Pt(110) exhibits a  $(1 \times 2)$  missing-row structure. No reconstruction is observed for Pt(111).

The surface reconstruction can be lifted by certain adsorbates; in particular, an elevated CO coverage leads to a lifting of the reconstruction of both the (110) and the (100) surface.

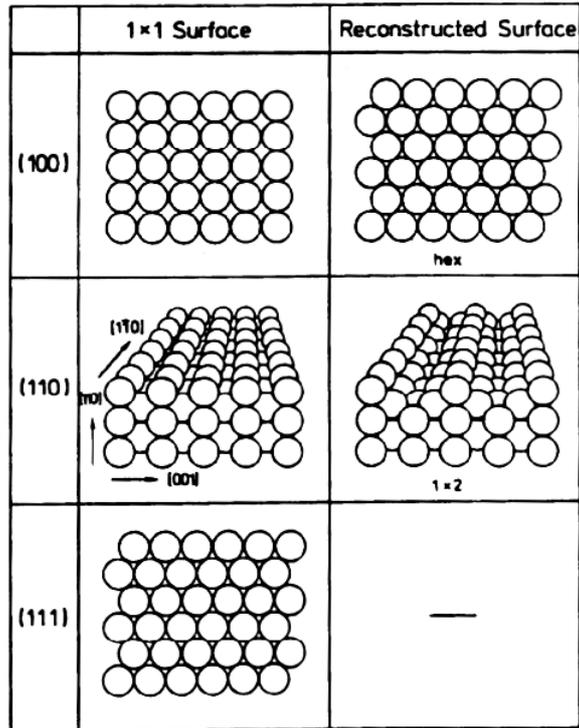


Figure 1.1: Reconstructed and  $1 \times 1$  surface structures of Pt(100), Pt(110) and Pt(111) (after Imbihl & Ertl (1995)).

The lifting of the reconstruction constitutes a phase transition. It starts at a certain critical coverage and the fraction of reconstructed surface then either decreases monotonously (Pt(110)) or abruptly jumps to a lower value (Pt(100)). The lifting of the reconstruction is reversible and in the case of Pt(100) hysteresis is observed. The critical coverage for Pt(110) (at a temperature of about 500 K) amounts to about 0.2 monolayers and at a coverage of 0.5 monolayers the reconstruction has completely been lifted. (A monolayer is defined as the number of atoms in the outermost layer.)

For the modelling of the reaction dynamics of CO oxidation on Pt the rate of change of the surface structure in the presence of changing adsorbate coverages is of utmost importance. It is a priori not clear how these structural changes should be included in a mathematical model.

### Adsorption

The probability that a collision of a molecule of a certain species with an empty surface leads to an adsorption event is called the sticking probability or sticking coefficient for the species. If adsorbates are already present on the surface, the probability of adsorption is modelled as a product of the sticking coefficient and an inhibition factor which is a function of the coverages. The sticking probability and the inhibition factor can, if only one species

is present in the gas phase, be determined by *thermal desorption spectroscopy* (Oura et al., 2003). However, the form of the inhibition factor might change if there are multiple species present in the gas phase. For CO oxidation on Pt it was found that the sticking coefficient of oxygen depends on the surface structure, while the sticking coefficient of CO is always approximately equal to one independent of the phase. The oxygen sticking coefficient has a value of 0.3 – 0.4 on the reconstructed ( $1 \times 2$ ) phase of Pt(110) and a value of 0.6 on the  $1 \times 1$  phase. The inhibition factor of CO has been measured to be  $1 - u^\xi$  for Pt(110), where  $\xi \approx 3.5$ . The exponent  $\xi$  is due to a precursor effect (see, e.g., Oura et al. (2003)). In the modelling below in Sections 1.2 and 1.3 this precursor effect is considered as a minor detail and neglected. The inhibition factor for oxygen adsorption is  $(1 - v)^2$  because an incoming oxygen molecule needs two neighbouring vacant sites for adsorption.

### Desorption

Desorption is a thermally activated process. In the absence of interactions between the adparticles the desorption rate is modelled as a product of an Arrhenius rate and a power of the coverage. The Arrhenius rate  $\nu \exp(-E/(RT))$ , where  $\nu$  is a frequency,  $E$  an activation energy and  $R$  the gas constant, contains the temperature dependence of the desorption process. It has been found experimentally that the rate of CO desorption on Pt(110) is to a good approximation proportional to the CO coverage. In the case of Pt(100) desorption of CO from the  $1 \times 1$  and the reconstructed surface must be distinguished because of the considerably higher binding energy on the reconstructed surface. The desorption of oxygen can be neglected below temperatures of about 700 K.

### Surface diffusion

On a microscopic level the diffusive hopping from site to site of an adsorbed particle is described by a random walk on a two-dimensional lattice. By rescaling the distance  $\lambda$  between two sites and the time  $\tau$  between two jumps in such a way that  $D = \lambda^2/(4\tau)$  is kept constant, one gets a Brownian motion process in the limit  $\lambda, \tau \rightarrow 0$ . The probability density to find a Brownian particle at time  $t$  at a point  $x$  in space satisfies a diffusion equation (heat equation) with diffusion coefficient  $D$  (Fick's law). Consequently, this equation also describes the particle density of a system of independently moving particles in the limit of large particle numbers. It can be shown that the same limit equation holds if the molecules are modelled as 'hard spheres', i.e., two molecules are not allowed to occupy the same site at the same time. (Such a particle system is called a *simple exclusion process* ; see, e.g., Kipnis & Landim (1999).) Therefore surface diffusion is often modelled on the macroscopic level by Fick's law. Other types of interaction between adparticles can be included by introducing a coverage dependence of the diffusion coefficient.

Simple Fickian diffusion has turned out to work quite well for CO on Pt(110). The diffusion coefficient at temperatures around 500 K is about  $10^{-13} - 10^{-12}$  m<sup>2</sup>/s. The diffusion of CO molecules is slightly anisotropic; it is faster along the [110]-oriented troughs (cf. Fig. 1.1). The diffusion of adsorbed oxygen is negligible.

## Reaction

The reaction of CO(ad) and O(ad) to CO<sub>2</sub>(g) is thermally activated. The rate is usually modelled in a first approximation as a product of an Arrhenius rate and the concentrations (mass action kinetics), although it certainly depends on the coverages in a more complicated way. (Cf. the discussion in Krischer et al. (1992).)

### 1.1.3 Spatio-temporal pattern formation

While Pt(111) exhibits merely bistable behaviour, a plethora of oscillations and patterns can be observed during CO oxidation on Pt(100) and, in particular, Pt(110). In the following we give an overview of the phenomenology of patterns on Pt(110).

#### Global oscillations

Global oscillations of the surface coverages have been investigated extensively by Eiswirth (1987) via work function measurements. Roughly speaking, the work function corresponds to the oxygen coverage because adsorbed oxygen contributes most. Oscillations have been observed for partial pressures  $p_{\text{CO}}$  and  $p_{\text{O}_2}$  from  $10^{-6}$  to  $10^{-3}$  mbar. Their phenomenology ranges from slow sinusoidal oscillations with large amplitude (Fig. 1.2) and relaxation oscillations involving two different time scales (Fig. 1.3) to fast sinusoidal oscillations with small amplitude. The fast sinusoidal oscillations sometimes give rise to a regime of chaotic dynamics via a series of period doubling bifurcations when decreasing the CO partial pressure  $p_{\text{CO}}$  (Fig. 1.4).

#### Spatio-temporal pattern formation at low pressures

With the development of the PEEM the observation of spatio-temporal pattern formation became possible. At pressures of the same order of magnitude as the global oscillations one can observe also spatio-temporal concentration patterns such as fronts, spirals, target patterns and standing waves (Ertl, 1991; Jakubith et al., 1990). An experimental bifurcation diagram with oxygen pressure fixed to  $p_{\text{O}_2} = 4 \times 10^{-4}$  mbar is reproduced in Fig. 1.5.

- If the CO partial pressure  $p_{\text{CO}}$  is increased starting from an O-covered surface, one first enters a bistable parameter range which is characterised by the spreading of CO islands on O-covered areas (Fig. 1.6a). Such islands nucleate at defects and may, at slightly higher  $p_{\text{CO}}$ , coexist with oxygen islands on CO-covered areas.

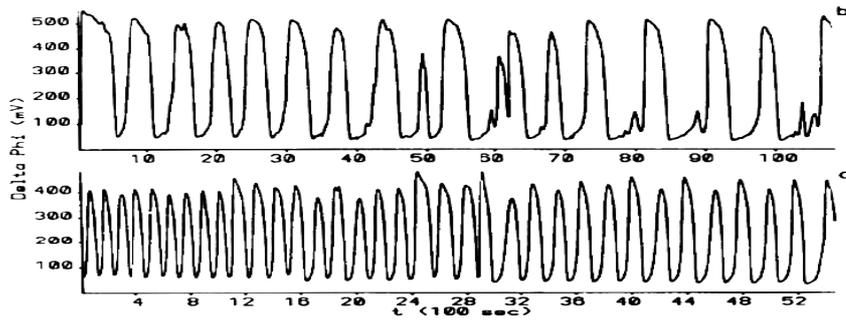


Figure 1.2: Slow sinusoidal oscillations with large amplitude at low temperatures. (a)  $p_{O_2} = 1.5 \times 10^{-4}$  mbar,  $p_{CO} = 1.125 \times 10^{-5}$  mbar,  $T = 460$  K. (b)  $p_{O_2} = 1.5 \times 10^{-4}$  mbar,  $p_{CO} = 0.975 \times 10^{-5}$  mbar,  $T = 470$  K. The work function  $\Delta\phi$  is plotted versus time (from Eiswirth (1987)).

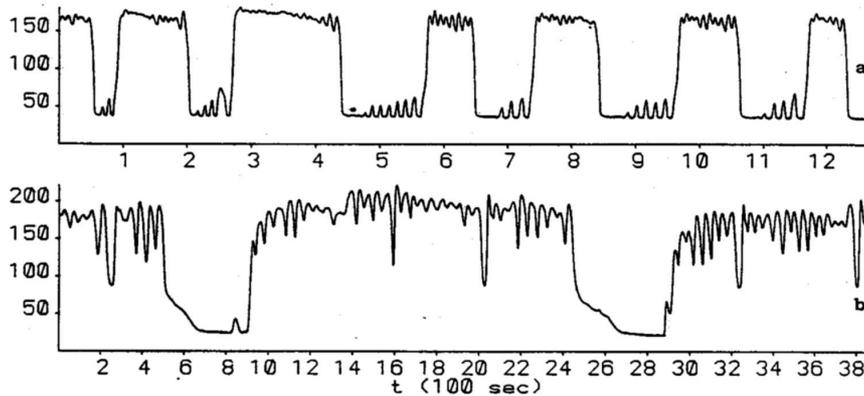


Figure 1.3: Relaxation oscillations at (a)  $p_{O_2} = 1.5 \times 10^{-4}$  mbar,  $p_{CO} = 4.5 \times 10^{-5}$  mbar,  $T = 520$  K and (b)  $p_{O_2} = 1.5 \times 10^{-5}$  mbar,  $p_{CO} = 0.6 \times 10^{-5}$  mbar,  $T = 500$  K (from Eiswirth (1987)).

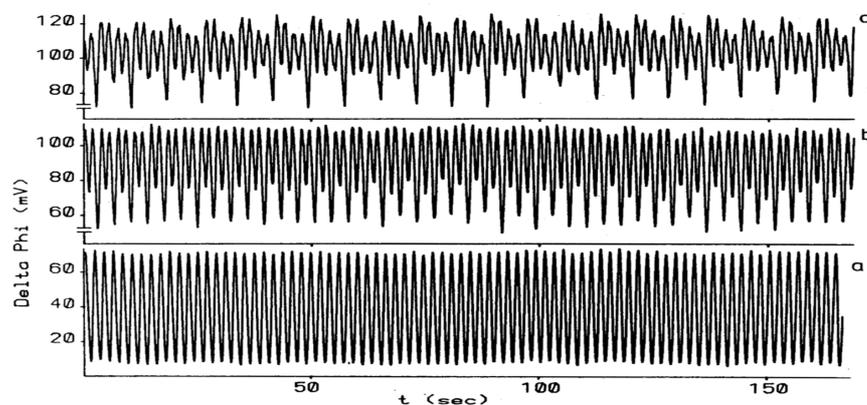


Figure 1.4: Fast sinusoidal oscillations with small amplitude and period doubling at  $p_{O_2} = 6 \times 10^{-5}$  mbar,  $T = 540$  K and (a)  $p_{CO} = 3.3 \times 10^{-5}$  mbar through (c)  $p_{CO} = 3.15 \times 10^{-5}$  mbar (from Eiswirth (1987)).

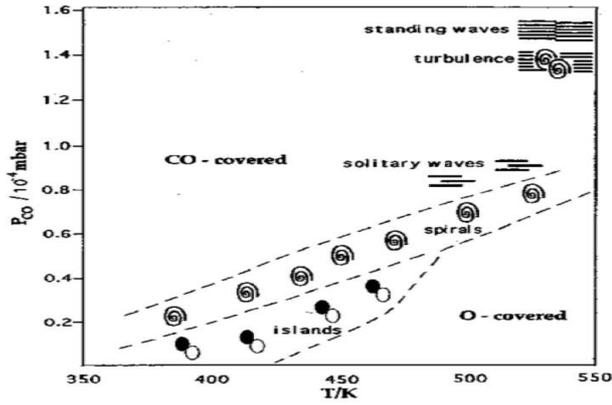


Figure 1.5: Experimental bifurcation diagram in the temperature  $T$  and the CO partial pressure  $p_{\text{CO}}$  with oxygen pressure fixed at  $p_{\text{O}_2} = 4 \times 10^{-4}$  mbar (after (Rotermund, 1997a)).

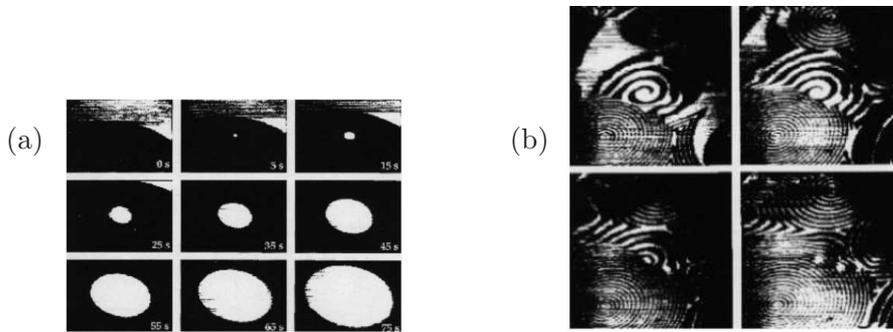


Figure 1.6: (a) Nucleation and growth of a CO island at  $p_{\text{O}_2} = 4 \times 10^{-4}$  mbar,  $T = 443$  K. The width of the depicted area is about  $200 \mu\text{m}$ . The time lapse between the pictures is 10 s. (b) PEEM snapshots of spiral waves at  $p_{\text{O}_2} = 4 \times 10^{-4}$  mbar,  $p_{\text{CO}} = 4.3 \times 10^{-5}$  mbar,  $T = 448$  K. The core of the spiral with the largest wavelength has a size of  $25 \times 14 \mu\text{m}^2$ . The time interval between the snapshots is 30 s (after (Rotermund, 1997a)).

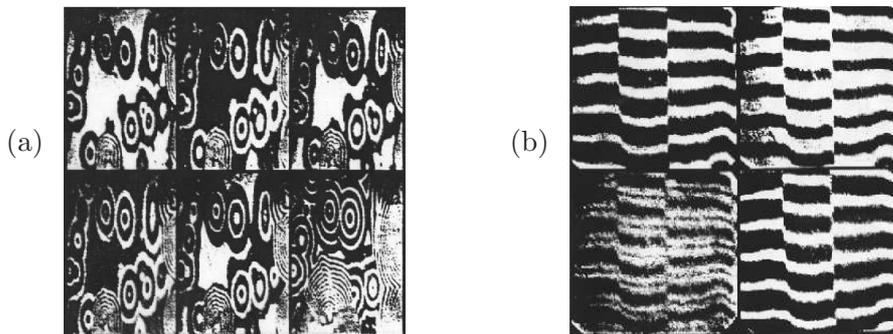


Figure 1.7: (a) Target patterns on an area of  $200 \times 300 \mu\text{m}^2$  at  $p_{\text{O}_2} = 3.2 \times 10^{-4}$  mbar,  $p_{\text{CO}} = 3 \times 10^{-5}$  mbar,  $T = 427$  K. The time lapse between the first five images is 4.1 s and 30 s between the two last images. (b) Standing waves with a period of 1.4 s on an area of  $300 \times 300 \mu\text{m}^2$  at  $p_{\text{O}_2} = 4.1 \times 10^{-4}$  mbar,  $p_{\text{CO}} = 1.75 \times 10^{-5}$  mbar,  $T = 550$  K, at time  $t = 0$  s, 0.08 s, 0.12 s, 0.46 s (after Jakubith et al. (1990)).

- ▶ Upon further increasing the CO pressure one enters a parameter region which is characterised by the existence of pulses and spirals (Figs. 1.6b and 1.7a), in other words, the system can in this parameter region be regarded as an *excitable medium* (see, e.g, Mikhailov (1994)).
- ▶ At even higher  $p_{\text{CO}}$  there is an oscillatory region, where mainly global fast sinusoidal oscillations with small amplitude are observed, but occasionally also standing wave patterns (stripes oscillating with opposite phase, Fig. 1.7b) and turbulent behaviour.

### Fluctuation-induced pattern formation at intermediate pressures

With increasing pressure in the gas phase smaller and smaller patches of the surface can be regarded as well mixed, the size of critical nuclei for the formation of islands and pulses is expected to decrease. Thus internal fluctuations due to the discrete nature of the reaction processes may become relevant. (See also the discussion in Section 1.3.)

An experimental observation at intermediate pressures, so-called raindrop patterns, which supports this reasoning is reproduced in Figs. 1.8 and 1.9. Starting from a reactive surface with a relatively high oxygen coverage, the CO partial pressure  $p_{\text{CO}}$  had been stepwise increased to a value just before the whole surface would switch to the CO-covered state. CO nuclei could be seen to originate at various places, forming a ring-shaped pattern that was subsequently destroyed (propagation failure). Their appearance seemed to be randomly distributed all over the catalyst surface (Rotermund, 1997b).

## 1.2 Macroscopic deterministic modelling

In this section we introduce a unified macroscopic deterministic model for CO oxidation on platinum single crystal surfaces at low to intermediate pressures. A complementary mesoscopic stochastic particle model will be presented in Section 1.3. Concrete parameter values are specified for Pt(110) at low pressures. The present work is a continuation of Reichert (2000) and Reichert et al. (2001), where we focused on the isothermal and spatially homogeneous case. Previous models based on the same mechanism that incorporate the surface structural phase transition can be found in Imbihl et al. (1985) and Andrade et al. (1989) for Pt(100), Krischer et al. (1992) for Pt(110), and Krömker (1997) for both surfaces. A summary of what follows has already been given in Starke et al. (2006).

### 1.2.1 Chemical reactions and surface diffusion

#### The reaction steps

In view of the experimental findings discussed in the previous section, the reaction processes on the surface should be described in terms of two species that can be adsorbed on either of

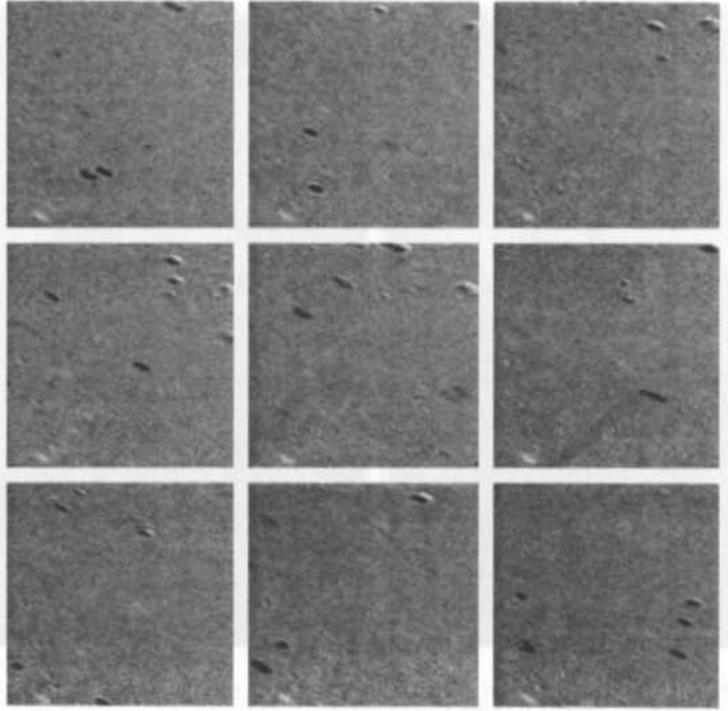


Figure 1.8: EMSI pictures of raindrops at  $p_{\text{O}_2} = 2.2 \times 10^{-2}$  mbar,  $p_{\text{CO}} = 4.9 \times 10^{-3}$  mbar,  $T = 530$  K. The time lapse between each snapshot is 0.4 s. The size of the depicted area is  $1.1 \times 0.9 \text{ mm}^2$  (Rotermund, 1997a).

(a)

(b)

Figure 1.9: (a) Part of a Pt(110) surface exhibiting raindrop patterns recorded with EMSI at  $p_{\text{CO}} = 7 \times 10^{-3}$  mbar and  $p_{\text{O}_2} = 2.2 \times 10^{-2}$  mbar. The time interval between the snapshots is 160 ms, and the depicted area is  $100 \times 70 \text{ }\mu\text{m}^2$ . (b) Space-time diagram of the raindrop ( $1.6 \text{ s} \times 100 \text{ }\mu\text{m}$ , Rotermund (1997b)).

two structural phases of the platinum surface. Therefore one should in principle work with six variables: two for the phases, and four for the two species adsorbed on each phase. One phase variable can be eliminated immediately because the sum of the fraction of both phases must be equal to one. Moreover, for our model we make the simplifying assumption that locally CO and oxygen are equally distributed over both phases. This assumption is certainly justified for Pt(110), since the mobility of CO and oxygen is about the same on each of the

(1)	Adsorption of CO	$\text{CO(g)} \rightarrow \text{CO(ad)}$
(2)	Desorption of CO	$\text{CO(ad)} \rightarrow \text{CO(g)}$
(3)	Reaction of CO(ad) and O(ad)	$\text{CO(ad)} + \text{O(ad)} \rightarrow \text{CO}_2\text{(g)}$
(4)	Adsorption of O <sub>2</sub>	$\text{O}_2\text{(g)} \rightarrow 2 \text{O(ad)}$
(5)	Reverse structural phase transition	$\star(\text{rec}) \rightarrow \star(1 \times 1)$
(6)	Structural phase transition	$\star(1 \times 1) \rightarrow \star(\text{rec})$

Table 1.1: Reaction steps of CO oxidation on Pt at low to intermediate pressures

two phases.

All relevant reaction steps for CO oxidation on Pt, as discussed in the previous section, are listed in Table 1.1. All steps proceed with rates (or velocities)  $K_i$  that are supposed to depend on the local concentrations  $u = [\text{CO(ad)}]$ ,  $v = [\text{O(ad)}]$ ,  $w = [\star(1 \times 1)]$ , and the local surface temperature  $T$ . Moreover, they depend on a set of control parameters  $\Lambda$  and a large number of system parameters that cannot be changed during experiments. The reaction rate  $K_i$  measures how often a certain reaction occurs per unit time. It should not be confused with the reaction rate constant. Instead of absolute temperature  $T$  we shall mostly work with a rescaled temperature  $\theta = (T - \bar{T})/\bar{T}$ , where  $\bar{T}$  is a reference temperature (the equilibrium temperature of an empty surface). The control parameters are thus the partial pressures of CO and O<sub>2</sub> in the gas phase denoted in the following by  $p_u$  and  $p_v$ , respectively, and the reference temperature  $\bar{T}$ , i.e.,  $\Lambda = (p_u, p_v, \bar{T})$ . System parameters that cannot be varied in experiments are, e.g., the sticking coefficients, the coefficient of heat conduction, etc.

### The rates

In the following we briefly describe the rate terms we use to model the reaction steps listed in Table 1.1. The vector of concentrations is denoted by  $\mathbf{u} = (u, v, w)$ , and  $x = (x_1, x_2)$  denotes a point in two-dimensional space. Note that we always use rescaled concentrations so that  $\rho_s u(x) dx_1 dx_2$  is either the number of CO molecules in the infinitesimal surface area  $dx_1 dx_2$  or the number of moles of CO molecules, depending on whether  $\rho_s$ , the density of surface atoms, is given in atoms/m<sup>2</sup> or mol/m<sup>2</sup>. Since we work with a rescaled temperature, Arrhenius rates  $k(T) = \nu \exp(-E/(RT))$  (where  $\nu$  is a frequency,  $E$  an activation energy and  $R$  the gas constant) have to be transformed to  $k(\theta, \bar{T})$  in the following way:

$$k(T) = \nu \exp\left(-\frac{E}{RT}\right) = \nu \exp\left(-\frac{E}{R\bar{T}}\right) \exp\left(\frac{E}{R\bar{T}} \frac{\theta}{1+\theta}\right) = k(\theta, \bar{T}). \quad (1.2)$$

Reaction steps (1), (2), and (4) from Table 1.1 describe adsorption and desorption of CO and oxygen. The rates are defined in Table 1.2.

If the adsorption rate of CO is taken to be proportional to  $1 - u$  as here (or  $1 - u^\xi$  as in Krischer et al. (1992)) independent of the oxygen coverage, one implicitly drops the

(1) **CO adsorption:**

$$K_1(\mathbf{u}, \theta, \Lambda) = k_1 p_u s_u (1 - u)$$

$k_1$	impingement rate of CO	$3.135 \times 10^5 \text{ mbar}^{-1} \text{ s}^{-1}$
$p_u$	CO partial pressure in the gas phase	[mbar]
$s_u$	sticking coefficient of CO	1.0

(2) **CO desorption:**

$$K_2(\mathbf{u}, \theta, \Lambda) = k_2(\theta, \bar{T}) u$$

$k_2$	Arrhenius rate	
$\nu_2$	frequency	$5 \times 10^{15} \text{ s}^{-1}$
$E_2$	energy	135 kJ/mol

(4) **Adsorption of O<sub>2</sub>:**

$$K_4(\mathbf{u}, \theta, \Lambda) = k_4 p_v (s_v w + \tilde{s}_v (1 - w)) ((1 - u)(1 - v))^2$$

$k_4$	impingement rate of O <sub>2</sub>	$2.929 \times 10^5 \text{ mbar}^{-1} \text{ s}^{-1}$
$s_v$	sticking coefficient of O <sub>2</sub> on (1 × 1) phase	0.6
$\tilde{s}_v$	sticking coefficient of O <sub>2</sub> on (rec) phase	0.3

Table 1.2: Rate terms for adsorption and desorption of CO and oxygen (steps (1), (2), and (4) from Table 1.1.)

conservation constraint imposed by a strict Langmuir-Hinshelwood mechanism on the number of adsorbed oxygen atoms, adsorbed CO molecules and vacant sites. On the other hand, it has been observed experimentally that the presence of oxygen on the surface does not noticeably influence the adsorption of CO (Imbihl & Ertl, 1995; Rotermund, 1997a). Consequently, we assume that CO molecules may be adsorbed both at empty and at O-covered sites, whereas the dissociative adsorption of an oxygen molecule can only take place at two neighbouring empty sites. The probability that two adjacent sites are occupied neither by CO nor by O is then  $((1 - u)(1 - v))^2$  instead of  $(1 - u - v)^2$ , the inhibition term used in Krischer et al. (1992). Since the former term is larger only by  $uv$ , the difference is negligible if  $u$  or  $v$  is small, which is typically the case at low pressures. At intermediate pressures the new term is in better agreement with experiments (Reichert et al., 2001). Note that the sticking coefficient of oxygen on the  $1 \times 1$  phase is considerably higher than on the reconstructed phase, which can lead to oscillatory, doubly metastable, and excitable behaviour of the reaction kinetics. The rate of desorption of CO molecules from the Pt(110) surface is modelled by a product of the CO concentration and an Arrhenius rate. In the case of Pt(100) the desorption of CO molecules from the reconstructed and the  $1 \times 1$  phase has to be distinguished because of the strong difference in the binding energies on the two phases.

**(3) Reaction:**

$$K_3(\mathbf{u}, \theta, \Lambda) = k_3(\theta, \bar{T}) uv$$

$k_3$	Arrhenius rate	
$\nu_3$	frequency	$5 \times 10^7 \text{ s}^{-1}$
$E_3$	energy	34 kJ/mol

Table 1.3: Rate term for reaction (step (3) from Table 1.1).

For the reaction rate we use (in lack of better knowledge) mass action kinetics, i.e., we assume that the reaction rate is given by the product of the concentrations multiplied with an Arrhenius rate (cf. Table 1.3).

A major change, as compared to previous work, concerns the rate term for the structural phase transition. In Krischer et al. (1992) it was assumed that, in the absence of oxygen in the gas phase, the equilibrium fraction of  $1 \times 1$  surface  $\bar{w}$  can be expressed as a function of the equilibrium CO coverage  $\bar{u}$ , i.e.,  $\bar{w} = f(\bar{u})$ , which is true for Pt(110). In the nonequilibrium case, the dynamics of  $w$  was modelled by  $\dot{w} \propto f(u) - w$ . However, this relaxation ansatz cannot be carried over to Pt(100) because hysteresis in the phase transition is observed experimentally, and thus  $\bar{w}$  cannot be expressed as a function of  $\bar{u}$ . This can be remedied by assuming  $\bar{u} = g(\bar{w})$  and taking  $\dot{w} \propto u - g(w)$  as in Krömker (1997), but such an approach is hard to justify in physical terms. Our ansatz is motivated by the fact that the phase transition proceeds via nucleation and growth. The probability for nucleation is determined solely by the CO coverage, but the growth of a phase is to some extent autocatalytic, which leads to a dependence of the rate of growth on the concentration of the phase itself. Therefore the growth rate of  $1 \times 1$  phase on a reconstructed surface, for instance, is chosen proportional to a weighted sum of  $u$  and  $w$ , each to some power  $\varepsilon$ . A highly nonlinear dependence of this rate on the local CO coverage has been observed experimentally by Hopkinson et al. (1993). They measured the growth rate of  $1 \times 1$  phase on a hex-R reconstructed Pt(100) surface to depend on the CO concentration on the hex-R phase to a power of about 4.5. It is plausible that the exponent for  $w$  should have approximately the same value. The reverse transition is modelled in an analogous way. In order not to introduce too many parameters, the same weight and exponent are used; the rate constants, however, are allowed to be different. The rate terms and parameters are defined in Table 1.4. Two additional important constants are listed in Table 1.5 for convenience.

**Surface diffusion of CO**

We assume simple Fickian diffusion of the CO molecules with a diffusion coefficient  $D_u$  of about  $10^{-12} \text{ m}^2/\text{s}$ . In other words, the diffusive flux of CO molecules is given by  $\mathbf{J}_u = -D_u \nabla u$ .

(5) **Reverse structural phase transition:**

$$K_5(\mathbf{u}, \theta, \Lambda) = k_5(\theta, \bar{T})((1 - \alpha)u^\varepsilon + \alpha w^\varepsilon) (1 - w)$$

$k_5$	Arrhenius	
$\nu_5$	frequency	$10^3 \text{ s}^{-1}$
$E_5$	energy	29 kJ/mol
$\alpha$	weight	0.1
$\varepsilon$	exponent	4.0

(6) **Structural phase transition:**

$$K_6(\mathbf{u}, \theta, \Lambda) = k_6(\theta, \bar{T})((1 - \alpha)(1 - u)^\varepsilon + \alpha(1 - w)^\varepsilon) w$$

$k_6$	Arrhenius rate	
$\nu_6$	frequency	$2 \times 10^2 \text{ s}^{-1}$
$E_6$	energy	29 kJ/mol
$\alpha$	weight	0.1
$\varepsilon$	exponent	4.0

Table 1.4: Parameters for surface reconstruction and lifting of the reconstruction (steps (5) and (6) from Table 1.1).

$R$	gas constant	8.314 J/mol kg
$N_A$	Avogadro number	$6.0221 \times 10^{23}$

Table 1.5: Additional important constants.

### 1.2.2 Mass balance

A mass balance yields the following system of differential equations for CO oxidation on Pt(110):

$$\begin{aligned} \partial_t u + \nabla \cdot \mathbf{J}_u &= K_1 - K_2 - K_3 \\ \partial_t v &= 2K_4 - K_3 \\ \partial_t w &= K_5 - K_6, \end{aligned} \tag{1.3}$$

or, explicitly,

$$\begin{aligned} \partial_t u &= k_1 p_u s_u (1 - u) - k_2(\theta, \bar{T}) u - k_3(\theta, \bar{T}) uv + D_u \Delta u \\ \partial_t v &= 2k_4 p_v (s_v w + \tilde{s}_v (1 - w))((1 - u)(1 - v))^2 - k_3(\theta, \bar{T}) uv \\ \partial_t w &= k_5(\theta, \bar{T})((1 - \alpha)u^\varepsilon + \alpha w^\varepsilon) (1 - w) \\ &\quad - k_6(\theta, \bar{T})((1 - \alpha)(1 - u)^\varepsilon + \alpha(1 - w)^\varepsilon) w. \end{aligned} \tag{1.4}$$

Bifurcation	Co-dimension	Abbreviation
Hopf bifurcation (supercritical)	1	h
Hopf bifurcation (subcritical)	1	h'
saddle-node	1	sn
saddle loop	1	sl
homoclinic bifurcation		
saddle-node/infinite period	1	sniper
saddle-node on a loop		
saddle node of periodic orbits	1	snp
cuspl	2	C
Takens-Bogdanov bifurcation	2	TB
degenerate Hopf bifurcation	2	DH
saddle-node loop	2	SNL
neutral saddle loop	2	NSL
trace 0 saddle loop		
Takens-Bogdanov-cuspl	3	TBC

Table 1.6: Commonly used denotations of bifurcations and their abbreviations. By a degenerate Hopf bifurcation we mean the one described in Guckenheimer (1986).

### 1.2.3 Numerical bifurcation analysis

In the following we present some aspects of the bifurcation structure of the kinetic part of system (1.4) with the parameters specified above for Pt(110). For more details and additional computations for Pt(100) see Reichert (2000); Reichert et al. (2001). The computations were performed using algorithms from the AUTO 97 package by Doedel et al. (1998). Abbreviations of the bifurcations found are listed in Table 1.6; for details see, e.g., Hale & Koçak (1991); Wiggins (1990); Guckenheimer (1986).

For Pt(110), bifurcation diagrams in  $p_u$  and  $p_v$  have been computed for several fixed crystal temperatures. At higher temperatures (500–560 K) the bifurcation diagram is organised by a cuspl and two Takens-Bogdanov points. In total there are 12 parameter regions with different dynamical behaviour (cf. Fig. 1.10), but only regions 1–5 are physically relevant because the others are too small to be detected in experiments. In regions 1, 2, and 3 there is only one attractor, a stable node or an asymptotically stable periodic orbit, respectively. The maximal width of the oscillatory region 2 at 540 K amounts to about 10% of the value of  $p_u$  at the supercritical Hopf bifurcation. It decreases towards higher as well as lower temperatures. In 4 and 5 two attractors coexist; in 5 there are two stable nodes, while in 4 the system asymptotically approaches either a stable node or a small asymptotically stable periodic orbit. The maximal width of region 4 is about 1% of  $p_u$  at 540 K, so it could possi-

bly be detected experimentally. The regions with nontrivial dynamics, such as bistability or oscillations, move towards higher partial pressures as temperature is increased.

Phase portraits of the dynamics in the different parameter regions are sketched in Fig. 1.10. In the pictures containing three fixed points the lower one can always be identified with a reactive, mainly reconstructed surface with a relatively high oxygen coverage (reactive state) and the upper one with a predominantly CO-covered  $1 \times 1$  surface (poisoned state). From a physical point of view the model presented here yields almost the same results as the one proposed by Krischer et al. (1992). The most important distinction is that here the two curves of saddle-node bifurcations do, for the investigated parameter region, not merge in a second cusp when  $p_u$  and  $p_v$  are increased, rather bistability persists even at atmospheric pressures, in accordance with experiment. In fact, this is due to the change of the adsorption kinetics of oxygen, as was checked by repeating the computation of the saddle-node curves with the term that was used in the model of Krischer et al. (1992). As in Krischer et al. (1992), the dynamics seems to be essentially two-dimensional; period doubling transitions to chaos that have been observed experimentally (Eiswirth, 1987) could not be found.

#### 1.2.4 Heat production and transfer

In this section we derive an evolution equation for the surface temperature from an energy balance. To this end, we assume that below an infinitesimal surface area  $dx_1 dx_2$  around a point  $x$  on the surface the crystal has temperature  $T(x)$  in a cube of Volume  $l_T dx_1 dx_2$ , where  $l_T$  is a characteristic depth. Furthermore, we assume that the bulk below such a cube has temperature  $T_b[T]$ , i.e., the bulk temperature is a functional of the surface temperature field. This assumption is due to the fact that the bulk temperature is regulated by a feedback mechanism according to the surface temperature in order to heat or cool the surface. The ambient gas phase is assumed to have constant temperature  $T_g$ . The energy balance of the cube can be written ( $U$  denoting internal energy)

$$\frac{dU}{dt} = \frac{dQ_{\text{chem}}}{dt} + \frac{dQ_{\text{rad}}}{dt} + \frac{dQ_{\text{cond}}}{dt}. \quad (1.5)$$

Here  $dQ_{\text{chem}}$ ,  $dQ_{\text{rad}}$  and  $dQ_{\text{cond}}$  denote the changes of internal energy through production and transfer of heat by chemical reactions, radiation and heat conduction, respectively.

##### Heat production through chemical reactions

Heat is produced by the chemical processes with rate

$$\frac{dQ_{\text{chem}}}{dt} = \rho_s (H_1 K_1 - H_2 K_2 + H_3 K_3 + H_4 K_4) dx_1 dx_2. \quad (1.6)$$

The energies  $H_1 - H_4$  and  $\rho_s$  are defined in Table 1.7. The units of the  $H_i$  and  $\rho_s$  have to be chosen such that  $[\rho_s][H_i] = \text{J/m}^2$ . The energies gained by adsorption and lost by desorption

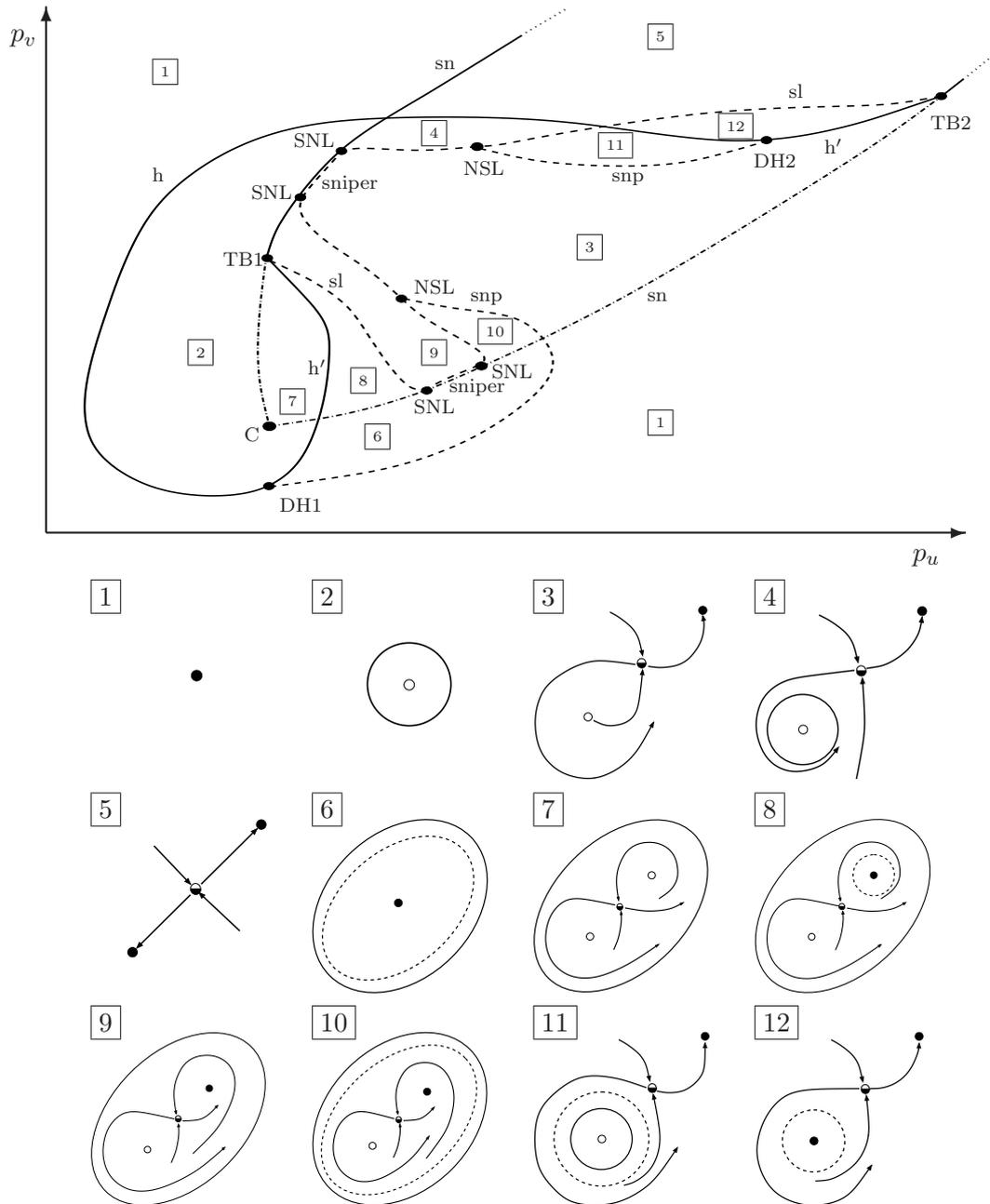


Figure 1.10: Sketches of the complete bifurcation diagram at 500–560 K (top) and phase portraits in different parameter regions (bottom) for Pt(110). Hopf bifurcations, and saddle-node bifurcations involving a stable node and a saddle with a one-dimensional unstable manifold are drawn with solid lines, saddle-node bifurcations involving a saddle with a one-dimensional and another with a two-dimensional unstable manifold are drawn with dash-dotted lines. The dashed curves indicate global bifurcations (cf. Table 1.6). In the phase portraits stable nodes are represented by filled circles, saddle points with a two-dimensional unstable manifold by empty circles, and saddle points with a one-dimensional unstable manifold by half-filled circles. Asymptotically stable periodic orbits are indicated by solid lines, unstable ones by dashed lines.

$H_1$	heat of adsorption of CO	$\approx 135$ kJ/mol
$H_2$	heat loss by desorption of CO	$\approx H_1$
$H_3$	heat of reaction	$\approx 20 - 30$ kJ/mol
$H_4$	heat of adsorption of an oxygen molecule	$\approx 230$ kJ/mol
$\rho_s$	density of surface atoms	$8.84 \times 10^{18}$ atoms/m <sup>2</sup>

Table 1.7: Parameters for heat production through chemical processes.

$\sigma$	Stefan-Boltzmann constant	$5.6705 \times 10^{-8}$ J/m <sup>2</sup> K s
$e_s$	integral emissivity of platinum	$\approx 0.05 - 0.1$
$e_g$	integral emissivity of gas	unknown
$a_s$	integral absorptivity of platinum	unknown

Table 1.8: Parameters for heat transfer by radiation.

are assumed to cancel each other approximately; the heat of reaction can be estimated by thermodynamic considerations. Note that in the case of Pt(100) we would have to distinguish between desorption from the  $1 \times 1$  and the reconstructed phase. The density  $\rho_s$  has been calculated with a grid constant  $a \approx 0.4$  nm and an average area occupied by a surface atom of  $(\sqrt{2}a^2)/2$ . If we assume for simplicity that each atom occupies a square, the ‘effective’ grid constant is 0.34 nm.

### Heat transfer by radiation

In order to calculate the contribution of radiation to the heat balance we assume that, in a first approximation, it is allowed to use integral emissivities and absorptivities, i.e., an integral version of Planck’s formula. Hence,

$$\frac{dQ_{\text{rad}}}{dt} = \sigma (a_s(T) e_g(T_g) T_g^4 - e_s(T) T^4) dx_1 dx_2. \quad (1.7)$$

(See Table 1.8 for the definition of the coefficients  $\sigma$ ,  $a_s(T)$ ,  $e_s(T)$  and  $e_g(T_g)$ ). It will turn out that in a first approximation we do not need to know the values of  $e_g$  and  $a_s$ .

### Heat conduction

Heat can be transferred from a cube below the surface to the bulk and to adjacent cubes parallel to the surface:

$$\frac{dQ_{\text{cond}}}{dt} = \lambda (l_T (\partial_{x_1}^2 T + \partial_{x_2}^2 T) - (T - T_b[T])/l_T) dx_1 dx_2, \quad (1.8)$$

where  $\lambda$  and  $l_T$  are defined in Table 1.9. The notation  $T_b[T]$  indicates that we think of the bulk temperature as a functional of the surface temperature because, in principle, it may be

$\lambda$	heat conduction coefficient of Pt	72 W/m K
$l_T$	characteristic thickness of surface layer	$\approx 5 \times 10^{-5}$ m

Table 1.9: Parameters for heat conduction.

$C$	specific heat capacity of Pt	130 J/kg K
$\rho_b$	bulk mass density of Pt	$21.09 \times 10^3$ kg/m <sup>3</sup>

Table 1.10: Additional parameters for the energy balance.

regulated according to the full detailed surface temperature field. In practice, of course, it depends only on an averaged surface temperature.

### 1.2.5 Energy balance

Gathering together the different contributions yields

$$\begin{aligned} \frac{dU}{dt} &= \rho_s (H_1 K_1 - H_2 K_2 + H_3 K_3 + H_4 K_4) dx_1 dx_2 \\ &\quad + \sigma (a_s(T) e_g(T_g) T_g^4 - e_s(T) T^4) dx_1 dx_2 \\ &\quad + \lambda \left( l_T (\partial_{x_1}^2 T + \partial_{x_2}^2 T) - (T - T_b[T])/l_T \right) dx_1 dx_2. \end{aligned} \quad (1.9)$$

Substituting

$$\frac{dU}{dt} = \frac{\partial T}{\partial t} C \rho_b l_T dx_1 dx_2 \quad (1.10)$$

in the equation above and dividing by  $C \rho_b l_T dx_1 dx_2$  yields an evolution equation for the surface temperature field  $T(x)$ . The parameters  $C$  and  $\rho_b$  are defined in Table 1.10. We assume now that the equilibrium temperature field without chemical reactions is given by  $T \equiv \bar{T}$ . We linearise the non-chemical contributions around  $T \equiv \bar{T}$ , assuming that the emissivity  $e_s$  and the absorptivity  $a_s$  depend only weakly on  $T$ . We have, up to first order,

$$T - T_b[T] = T - \bar{T} + \bar{T} - T_b[T \equiv \bar{T}] - \int_G \left. \frac{\delta T_b}{\delta T} \right|_{T \equiv \bar{T}} (T - \bar{T}) dx_1 dx_2. \quad (1.11)$$

Hence,

$$\begin{aligned} \frac{\partial(T - \bar{T})}{\partial t} &= -\frac{4\sigma e_s(\bar{T}) \bar{T}^3}{C \rho_b l_T} (T - \bar{T}) + \frac{\lambda}{C \rho_b} \Delta (T - \bar{T}) \\ &\quad - \frac{\lambda}{C \rho_b l_T^2} (T - \bar{T}) + \frac{\lambda}{C \rho_b l_T} \int_G \left. \frac{\delta T_b}{\delta T} \right|_{T \equiv \bar{T}} (T - \bar{T}) dx_1 dx_2 + \text{h.o.t.}, \end{aligned} \quad (1.12)$$

where  $\Delta = \partial_{x_1}^2 + \partial_{x_2}^2$ , and the two-dimensional domain representing the surface has been denoted by  $G$ . Thus the equation for the rescaled temperature  $\theta = (T - \bar{T})/\bar{T}$  reads (reincluding

the chemical terms)

$$\begin{aligned} \frac{\partial \theta}{\partial t} = & \frac{\rho_s}{C \rho_b l_T \bar{T}} (H_1 K_1 - H_2 K_2 + H_3 K_3 + H_4 K_4) \\ & - \left( \frac{4 \sigma e_s(\bar{T}) \bar{T}^3}{C \rho_b l_T} + \frac{\lambda}{C \rho_b l_T^2} \right) \theta + D_\theta \Delta \theta + \frac{D_\theta}{l_T} \int_G \frac{\delta T_b}{\delta T} \Big|_{T=\bar{T}} \theta \, dx_1 \, dx_2, \end{aligned} \quad (1.13)$$

where  $D_\theta = \lambda / (C \rho_b)$ . Next, we define

$$\beta(\bar{T}) = \frac{\rho_s}{C \rho_b l_T \bar{T}}. \quad (1.14)$$

Note that the units have to be chosen such that  $[\beta] = [H_i]^{-1}$ . Furthermore we set

$$\gamma(\bar{T}) = \frac{4 \sigma e_s(\bar{T}) \bar{T}^3}{C \rho_b l_T} + \frac{\lambda}{C \rho_b l_T^2}. \quad (1.15)$$

Neglecting the integral term, we finally get the following equation for the rescaled surface temperature:

$$\partial_t \theta = \beta(\bar{T}) (H_1 K_1 - H_2 K_2 + H_3 K_3 + H_4 K_4) - \gamma(\bar{T}) \theta + D_\theta \Delta \theta, \quad (1.16)$$

or, explicitly,

$$\begin{aligned} \partial_t \theta = & \beta(\bar{T}) \left( H_1 k_1 p_u s_u (1 - u) - H_2 k_2(\theta, \bar{T}) u + H_3 k_3(\theta, \bar{T}) uv \right. \\ & \left. + H_4 k_4 p_v (s_v w + \tilde{s}_v (1 - w)) ((1 - u)(1 - v))^2 \right) \\ & - \gamma(\bar{T}) \theta + D_\theta \Delta \theta. \end{aligned} \quad (1.17)$$

### 1.2.6 Summary

Finally, the full system of reaction-diffusion equations coupled to an equation for the surface temperature field reads

$$\partial_t u = k_1 p_u s_u (1 - u) - k_2(\theta, \bar{T}) u - k_3(\theta, \bar{T}) uv + D_u \Delta u \quad (1.18a)$$

$$\partial_t v = 2 k_4 p_v (s_v w + \tilde{s}_v (1 - w)) ((1 - u)(1 - v))^2 - k_3(\theta, \bar{T}) uv \quad (1.18b)$$

$$\begin{aligned} \partial_t w = & k_5(\theta, \bar{T}) ((1 - \alpha) u^\varepsilon + \alpha w^\varepsilon) (1 - w) \\ & - k_6(\theta, \bar{T}) ((1 - \alpha)(1 - u)^\varepsilon + \alpha(1 - w)^\varepsilon) w \end{aligned} \quad (1.18c)$$

$$\begin{aligned} \partial_t \theta = & \beta(\bar{T}) \left( H_1 k_1 p_u s_u (1 - u) - H_2 k_2(\theta, \bar{T}) u + H_3 k_3(\theta, \bar{T}) uv \right. \\ & \left. + H_4 2 k_4 p_v (s_v w + \tilde{s}_v (1 - w)) ((1 - u)(1 - v))^2 \right) \\ & - \gamma(\bar{T}) \theta + D_\theta \Delta \theta. \end{aligned} \quad (1.18d)$$

### 1.3 Mesoscopic stochastic modelling

The PDE model introduced in the previous section obviously does not account for fluctuations due to the discrete nature of the underlying reaction steps. On the other hand, at low pressures in the gas phase ( $\approx 10^{-4}$  mbar) each adsorbed CO molecule changes its site about  $10^6$  times before the next particle impinges. This implies that the surface can be regarded as well mixed on a length scale of about  $1 \mu\text{m}$  so that fluctuations are averaged out. With increasing pressure, however, smaller and smaller patches of the surface can be regarded as well mixed and the size of a critical nucleus (a minimal perturbation that would trigger a pulse or a front) decreases. The deterministic PDE models are expected to fail and stochastic effects may become relevant.

To estimate the size of a well-mixed patch we assume the diffusion coefficient  $D_u$  of CO to be known by measurement. A characteristic time scale for the mixing on the surface is given by  $\tau = 1/(k_1 p_u + 2 k_4 p_v)$ , the average time required for the impingement of one monolayer of molecules from the gas phase. (Recall that  $k_1$  and  $k_4$  are the impingement rates of CO and  $\text{O}_2$ , respectively, cf. Table 1.2.) The relationship  $l^2 = 4D_u\tau$  defines a characteristic length scale  $l$  on which the surface may be regarded as well mixed. The number  $n$  of sites in such a well-mixed patch is then about  $l^2/a^2 = (4D_u\tau)/a^2$ , where  $a$  is the ‘effective’ lattice constant. At a pressure of about  $10^{-4}$  mbar, for instance, the calculation above (with  $D_u \approx 10^{-12} \text{ m}^2/\text{s}$ ) yields  $l \approx 1 \mu\text{m}$  and  $n \approx 10^6$ , values that correspond well to experimental observations. Heuristically,  $1/\sqrt{n}$  is a measure for the strength of local coverage fluctuations. While at low pressures the strength of fluctuations is less than  $10^{-3}$  (cf. Fig. 1.11), it reaches the order of 1% at  $10^{-2}$  mbar, the pressure range where raindrop patterns are observed. This suggests that the formation of raindrops is, at least partially, a stochastic effect. Thus a stochastic model is needed to complement the deterministic differential equations model introduced in the previous section.

The appropriate type of stochastic model is a mesoscopic stochastic particle model of the sort described, e.g., in Nicolis & Prigogine (1977); Haken (1983); van Kampen (1992); Gardiner (2004). In order to set up the mesoscopic stochastic particle model, we think of the Pt surface, which for simplicity is assumed to be quadratic, as being divided into  $N^2$  cells of mesoscopic size  $l^2$ . Here  $l$  should be chosen smaller than the characteristic length scale discussed above. Each cell is supposed to contain  $n$  adsorption sites. The state of the surface is, as far as the occupation by adsorbed particles is concerned, described by the vector

$$\mathbf{U}(i, t) = (U(i, t), V(i, t), W(i, t)), \quad i = (i_1, i_2) \in \mathcal{G}. \quad (1.19)$$

Here  $U(i, t)$  denotes the number of CO molecules in cell  $i$  at time  $t$ ,  $V(i, t)$  the number of oxygen atoms,  $W(i, t)$  the number of sites in a  $1 \times 1$  surface structure, and  $\mathcal{G} = \{1, \dots, N\}^2$ . If the surface temperature is to be included in the model, we denote by  $\theta(i, t)$  the rescaled temperature in the volume  $l_T l^2$  below the cell with index  $i$ . The random dynamics of  $(\mathbf{U}, \theta)$

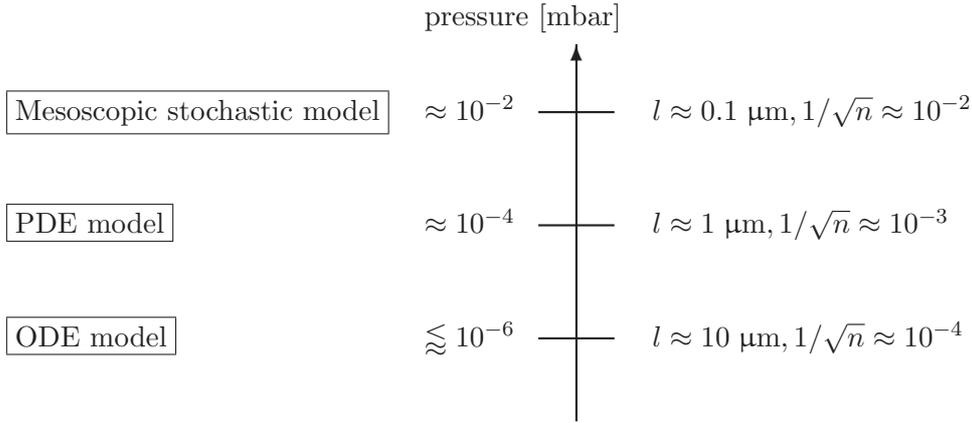


Figure 1.11: Range of validity of different types of models for CO oxidation on Pt.

is characterised by intensities (probabilities per unit time)  $q((\mathbf{U}, \theta), (\tilde{\mathbf{U}}, \tilde{\theta}))$  for jumps in the state space  $S = (\mathbb{N}_0^3 \times \mathbb{R})^{N^2}$  from the present state  $(\mathbf{U}, \theta)$  to another state  $(\tilde{\mathbf{U}}, \tilde{\theta})$  and a rate for the deterministic change of the temperature  $\theta$ . Let  $\chi_{u,i}$  be the state which corresponds to one CO molecule in cell  $i$ , and  $\chi_{v,i}, \chi_{w,i}$  the corresponding states for oxygen atoms and  $1 \times 1$  sites, respectively. Moreover, let  $\chi_{\theta,i}$  be the discrete temperature field with temperature 1 below cell  $i$  and temperature 0 below all other cells.

### 1.3.1 Jumps caused by reaction and diffusion events

All possible transitions due to reaction and diffusion events can be written as

$$(\mathbf{U}, \theta) \rightarrow (\tilde{\mathbf{U}}, \tilde{\theta}) = (\mathbf{U} + \boldsymbol{\delta}, \theta + \varepsilon), \quad (1.20)$$

where  $(\boldsymbol{\delta}, \varepsilon)$  is taken from a finite set of possible transitions  $\mathcal{T}$ .

#### Jumps caused by reaction events

All jumps due to a reaction event in a particular cell  $i$  are listed in Table 1.11. These transitions are assumed to occur with rates that correspond to the reaction rates of the deterministic model. We suppose, for instance, that the deterministic adsorption rate and the intensity for an adsorption event in cell  $i$  in the stochastic model are related by

$$q((\mathbf{U}, \theta), (\mathbf{U} + \chi_{u,i}, \theta + \frac{1}{n}\beta(\bar{T})H_1 \chi_{\theta,i})) = n K_1(\frac{1}{n}\mathbf{U}(i, t), \theta(i, t), \Lambda). \quad (1.21)$$

The same kind of relationship is assumed to hold for all other reaction steps.

In the isothermal case, i.e., at sufficiently low pressures, the temperature  $\theta$  and its changes  $\varepsilon$  may be omitted. Instead of distinguishing the temperature below each cell we may, for

	Process	$\delta$	$\varepsilon$
(1)	Adsorption of CO	$\chi_{u,i}$	$\frac{1}{n}\beta(\bar{T})H_1\chi_{\theta,i}$
(2)	Desorption of CO	$-\chi_{u,i}$	$-\frac{1}{n}\beta(\bar{T})H_2\chi_{\theta,i}$
(3)	Reaction of CO(ad) and O(ad)	$-\chi_{u,i} - \chi_{v,i}$	$\frac{1}{n}\beta(\bar{T})H_3\chi_{\theta,i}$
(4)	Adsorption of O <sub>2</sub>	$2\chi_{v,i}$	$\frac{1}{n}\beta(\bar{T})H_4\chi_{\theta,i}$
(5)	Inverse structural phase transition	$\chi_{w,i}$	0
(6)	Structural phase transition	$-\chi_{w,i}$	0

Table 1.11: Jumps caused by reaction events in a particular cell  $i$ .

sufficiently small surfaces, assume that heat diffusion is infinitely fast. The release of heat due to a reaction process in a particular cell  $i$  then leads to a change of the global surface temperature  $\theta$ . The adsorption of a CO molecule somewhere on the surface, for instance, causes a temperature change by  $\varepsilon = \beta(\bar{T})H_1/(N^2 n)$ .

### Jumps caused by diffusion of CO

The diffusion of CO molecules is modelled by the following transitions. Let  $e_1 = (1, 0)$  and  $e_2 = (0, 1)$ . Then, if cell  $i$  is not a boundary cell,

$$q((\mathbf{U}, \theta), (\mathbf{U} - \chi_{u,i} + \chi_{u,(i \pm e_k)}, \theta)) = n \frac{d}{4} \frac{U(i)}{n}, \quad k = 1, 2, \quad (1.22)$$

where  $d$  is the hopping rate of a CO molecule to a neighbouring cell. Transitions involving the boundary cells have to be specified separately according to the desired boundary conditions. In the simulations in Chapter 4, for instance, we use periodic boundary conditions and thus

$$q((\mathbf{U}, \theta), (\mathbf{U} - \chi_{u,i} + \chi_{u,(i \pm e_k \bmod N)}, \theta)) = n \frac{d}{4} \frac{U(i)}{n}, \quad k = 1, 2, \quad (1.23)$$

for all  $i \in \mathcal{G} = \{1, \dots, N\}^2$ .

### 1.3.2 Temperature drift

Recall that  $l$  denotes the edge length of a cell. Between two jumps caused by reaction or diffusion events the temperature variables  $\theta(i, t)$  move ‘deterministically’ with rate

$$b_i(\theta(t)) = -\gamma(\bar{T})\theta(i, t) + D_\theta \sum_{k=1,2} \frac{1}{l^2} (\theta(i - e_k, t) - 2\theta(i, t) + \theta(i + e_k, t)). \quad (1.24)$$

Again the rates for the boundary cells have to be specified according to the desired boundary conditions. If heat diffusion is assumed infinitely fast, the second term is omitted. Such a deterministic motion is in probabilistic language called a *drift*.

### 1.3.3 Summary

The existence of a stochastic process that corresponds to the transition intensities above is discussed for the isothermal case in a more general setting in Chapters 2 and 3. A mathematically rigorous discussion of the existence of the process including the temperature drift is left aside here but should cause no principal problems. We only remark that the Markov process corresponding to the transition intensities and drift terms specified above is characterised, at least heuristically, by its generator  $L$ . The generator of a Markov process is an operator that contains all information about the infinitesimal behaviour of the process in condensed form. In our case, it describes for all real-valued functions  $g$  on the state space  $S$  in its domain of definition how the conditional expectations

$$E[g(\mathbf{U}(t+h), \theta(t+h)) - g(\mathbf{U}(t), \theta(t)) | \mathbf{U}(t), \theta(t)] \quad (1.25)$$

behave for small  $h$ . That is,

$$E[g(\mathbf{U}(t+h), \theta(t+h)) - g(\mathbf{U}(t), \theta(t)) | \mathbf{U}(t), \theta(t)] = Lg(\mathbf{U}(t), \theta(t)) h + o(h). \quad (1.26)$$

To write down the generator in compact form we first construct from the transition intensities an intensity function  $c$  and a transition kernel  $\mu$  by setting

$$c(\mathbf{U}, \theta) = \sum_{\substack{(\boldsymbol{\delta}, \varepsilon) \in \mathcal{T}: \\ (\mathbf{U} + \boldsymbol{\delta}, \theta + \varepsilon) \in S}} q((\mathbf{U}, \theta), (\mathbf{U} + \boldsymbol{\delta}, \theta + \varepsilon)), \quad (1.27)$$

and

$$\mu(\mathbf{U}, \theta; B) = c(\mathbf{U}, \theta)^{-1} \sum_{\substack{(\boldsymbol{\delta}, \varepsilon) \in \mathcal{T}: \\ (\mathbf{U} + \boldsymbol{\delta}, \theta + \varepsilon) \in B}} q((\mathbf{U}, \theta), (\mathbf{U} + \boldsymbol{\delta}, \theta + \varepsilon)) \quad (1.28)$$

for subsets  $B \subset S$ . The probabilistic interpretation of  $c$  and  $\mu$  is that the stochastic process leaves the state  $(\mathbf{U}, \theta)$  by a jump with rate  $c(\mathbf{U}, \theta)$ , and the jump terminates with probability  $\mu(\mathbf{U}, \theta; B)$  in the subset  $B$  of the state space  $S$ . In the isothermal case  $\theta$  and  $\varepsilon$  are of course omitted. Hence, if we assume constant temperature, the generator is given by

$$\begin{aligned} Lg(\mathbf{U}) &= c(\mathbf{U}) \int_S (g(\tilde{\mathbf{U}}) - g(\mathbf{U})) \mu(\mathbf{U}; d\tilde{\mathbf{U}}) \\ &= \sum_{\substack{\boldsymbol{\delta} \in \mathcal{T}: \\ \mathbf{U} + \boldsymbol{\delta} \in S}} q(\mathbf{U}, \mathbf{U} + \boldsymbol{\delta}) (g(\mathbf{U} + \boldsymbol{\delta}) - g(\mathbf{U})). \end{aligned} \quad (1.29)$$

It has the typical form of a generator of a pure jump Markov process (see, e.g., Revuz & Yor (2005)). In the case of nonconstant temperature we have

$$\begin{aligned}
 Lg(\mathbf{U}, \theta) &= c(\mathbf{U}, \theta) \int_S (g(\tilde{\mathbf{U}}, \tilde{\theta}) - g(\mathbf{U}, \theta)) \mu(\mathbf{U}, \theta; d(\tilde{\mathbf{U}}, \tilde{\theta})) \\
 &\quad + \sum_{i \in \mathcal{G}} b_i(\theta) \frac{\partial g}{\partial \theta(i)}(\mathbf{U}, \theta) \\
 &= \sum_{\substack{(\delta, \varepsilon) \in \mathcal{T}: \\ (\mathbf{U} + \delta, \theta + \varepsilon) \in S}} q((\mathbf{U}, \theta), (\mathbf{U} + \delta, \theta + \varepsilon)) (g(\mathbf{U} + \delta, \theta + \varepsilon) - g(\mathbf{U}, \theta)) \\
 &\quad + \sum_{i \in \mathcal{G}} b_i(\theta) \frac{\partial g}{\partial \theta(i)}(\mathbf{U}, \theta).
 \end{aligned} \tag{1.30}$$

A generator of this type characterises a Markov jump process with inter-jump drift.

## Discussion

In the present chapter we have introduced, after a discussion of the experimental background, a deterministic macroscopic PDE model and a complementary mesoscopic stochastic particle model for CO oxidation on Pt surfaces. A numerical bifurcation analysis for the kinetic part of the deterministic model shows good agreement with experimental results. The stochastic model will be used in Chapter 4 for the simulation of raindrop patterns (Figs. 1.8 and 1.9). Various spatio-temporal patterns (fronts, travelling pulses, spirals and chemical turbulence) have been investigated numerically by Bär (1993) in his thesis and related work on the basis of the model by Krischer et al. (1992). Since the bifurcation structure of our model is similar, we expect that it is capable of reproducing these patterns as well. However, this has not been checked systematically.

The standing wave patterns (Fig. 1.7b) have been simulated with a model including a subsurface oxygen species (oxygen atoms below the surface) and a global coupling via the gas phase by von Oertzen et al. (2000). However, it is not fully clarified if a gas phase coupling is really necessary, since the formation of standing waves could be explained in the framework of reaction-diffusion equations by a Turing-Hopf instability (Krömker, 1997). So far there exists (to our knowledge) no realistic model that reproduces the slow sinusoidal oscillations at low pressures or the period-doubling transition to chaos.

So far we have looked at CO oxidation on platinum exclusively from the complex systems point of view. Apart from its role as standard system for the study of pattern formation, CO oxidation on Pt is of great practical importance, since the oxidation of CO is one of the objectives of automobile exhaust catalysts, so-called three-way catalytic converters (TWCs). The other objectives are oxidation of unburnt hydrocarbons ( $C_xH_y$ ) and reduction of nitrogen oxides ( $NO_x$ ). Both oxidation reactions are known to be catalysed by platinum, while the reduction of  $NO_x$  is catalysed by rhodium. Current TWCs consist of a honeycomb-shaped

ceramic body with channels axially orientated in the direction of flow of the exhaust gas (Hayes & Kolaczkowski, 1997; Thomas & Thomas, 1996). The channel walls are covered by a thin layer of porous material (the washcoat). Small crystallites of platinum and rhodium, on the surfaces of which the reactions take place, adhere to the interior surfaces of the porous washcoat. Owing to the complex geometry of the TWC, its performance cannot be deduced directly from knowledge about the reactions on single crystal surfaces. A thorough understanding of single crystal experiments is, however, mandatory for understanding how the TWC works and for optimisation. The simulation of pattern formation on single crystal surfaces provides a good means for testing and validating different models.

The introduction of two models, a macroscopic deterministic model and a mesoscopic stochastic particle model, naturally raises the question whether these two models are consistent. This problem shall be dealt with in a more general setting in the following two chapters.

## Chapter 2

# Law of large numbers for linear models

### Overview

In the previous chapter we have modelled the CO oxidation reaction on Pt single crystal surfaces both with a system of partial differential equations (PDEs) and a corresponding mesoscopic stochastic particle model. In this chapter we start to investigate in a general setting how macroscopic deterministic PDE models and mesoscopic stochastic particle models are related. Our aim is to prove rigorously that the PDE models approximate the corresponding stochastic particle models in the limit of large particle numbers.

The stochastic particle models we call mesoscopic are, in essence, a combination of a continuous-time version of the classical urn model by P. and T. Ehrenfest for diffusion through a membrane (see, e.g., Karlin & Taylor (1975)) and the standard stochastic model for chemical reactions (see, e.g., van Kampen (1992)). They have been described since the early seventies by many authors in physics (Nicolis & Prigogine, 1977; Gillespie, 1977; Haken, 1983; Gardiner, 2004; van Kampen, 1992) and mathematics (Kurtz, 1981; Arnold & Theodosopulu, 1980; Kotelenz, 1986, 1988; Blount, 1991, 1993, 1994; Guias, 2002). In the physical literature the model is often simply called ‘the’ stochastic model for chemical reactions; in mathematics models of this type are also known as ‘density-dependent population processes’. Similar models for coupled waiting lines, so-called queueing networks, are dealt with in communication theory (Brémaud, 1999).

By terming those particle models ‘mesoscopic’ we wish to point out that they are set up at mesoscopic time and length scales. In particular, as could be seen in the preceding chapter, we do *not* take into account *explicitly* the interactions between individual particles; rather the intensities for births and deaths of particles depend on the local densities. The presence of a mechanism that ensures a rapid local stirring is assumed.

Our method for deriving laws of large numbers generally proceeds in two steps. We first

study the convergence of a semi-discrete finite-difference approximation of the limit equation where the spatial derivatives are replaced by finite differences. Having established the convergence of the semi-discrete approximation, the second step in the proof consists in estimating the distance between the approximation and a particle density associated to the stochastic particle model in an appropriate norm. This procedure is motivated by the observation that the particle density generally satisfies a stochastic differential equation that can be regarded as a spatially semi-discretised finite-difference approximation of the macroscopic PDE perturbed by a martingale noise term. In previous treatments (Kotelenez, 1986, 1988; Blount, 1991, 1993, 1994; Guias, 2002) laws of large numbers have been shown for linear and certain nonlinear models by means of semigroup methods. In particular, the solutions of the limit equations are characterised as the mild solutions that one obtains from the semigroup approach to linear and semilinear parabolic PDEs. Our method, which is inspired by Oelschläger's treatment of 'moderately' interacting particle systems (Oelschläger, 1989), is related to the variational approach to parabolic PDEs. The solution of the limit equation is an appropriately defined weak solution the existence of which is usually established with Hilbert-space methods. (See also the discussion in the introduction.)

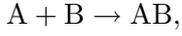
One major obstacle that has to be overcome in our approach is, as already mentioned, the approximation of such weak solutions with a semi-discrete finite-difference method. Although at first sight spatial semi-discretisation with finite-differences seems to be a legitimate approach to solving parabolic equations that appears natural in our context, it is rarely used in the literature. (See however Lions (1969), Chapter 4, for an example.) The reason is, of course, that the Faedo-Galerkin method or semi-discretisation in time (Rothe's method) are usually much more convenient. Nevertheless we have at our disposal the methodology that has been developed for the analysis of fully discrete finite-difference schemes (Raviart, 1967; Temam, 1973, 2001; Zeidler, 1990c).

For didactic purposes we restrict the discussion in the present chapter to linear models. Nonlinearities will be treated in Chapter 3. In the first part of the following section we introduce the general type of macroscopic PDE model that will later appear as deterministic limit of the mesoscopic stochastic particle models. The macroscopic model is, at this point, introduced only on the grounds of thermodynamical arguments. In the second part we give a detailed description of the linear stochastic particle models and associated particle densities that are obtained by rescaling the original model. In order not to obscure the simple main ideas by too much notation, we always work in parallel with an example model involving only one species and a general linear model. To prepare the first step of the derivation of the law of large numbers we then discuss in some detail the solution of the limit PDE in Section 2.2 and introduce an approximation in terms of an analogous equation with discretised spatial derivatives. In Section 2.3 we carry out the proof of the law of large numbers for the example model, and in Section 2.4 the general linear model is treated in a similar way.



### 2.1.1 The macroscopic PDE model

In a continuum approach the reactions (2.1) are assumed to proceed locally with rates  $K_i : \mathbb{R}^{n_s} \rightarrow \mathbb{R}$ ,  $i = 1, \dots, n_r$ , that are functions of the local concentrations  $[C_j]$  of the species  $C_j$ ,  $j = 1, \dots, n_s$ . The reaction rate of a bimolecular reaction



for instance, is typically modelled by

$$K = k[A][B].$$

Here the product  $[A][B]$  is a measure for the probability to find a molecule of species A close to a molecule of species B. The constant  $k$  is the *reaction rate constant* which, unfortunately, is also called reaction rate by some authors. In the general case the rate of reaction  $i$  is often modelled by

$$K_i = k_i \prod_{j=1}^{n_s} [C_j]^{n_{ij}}. \quad (2.4)$$

However, this ansatz is not universal if we allow chemical reactions in a broad sense. For example, it does not cover the adsorption of CO or oxygen molecules on a platinum surface. The range of validity of the ansatz (2.4) is discussed in van Kampen (1992).

On the macroscopic level the dynamics of the concentrations  $u_j = [C_j]$  is described by a system of  $n_s$  mass-balance equations on the space-time domain  $Q_T = G \times (0, T)$ ,  $T > 0$  being the time of observation:

$$\partial_t u_j + \nabla \cdot \mathbf{J}_j(x, \mathbf{u}, \nabla \mathbf{u}) = f_j(x, \mathbf{u}), \quad j = 1, \dots, n_s. \quad (2.5)$$

Here  $\mathbf{u} = (u_1, \dots, u_{n_s})$ , and  $\nabla \mathbf{u} = ((\nabla u_1)^T, \dots, (\nabla u_{n_s})^T)^T$ . In addition, appropriate boundary and initial conditions have to be specified. We assume that the reaction functions  $f_j$  and the fluxes  $\mathbf{J}_j$  do not depend explicitly on the space variable  $x$ . The reaction functions  $f_j : \mathbb{R}^{n_s} \rightarrow \mathbb{R}$  are obtained from the (appropriately rescaled) reaction rates  $K_i$  in the following way. We first define the matrix  $(\nu_{ij}) \in \mathbb{Z}^{n_r \times n_s}$  by  $\nu_{ij} = \tilde{n}_{ij} - n_{ij}$ . Then

$$f_j(\mathbf{u}) = \sum_{i=1}^{n_r} \nu_{ij} K_i(\mathbf{u}). \quad (2.6)$$

The vector functions  $\mathbf{J}_j : \mathbb{R}^{n_s} \times \mathbb{R}^{n_s \times m} \rightarrow \mathbb{R}^m$ ,  $j = 1, \dots, n_s$ , are appropriate ‘constitutive laws’ for the diffusive mass flux.

The reaction functions  $f_j$  and the fluxes  $\mathbf{J}_j$  have to be chosen appropriately by the modeller for the specific system under consideration. In principle, the choice is limited for physical reasons by a condition of positive entropy production (see, e.g., Hutter & Jöhnk (2004)), but (to our knowledge) there is no generally accepted and easily applicable condition available.

Above we have introduced the reaction rates  $K_i$  as functions mapping  $\mathbb{R}^{n_s}$  to  $\mathbb{R}$ . This is convenient mathematically, e.g., for proving the existence of solutions of special cases of Eq. (2.5). However, common sense tells us that the  $K_i$  should simply be defined as mappings from  $(\mathbb{R}_0^+)^{n_s}$  to  $\mathbb{R}_0^+$ . Throughout this work we assume that the reaction rates  $K_i$  (and hence the reaction functions  $f_j$ ) are defined on  $\mathbb{R}^{n_s}$ , but that their restrictions to  $(\mathbb{R}_0^+)^{n_s}$  (or a subset thereof) satisfy certain physically reasonable conditions (cf. conditions (C1) and (C2) below).

Sometimes the number of equations can be reduced further by making use of conservation constraints. In the CO oxidation scheme (2.2) above, for instance,  $[\star(1 \times 1)] + [\star(\text{rec})]$  should obviously be conserved. Thus one of the two phase variables can be eliminated.

In general, the reaction functions  $f_j(\mathbf{u})$  and the diffusion operators  $\nabla \cdot \mathbf{J}_j(\mathbf{u}, \nabla \mathbf{u})$  are nonlinear. However, as already announced, we concentrate in the present chapter on linear models.

### A simple example model

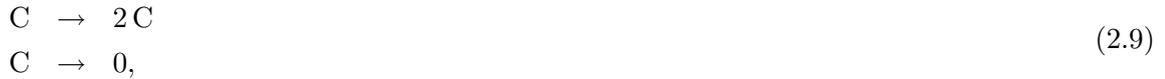
In order to introduce our method of proving laws of large numbers with a minimum of technical and notational difficulties, we treat in Section 2.3 a simple system involving only one species in a one-dimensional reactor which is represented by the interval  $(0, L)$ . Particles of the species are born and die with linear rates, and thus the reaction function reads

$$f(u) = k_1 u - k_2 u, \quad k_1, k_2 > 0. \quad (2.7)$$

Moreover, we assume simple Fickian diffusion, i.e.,

$$J = -D \nabla u, \quad (2.8)$$

where  $D > 0$  is the diffusion coefficient. The reaction function  $f$  may be thought of as being derived from the scheme



assuming that the first reaction proceeds with rate  $K_1(u) = k_1 u$  and the second reaction with rate  $K_2(u) = k_2 u$ . The macroscopic PDE is thus given by

$$\partial_t u - D \Delta u = k_1 u - k_2 u \quad \text{on } Q_T = (0, L) \times (0, T). \quad (2.10)$$

### The general linear model

In the general case the  $n_r$  reactions proceed with rates  $K_i : \mathbb{R}^{n_s} \rightarrow \mathbb{R}$  that are in this chapter assumed to be linear functions of the concentrations  $\mathbf{u}$ , i.e.,

$$K_i(\mathbf{u}) = k_{i,1} u_1 + \dots + k_{i,n_s} u_{n_s} \quad (2.11)$$

with constants  $k_{ij} \in \mathbb{R}$ ,  $i = 1, \dots, n_r$ ,  $j = 1, \dots, n_s$ . Furthermore, we make the following assumptions (for  $i = 1, \dots, n_r$ ,  $j = 1, \dots, n_s$ ).

$$K_i(\mathbf{w}) \geq 0 \text{ for all } \mathbf{w} \in (\mathbb{R}_0^+)^{n_s}. \quad (\text{C1})$$

$$\text{If } \nu_{ij} < 0 \text{ then } K_i(\mathbf{w}) = 0 \text{ for all } \mathbf{w} \in (\mathbb{R}_0^+)^{n_s} \text{ with } w_j = 0. \quad (\text{C2})$$

Conditions (C1) and (C2) are obvious from a physical point of view and considerably restrict the class of admissible linear reaction rates. For the example model they were automatically satisfied. Condition (C1) implies that all  $k_{ij}$  are greater than or equal to zero. The second condition implies that if for a particular reaction  $i$  there is a  $j$  such that  $\nu_{ij}$  is negative, then the rate  $K_i$  is either identically zero (i.e., the reaction can be deleted from the scheme) or proportional to  $u_j$ . The reaction functions  $f_j$  are given by Eq. (2.6). Furthermore we assume that the particles of each species perform a simple Fickian diffusion with diffusion coefficient  $D_j > 0$ ,  $j = 1, \dots, n_s$ . The system of PDEs for the concentrations that describes the time evolution of the system on the macroscopic level is thus given by

$$\partial_t u_j - D_j \Delta u_j = f_j(\mathbf{u}) \quad \text{on } Q_T = G \times (0, T), \quad (2.12)$$

$j = 1, \dots, n_s$ , where  $G \subset \mathbb{R}^m$  is the bounded domain representing the chemical reactor.

### 2.1.2 The mesoscopic stochastic particle model

The mesoscopic stochastic particle models for reaction-diffusion systems we aim to study are motivated by the stochastic model for CO oxidation on platinum from the previous chapter (Section 1.3). Thus, for physical motivation and heuristic discussions we shall keep to the terminology of catalytic surface reactions. If we think of a surface reaction the ‘reactor’ consists of a metal surface, and molecules from a gaseous or liquid phase are adsorbed at certain adsorption sites that form a two-dimensional lattice. The adsorbed atoms or molecules may perform a diffusive hopping from site to site on the surface and take part in chemical reactions. Two types of particle models are commonly used for a quantitative description of such systems. For the following discussion we shall refer to them as ‘mesoscopic’ and ‘microscopic’ particle models.

#### Mesoscopic versus microscopic particle models

The state space of a model we would call microscopic is, roughly speaking, a subset of  $C^{\mathbb{Z}^m}$ , where a grid point of  $\mathbb{Z}^m$  is identified with an adsorption site, and the set  $C$  contains all configurations a site may attain. In the simplest case  $C = \{0, 1\}$  where 1 stands for occupied and 0 for empty. Hence, the state contains information about each individual site; we know if it is occupied at a certain moment in time or not and by which sort of particles. If a particle changes its position on the lattice or is transformed by a chemical reaction, one or several

sites change their configuration. The rates of change may, in principle, depend on the state of the whole system, i.e., on the configuration of each individual site. In practice, of course, a particular site is only influenced by sites in a certain neighbourhood. Models of this type are known in the physical literature as ‘lattice-gas models’ or (dynamic) ‘Monte-Carlo models’ and in mathematics as ‘interacting particle systems’ (Kipnis & Landim, 1999; de Masi & Presutti, 1991; Liggett, 2005) or ‘stochastic spatial models’ (Durrett, 1999).

Heuristically, a macroscopic description of the evolution of a many-particle system in terms of a PDE always requires a rapid stirring mechanism to ensure a ‘local equilibrium’. While in the mathematical treatment of microscopic particle models it is a major problem to define precisely the meaning of local equilibrium, an equilibrium assumption enters already in the formulation of the mesoscopic models. The state space of a mesoscopic particle model may again be defined as subset of a set of the form  $C^{\mathbb{Z}^m}$ . However, here a grid point in  $\mathbb{Z}^m$  represents a cell or compartment of mesoscopic size, and the set  $C$  is the set of possible particle numbers in a cell for each species. That is,  $C = \mathbb{N}_0^{n_s}$  or a subset thereof. Such cells may contain a relatively large number of sites in the sense of the microscopic models. They are assumed to be always well-mixed due to a sufficiently rapid stirring mechanism (e.g., fast diffusion of one or several species). The rates at which particles are created or destroyed in a particular cell are assumed to be functions of the particle densities in the cell (which corresponds to a local equilibrium assumption). Roughly speaking, this corresponds to the assumption that the probability to find a particle of a certain species at a particular site inside a cell is equal to the density of that species. Thus, information about the detailed configuration of particles is not available if one uses the mesoscopic approach; one only knows how many particles of each species are at a certain moment in time located in a certain cell.

The mathematical structure of both model types is the same. Technically, both models are Markov chains in continuous time with a finite or countable state space (assuming that we consider only a finite array of cells or sites). The main difference between both model types is that a certain site of a microscopic model can typically attain only a few different configurations, whereas the number of possible configurations of a cell of a mesoscopic model may be large or infinite.

At the present stage a derivation from first principles (i.e., quantum mechanics) seems to be out of reach for either of the stochastic models for surface reactions discussed above. Yet an appropriate particle model of the microscopic type may be considered as a satisfactory mathematical description of a surface reaction on the molecular level. However, the microscopic models generally bear two disadvantages. The first one is that their analysis is notoriously difficult. Although considerable progress has been made during the past decades, the derivation of macroscopic limit equations by a law of large numbers still doesn’t seem feasible for most ‘realistic’ models. The second disadvantage is computational: even with today’s compute power only relatively small patches of a surface can be simulated.

Mesoscopic models, in contrast, are generally easier to analyse and allow for the derivation of results of a more general nature. Although they do not describe the system in detail on the molecular level, they provide plausible models for the fluctuations caused by the discrete nature of the reaction processes occurring on the surface. Furthermore, the simulation of mesoscopic models is not limited to small surface areas.

In the present work we thus follow the mesoscopic approach. In summary, we think of the surface (or any other chemical reactor) as being divided into well-mixed cells of mesoscopic size each containing a possibly large number of sites. The chemical reactions inside the cells and the exchange of particles between adjacent cells is random and modelled by jump (birth-death) processes. The intensities of births and deaths are supposed to depend only on the particle densities in the cells, neglecting any details about the particle configuration inside the cell.

Conceptually related to the mesoscopic stochastic particle models discussed above are so-called ‘moderately’ interacting particle systems (Oelschläger, 1989).

### The scales

In order to set up the model, a careful discussion of the characteristic length and time scales that can be identified in the physical system is required. These scales will appear as parameters in the mathematical model and are varied in the course of the derivation of the law of large numbers.

In a reaction-diffusion system typically three different characteristic length scales can be identified: the total size of the system  $L$ , a ‘diffusion length’  $l$  which corresponds to the cell size (a length scale on which the system can be considered as well mixed) and the distance between two neighbouring sites  $\lambda$  (or another typical scale for the mean inter-particle distance). We postulate

$$\lambda \ll l \ll L,$$

which is certainly a reasonable assumption for many systems. The micro-scale  $\lambda$  will not appear explicitly in the mesoscopic models. These three length scales lead in a natural way to two ratios,

$$N = L/l \gg 1 \quad \text{and} \quad n = l/\lambda \gg 1,$$

which in one space dimension correspond to the number of cells and the number of sites per cell, respectively. The law of large numbers we are aiming at can be regarded as an idealisation obtained by letting both ratios tend to infinity. In our approach we keep the system size  $L$  fixed. Hence, the cell size  $l$  and the distance between two sites  $\lambda$  must go to zero because otherwise the number of cells and the number of particles per cell cannot become infinite. Alternatively, we could fix  $\lambda$  and let  $l$  and  $L$  tend to infinity.

In a similar way one can identify three different time scales: a time scale which corresponds to the hopping rate from site to site  $\delta$  and does not appear explicitly in the mesoscopic model, a time scale which corresponds to the ‘hopping rate’  $d$  from cell to cell, and, finally, the time of observation  $T$ . We assume

$$1/\delta \ll 1/d \ll T.$$

Later in this section we shall introduce families of stochastic particle models with associated particle densities  $u_l$  that are labelled by the scale parameter  $l$ . The law of large numbers says that  $u_l$  approaches the solution of a limit PDE for  $l \rightarrow 0$ . In view of the discussion above, the parameters that may independently be varied to achieve this aim are  $l$  ( $l = L/N$ ,  $L$  being fixed),  $n$  and  $d$ . Thus, we shall consider sequences of particle models with varying parameters  $l$ ,  $n$ , and  $d$ .

The question now arises whether these scale parameters are independent or if there exist relations among them that have to be satisfied to get a reasonable limit. We shall see that it is necessary to assume the familiar relation

$$dl^2 \sim d/N^2 \sim D \tag{2.13}$$

between  $d$  and  $l$ , where  $D$  denotes the macroscopic diffusion coefficient. It will become apparent in the course of the derivation of the law of large numbers that it is convenient to postulate a second scaling relation concerning the coupling of  $d$  and  $n$  (or  $l$  and  $n$ , in view of (2.13)).

We remark that one may also consider the situation  $\lambda \lesssim l$ . This amounts to deriving a law of large numbers for  $N \rightarrow \infty$  while keeping  $n$  constant. A result of this type is proved in Blount (1994).

### The stochastic example model

We now introduce the mesoscopic stochastic particle model that corresponds to the macroscopic PDE (2.10). We take the occasion to recall some definitions and important results from the theory of stochastic processes that will be required later.

In the mesoscopic picture the state  $U(t)$  of the system at time  $t \geq 0$  is given by the collection of particle numbers  $U(i, t)$  in the cells  $i = 1, \dots, N$ ,  $N \in \mathbb{N}$  being the total number of cells:

$$U(t) = (U(i, t))_{i=1, \dots, N} \in S = \mathbb{N}_0^N. \tag{2.14}$$

The infinitesimal characteristics of the dynamics of the stochastic process  $(U(t))_{t \geq 0}$  are the transition intensities  $q(U, \tilde{U})$  from  $U \in S$  to other states  $\tilde{U}$ . The so-called  $Q$ -matrix  $q(\cdot, \cdot)$  is real-valued and nonnegative if  $\tilde{U} \neq U$ , and  $q(U, \tilde{U})h$  is interpreted as the probability to jump

from  $U$  to  $\tilde{U}$  in a small time interval  $h$ . Let  $\chi_i$ ,  $i = 2, \dots, N-1$ , be the  $i$ th ‘unit vector’, the state corresponding to one particle in cell  $i$  and zero particles in all remaining cells, and let  $\chi_1 = \chi_N \equiv 0$ , the state corresponding to zero particles in all cells. For the sake of simplicity we shall impose ‘Dirichlet’ boundary conditions, i.e., the particle numbers in cells 1 and  $N$  are always zero:  $U(1, \cdot) = U(N, \cdot) \equiv 0$ . The following transitions may occur.

- ▶ A particle may leave cell  $i = 2, \dots, N-1$  and jump to cell  $i-1$  or  $i+1$ .

$$\begin{aligned} q(U, U - \chi_i + \chi_{i-1}) &= n \frac{d}{2} \frac{U(i)}{n}, \\ q(U, U - \chi_i + \chi_{i+1}) &= n \frac{d}{2} \frac{U(i)}{n}, \end{aligned} \tag{2.15}$$

where  $d > 0$  is the ‘hopping rate’ of a particle to a neighbouring cell.

- ▶ A particle in cell  $i = 2, \dots, N-1$  may give birth to another one.

$$q(U, U + \chi_i) = n k_1 \frac{U(i)}{n}. \tag{2.16}$$

- ▶ A particle in cell  $i = 2, \dots, N-1$  may die.

$$q(U, U - \chi_i) = n k_2 \frac{U(i)}{n}. \tag{2.17}$$

Here the constants  $k_1$  and  $k_2$  are the same as in the PDE model.

The intensity for all other possible transitions is zero.

Note that all transitions with intensity nonzero are of the form  $U \rightarrow \tilde{U} = U + \delta$  for  $\delta$  from a finite set  $\mathcal{T} \subset \mathbb{Z}^N$ . Moreover, there is a constant  $C$  such that

$$q(U, U + \delta) \leq C \|U\|_{\mathbb{R}^N} \quad \text{for all } \delta \in \mathcal{T}, \tag{2.18}$$

where  $\|\cdot\|_{\mathbb{R}^N}$  denotes the Euclidian norm in  $\mathbb{R}^N$ . The Euclidian norm of the ambient space  $\mathbb{R}^N$  induces a metric on the countable state space  $S = \mathbb{N}_0^N$ . This metric, in turn, induces the discrete topology on  $S$  and of course makes it a locally compact, separable, metric space, the conditions required by the semigroup approach to the construction of Markov processes. It follows from a general theorem (Theorem 3.1 in Chapter 8 of Ethier & Kurtz (1986), cf. Corollary 2) that there exists a stochastic process with state space  $S$  having transition intensities as specified above. More precisely, given an initial distribution for  $U(0)$  on the measurable space  $(S, 2^S)$  there exists a probability space  $(\Omega, \mathcal{A}, P)$  (the probability measure  $P$  depending on the initial distribution) and a family of random variables

$$(U(t))_{t \geq 0} = (U(1, t), \dots, U(N, t))_{t \geq 0} \tag{2.19}$$

with values in  $S = \mathbb{N}_0^N$  such that the process  $(U(t))$  has the prescribed infinitesimal behaviour. To make this statement more precise, let  $\mathcal{F}_t = \sigma(U(s), s \leq t) \subset \mathcal{A}$  be the  $\sigma$ -field induced

by the random variables  $U(s)$ ,  $s \leq t$ , i.e., the  $\sigma$ -field generated by subsets of  $\Omega$  of the form  $\{U(t_1) \in B_1, \dots, U(t_k) \in B_k\}$ , where  $0 \leq t_1 \leq \dots \leq t_k \leq t$ ,  $B_1, \dots, B_k \subset S$ , and  $k \in \mathbb{N}$ . In other words,  $(\mathcal{F}_t)_{t \geq 0}$  is the *filtration* (nondecreasing family of  $\sigma$ -fields of sets in  $\mathcal{A}$ ) induced by  $(U(t))_{t \geq 0}$ . Then the probability for a diffusive jump to the left, for instance, behaves as

$$P\left[U(t+h) = U(t) - \chi_i + \chi_{i-1} \mid \mathcal{F}_t\right] = q(U(t), U(t) - \chi_i + \chi_{i-1})h + o(h) \quad (h \rightarrow 0),$$

$i = 2, \dots, N-1$ . Furthermore,  $(U(t))$  has the *Markov property*, i.e., for  $t, s > 0$ ,  $B \subset S$ ,

$$P\left[U(t+s) \in B \mid \mathcal{F}_t\right] = P\left[U(t+s) \in B \mid U(t)\right]. \quad (2.20)$$

The process  $(U(t))$  may be constructed having right-continuous paths with left limits. This implies that  $(U(t))$  is *progressively measurable*, which means that the function  $(s, \omega) \mapsto U(s, \omega)$  regarded as mapping from  $[0, t] \times \Omega$  to  $S$  is measurable  $\mathcal{B}[0, t] \times \mathcal{F}_t / 2^S$  for all  $t > 0$ . Here  $\mathcal{B}[0, t]$  denotes the Borel  $\sigma$ -field on the interval  $[0, t]$ , and  $\mathcal{B}[0, t] \times \mathcal{F}_t$  is the product  $\sigma$ -field. In particular, we are allowed to work with double integrals over  $[0, t] \times \Omega$ .

Let  $\hat{C}(S)$  be the space of continuous real-valued functions on  $S$  vanishing at infinity equipped with the supremum norm. (In our case, since the state space  $S$  is endowed with the discrete topology, the continuity assumption is not a restriction.) The Markov process  $(U(t))$  corresponds to a *Feller semigroup* (see, e.g., Ethier & Kurtz (1986), pp. 162) of operators on  $\hat{C}(S)$  with generator  $L$  given by

$$Lg(U) = \sum_{\tilde{U} \neq U} q(U, \tilde{U}) (g(\tilde{U}) - g(U)), \quad g \in \hat{C}(S). \quad (2.21)$$

This means that there is a strongly continuous, positive, contraction semigroup of bounded linear operators  $(T(t))_{t \geq 0}$  on  $\hat{C}(S)$  with (conservative) generator  $L$  such that

$$E\left[g(U(t+s)) \mid \mathcal{F}_t\right] = T(s)g(U(t)), \quad g \in \hat{C}(S), \quad t, s \geq 0. \quad (2.22)$$

In general, the generator  $L$  is defined only on a subset of  $\hat{C}(S)$ .

Recall that a real-valued stochastic process  $(X(t))_{t \geq 0}$  with  $E[|X(t)|] < \infty$  for all  $t \geq 0$  is called a *martingale* with respect to a filtration  $(\mathcal{F}_t)_{t \geq 0}$  if it is adapted (i.e.,  $X(t)$  is  $\mathcal{F}_t$ -measurable for all  $t \geq 0$ ) and satisfies

$$E[X(t+s) \mid \mathcal{F}_t] = X(t) \quad \text{for all } t, s \geq 0. \quad (2.23)$$

A nonnegative real-valued random variable  $\tau$  is called a *stopping time* (with respect to the filtration  $(\mathcal{F}_t)$ ) if  $\{\tau \leq t\} \in \mathcal{F}_t$  for all  $t \geq 0$ . Equation (2.24) below, also known as *Dynkin's formula*, is an essential ingredient in the proof of the law of large numbers. It identifies martingales related to a Markov process and holds generally for  $((\mathcal{F}_t)$ -) Markov processes corresponding to a Feller semigroup with generator  $L$ . Let  $g$  be a function in  $\hat{C}(S)$  (or, in the general case, from the domain of  $L$ ). Then the process

$$M_g(t) = g(U(t)) - g(U(0)) - \int_0^t Lg(U(s)) ds, \quad t \geq 0, \quad (2.24)$$

is a  $((\mathcal{F}_t)$ -) martingale. Moreover, the stopped process  $(M_g(t \wedge \tau))_{t \geq 0}$  is a martingale for each  $((\mathcal{F}_t)$ -) stopping time  $\tau$ . For more details on these topics see the relevant chapters of Ethier & Kurtz (1986); Kallenberg (2002); Revuz & Yor (2005).

In order to obtain a particle density, we now rescale the stochastic process  $(U(t))$ . Let  $l = L/N$  and set  $z_i = (i - 1/2)l$ ,  $i = 1, \dots, N$ . We denote by  $c_l(z_i)$  the open interval of length  $l$  with  $z_i$  as midpoint, i.e.,  $c_l(z_i) = (z_i - l/2, z_i + l/2)$ . The union of the intervals  $c_l(z_i)$  is, except for a finite number of points, the interval  $G = (0, L)$  which represents the chemical reactor. Let  $\mathbb{1}_{c_l(z_i)}(\cdot)$  be the indicator function of the interval  $c_l(z_i)$ . It is natural to associate to  $U(t)$  the step-function valued particle density  $u_l(\cdot, t) : \mathbb{R}^m \rightarrow \frac{1}{n}\mathbb{N}_0$  given by

$$x \mapsto u_l(x, t) = \sum_{i=1}^N \frac{U(i, t)}{n} \mathbb{1}_{c_l(z_i)}(x). \quad (2.25)$$

(Recall that  $n$  represents the number of sites per cell.) That is, the particle number in each cell is scaled by  $1/n$  and the index  $i$  numbering the cells is ‘scaled’ by  $l$ . Here we have labelled the particle density by the parameter  $l$ , but keep in mind that it depends on the parameters  $l = L/N$ ,  $n$ , and  $d$ .

In order to specify precisely the state space of the particle density process  $(u_l(t))_{t \geq 0}$ , we now introduce a few notions the utility of which will become clearer during the derivation of the law of large numbers. They are useful also for the discussion of the general linear model in the next paragraph.

It seems natural to consider the step-function valued particle density defined above as an element of the space  $L^2$  (or any other  $L^p$ ). In fact, it will turn out to be useful to regard it as an element of a discrete version  $\mathcal{L}^2(\mathcal{G}_l)$  of the Lebesgue space  $L^2(G)$ . Let for the following discussion (cf. Zeidler (1990c), Chapter 35)  $G$  be a bounded domain in  $\mathbb{R}^m$ . In order to define the discrete Lebesgue space  $\mathcal{L}^2(\mathcal{G}_l)$ , we choose a cubic lattice in  $\mathbb{R}^m$  with grid mesh  $h \in I = (0, h_0] \subset \mathbb{R}^+$ . More precisely, for some fixed  $z_0 \in \mathbb{R}^m$  we define the set of vertices  $\mathcal{Z}_h(z_0)$  by

$$\mathcal{Z}_h(z_0) = \left\{ z \in \mathbb{R}^m : z = h z_0 + i_1 h e_1 + \dots + i_m h e_m, (i_1, \dots, i_m) \in \mathbb{Z}^m \right\}, \quad (2.26)$$

where  $e_k$  denotes the  $k$ th unit vector in  $\mathbb{R}^m$ . The  $k$ th coordinate of a vertex is thus an integer multiple of  $h$  shifted by  $h z_{0,k}$ . To each vertex  $z \in \mathcal{Z}_h(z_0)$  we assign an open cube  $c_h(z) \subset \mathbb{R}^m$  with edges parallel to the coordinate axis having edge length  $h$  and  $z$  as midpoint.

**Definition 2.1.1.** The set  $\mathcal{G}_h$  of *interior lattice points* of  $G$  generated by the lattice  $\mathcal{Z}_h(z_0)$  is defined as

$$\mathcal{G}_h = \{ z \in \mathcal{Z}_h(z_0) : c_h(z) \subset G \}.$$

**Definition 2.1.2.** By a *lattice function* we understand a function  $u_h : \mathcal{Z}_h(z_0) \rightarrow \mathbb{R}$ , i.e., a function that assigns a real number to each vertex  $z \in \mathcal{Z}_h(z_0)$ . The *extended version* of a

lattice function is the step function  $u_h : \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $x \mapsto \sum_{z \in \mathcal{Z}_h(z_0)} u_h(z) \mathbb{1}_{c_h(z)}(x)$ , where  $\mathbb{1}_{c_h(z)}$  is the indicator function of the open cube  $c_h(z)$ .

Next, we endow the space of lattice functions that vanish outside  $\mathcal{G}_h$  with a scalar product.

**Definition 2.1.3.** The *discrete Lebesgue space*  $\mathcal{L}^2(\mathcal{G}_h)$  is the space of lattice functions that are zero outside  $\mathcal{G}_h$  equipped with the scalar product

$$(u_h, v_h)_{\mathcal{L}^2(\mathcal{G}_h)} = h^m \sum_{z \in \mathcal{G}_h} u_h(z) v_h(z) = \int_{\mathbb{R}^m} u_h(x) v_h(x) dx = \int_G u_h(x) v_h(x) dx. \quad (2.27)$$

**Remark 2.1.4.** Note that a lattice function in  $\mathcal{L}^2(\mathcal{G}_h)$ , its extended version and the restriction of the extended version to the domain  $G$  are, in principle, three different objects. However, to keep our notation reasonably simple, we usually do not use different symbols.

We now come back to the stochastic example model. By taking  $z_0 = 1/2$  and grid mesh  $l$  we see that the midpoints  $z_i$ ,  $i = 1, \dots, N$ , of the cells introduced above constitute the set  $\mathcal{G}_l$  of interior lattice points of the interval  $G = (0, L)$  generated by the lattice  $\mathcal{Z}_l(z_0)$ . Hence, we can identify the particle density  $u_l$  defined above as extended version of a lattice function from the space  $\mathcal{L}^2(\mathcal{G}_l)$ . Finally, we choose as state space  $S_l$  of the particle density process  $(u_l(t))_{t \geq 0}$  the discrete subset of lattice functions in  $\mathcal{L}^2(\mathcal{G}_l)$  that take values in  $\frac{1}{n} \mathbb{N}_0$ . The norm in  $\mathcal{L}^2(\mathcal{G}_l)$  induces a metric on  $S_l$  which obviously makes it a locally compact, separable, metric space as well. The particle density process  $(u_l(t))$  is a Markov process in its own right with respect to the filtration  $(\mathcal{F}_{l,t})$  induced by itself, and we denote its transition intensities by  $q_l(\cdot, \cdot)$ . They are related to the transition intensities of the unscaled particle model by

$$q_l(u_l, \tilde{u}_l) = q(n u_l, n \tilde{u}_l), \quad (2.28)$$

where  $n u_l$  and  $n \tilde{u}_l$  are interpreted as elements of  $\mathbb{N}_0^N$ . Furthermore,  $(u_l(t))$  corresponds to a Feller semigroup generated by

$$L_l g(u_l) = \sum_{\tilde{u}_l \neq u_l} q_l(u_l, \tilde{u}_l) (g(\tilde{u}_l) - g(u_l)), \quad g \in \hat{C}(S_l), \quad (2.29)$$

and all properties of the unscaled process  $(U(t))$  discussed above (regularity of paths, measurability, Dynkin's formula, etc.) carry over to the particle density process.

### The general linear model

We immediately formulate the particle density process, i.e., the rescaled version of the mesoscopic stochastic particle model for the general linear case. We now assume that the chemical reactor is represented by a bounded domain  $G \subset \mathbb{R}^m$ . Let us choose as above a family of cubic lattices  $\mathcal{Z}_h(z_0)$  in  $\mathbb{R}^m$  with grid mesh  $h \in I = (0, h_0] \subset \mathbb{R}^+$ , and let  $\mathcal{G}_h$  be the set of

interior lattice points of  $G$  generated by  $\mathcal{Z}_h(z_0)$ . The cells now correspond to open cubes in  $\mathbb{R}^m$  with edge length  $l$  around the points  $z \in \mathcal{G}_l$ . The state space  $S_l$  of the particle density process  $\mathbf{u}_l(t) = (u_{l,1}(t), \dots, u_{l,n_s}(t))$ ,  $t \geq 0$ , is the set of vector-valued lattice functions from the space  $(\mathcal{L}^2(\mathcal{G}_l))^{n_s}$  that take values in  $\frac{1}{n}\mathbb{N}_0^{n_s}$ .

**Definition 2.1.5.** The set of lattice points  $\mathcal{G}_h^1$  is defined as

$$\mathcal{G}_h^1 = \{z \in \mathcal{G}_h : z \pm h e_k \in \mathcal{G}_h, k = 1, \dots, m\}.$$

Let, for  $z \in \mathcal{G}_l^1$  and  $j = 1, \dots, n_s$ ,  $\chi_{j,z}$  be the state with particle density one of species  $j$  in cell  $z$  and zero elsewhere. For  $z \in \mathcal{G}_l \setminus \mathcal{G}_l^1$  we define  $\chi_{j,z}$  identically zero. The transition intensities  $q_l(\cdot, \cdot)$  are now the following.

- A particle of species  $j$  may leave cell  $z \in \mathcal{G}_l^1$  and jump to  $z \pm l e_k$ .

$$\begin{aligned} q_l(\mathbf{u}_l, \mathbf{u}_l - \frac{1}{n}\chi_{j,z} + \frac{1}{n}\chi_{j,(z-le_k)}) &= n \frac{d_j}{2m} u_{l,j}(z), \\ q_l(\mathbf{u}_l, \mathbf{u}_l - \frac{1}{n}\chi_{j,z} + \frac{1}{n}\chi_{j,(z+le_k)}) &= n \frac{d_j}{2m} u_{l,j}(z), \end{aligned} \quad (2.30)$$

where  $d_j > 0$  is the hopping rate of species  $j$ .

- The number of particles in cell  $z \in \mathcal{G}_l^1$  changes according to reaction  $i$ .

$$q_l(\mathbf{u}_l, \mathbf{u}_l + \frac{1}{n}\sum_{j=1}^{n_s} \nu_{ij} \chi_{j,z}) = n K_i(\mathbf{u}_l(z)) \quad \text{if } \mathbf{u}_l + \frac{1}{n}\sum_{j=1}^{n_s} \nu_{ij} \chi_{j,z} \in S_l. \quad (2.31)$$

The intensity for other possible transitions is zero. As for the example model of the previous paragraph, the matrix of transition intensities characterises a Markov jump process in  $S_l$  that corresponds to a Feller semigroup generated by

$$L_l g(\mathbf{u}_l) = \sum_{\tilde{\mathbf{u}}_l \neq \mathbf{u}_l} q_l(\mathbf{u}_l, \tilde{\mathbf{u}}_l) (g(\tilde{\mathbf{u}}_l) - g(\mathbf{u}_l)), \quad g \in \hat{C}(S_l). \quad (2.32)$$

This follows again from Theorem 3.1 in Chapter 8 of Ethier & Kurtz (1986).

## 2.2 The macroscopic PDE and a semi-discrete approximation

### 2.2.1 Weak formulation of the PDE

Our final aim is to show that the particle density of the example model and the general linear model defined in the previous section approximate the solutions of the PDEs (2.10) and (2.12). First of all, however, we have to specify more precisely the notion of solution we are going to adopt.

### The example model

Let  $G = (0, L)$  and recall that  $Q_T = G \times (0, T)$ , where  $T > 0$  is the time of observation. The classical initial-boundary value problem with Dirichlet boundary conditions for Eq. (2.10) is to seek a sufficiently smooth function  $u$  that solves

$$\begin{cases} \partial_t u - D \partial_x^2 u = k_1 u - k_2 u & \text{on } Q_T \\ u = 0 & \text{on } \partial G \times [0, T] \\ u(\cdot, 0) = u_0(\cdot) & \text{on } G. \end{cases} \quad (2.33)$$

Since, in general, the classical point of view is too restrictive, we turn to a generalised formulation of the problem. The weak formulation (also called variational formulation) is obtained as follows. Assume for the following calculations that  $u$  is a sufficiently smooth classical solution. After multiplying the equation with a test function  $v \in C_0^\infty(G)$  and integrating over  $G$  we get by an integration by parts that

$$\frac{d}{dt} \int_G u v \, dx + D \int_G \partial_x u \partial_x v \, dx = (k_1 - k_2) \int_G u v \, dx \quad (2.34)$$

for all  $v \in C_0^\infty(G)$  and  $t \in (0, T)$ . If, in turn, a sufficiently smooth function  $u$  solves (2.34), it is also a solution of (2.33). Note that the integral terms

$$a(u, v) = D \int_G \partial_x u \partial_x v \, dx \quad (2.35)$$

and

$$(f(u), v)_{L^2(G)} = (k_1 - k_2) \int_G u v \, dx \quad (2.36)$$

can be viewed as bilinear forms on  $H_0^1(G) \times H_0^1(G)$ . Here  $H_0^1(G)$  denotes (also for a general bounded domain in  $\mathbb{R}^m$  with Lipschitz boundary) the Sobolev space of functions in  $L^2(G)$  that have weak derivatives in  $L^2(G)$  and vanish on the boundary of  $G$  in the trace sense (see, e.g., Zeidler (1990b), Chapter 21). Recall that the spaces  $H_0^1(G)$ ,  $L^2(G)$  and  $H^{-1}(G) = (H_0^1(G))^*$  form a Gelfand triple:

$$H_0^1(G) \hookrightarrow L^2(G) \cong (L^2(G))^* \hookrightarrow H^{-1}(G) \quad (2.37)$$

(see, e.g., Zeidler (1990b), Chapter 23). The space  $H^1(0, T; H_0^1(G), L^2(G))$  is defined as the space of functions in  $L^2(0, T; H_0^1(G))$  that have generalised time derivatives in  $L^2(0, T; H^{-1}(G))$ . Here and in the sequel integrals of Banach-space valued functions are always interpreted in the sense of the Bochner integral (see, e.g., Da Prato & Zabczyk (1992)). The notation indicates that  $H_0^1(G)$  is embedded into  $H^{-1}(G)$  via  $L^2(G)$  as in (2.37). In the following we often skip the domain  $G$  in the notation if there is no risk of ambiguity. The generalised time derivative of a function  $u \in L^2(0, T; H_0^1)$  is defined here as a function  $u'$  in  $L^2(0, T; H^{-1})$  such that the equation

$$\int_0^T \langle u'(t), v \rangle_{H_0^1} \varphi(t) dt = - \int_0^T (u(t), v)_{L^2} \varphi'(t) dt \quad (2.38)$$

is satisfied for arbitrary  $\varphi \in C_0^\infty(0, T)$  and  $v \in H_0^1$ . Let  $b(\cdot, \cdot)$  be the bilinear form given by

$$b(u, v) = a(u, v) - (f(u), v)_{L^2}, \quad u, v \in H_0^1. \quad (2.39)$$

The weak problem corresponding to (2.33) is then (cf. Eq. (2.34)) to look for a function  $u \in H^1(0, T; H_0^1, L^2)$  that satisfies

$$\frac{d}{dt} (u(t), v)_{L^2} + b(u(t), v) = 0 \quad (2.40a)$$

for all  $v \in H_0^1$  and a.e.  $t \in (0, T)$ , and

$$u(0) = u_0 \in L^2. \quad (2.40b)$$

Here the equation involving the time derivative is supposed to hold in  $\mathcal{D}'(0, T)$ , i.e., in the scalar distribution sense. Note that the bilinear form  $b(\cdot, \cdot)$  is bounded and satisfies a Gårding inequality:

$$b(u, u) \geq \alpha \|u\|_{H_0^1}^2 - \beta \|u\|_{L^2}^2 \quad (2.41)$$

for constants  $\alpha, \beta > 0$ . This follows from the the linearity of the reaction function  $f$  and the fact that the bilinear form  $a(\cdot, \cdot)$  is bounded and coercive, i.e.,

$$a(u, v) \leq C \|u\|_{H_0^1} \|v\|_{H_0^1} \quad (2.42)$$

and

$$a(u, u) \geq \alpha \|u\|_{H_0^1}^2 \quad (2.43)$$

for constants  $C, \alpha > 0$ . These inequalities, in turn, follow immediately from the Poincaré inequality. Hence a general result on linear first-order evolution equations (see, e.g., Corollary 23.26 to Theorem 23.A in Zeidler (1990b)) ensures that (2.40) has a unique solution. Since the space  $H^1(0, T; H_0^1, L^2)$  is continuously embedded in  $C([0, T], L^2)$ , the initial condition makes sense.

The weak problem (2.40) can be interpreted as operator equation. To see this note that the bilinear form  $a(\cdot, \cdot)$  induces a bounded linear operator  $A : H_0^1 \rightarrow H^{-1}$  by

$$\langle Au, v \rangle_{H_0^1} = a(u, v), \quad u, v \in H_0^1. \quad (2.44)$$

The operator  $A$  is (strongly) *monotone*, i.e.,

$$\langle Au - Av, u - v \rangle_{H_0^1} \geq \alpha \|u - v\|_{H_0^1}^2 \geq 0 \quad \text{for all } u, v \in H_0^1. \quad (2.45)$$

Moreover, the reaction function  $f$  induces a bounded linear operator (a so-called *Nemyckii operator*, see, e.g., Chapter 26 of Zeidler (1990c))  $F : L^2 \rightarrow L^2$  by  $(Fu)(\cdot) = f(u(\cdot))$ . Therefore the operator equation

$$u' + Au = Fu, \quad u(0) = u_0 \in L^2, \quad (2.46)$$

in the space  $H^1(0, T; H_0^1, L^2)$  is equivalent to the weak formulation (2.40) above. For more details see Zeidler (1990b), Chapter 23.

The general linear model

We now suppose that  $G$  is a bounded domain in  $\mathbb{R}^m$  with Lipschitz boundary. Recall that we are dealing with  $n_r$  equations for  $n_s$  species and that the initial-boundary value problem for the PDE system (2.12) with Dirichlet boundary conditions is given by

$$\begin{cases} \partial_t u_j - D_j \Delta u_j = f_j(\mathbf{u}) & \text{on } Q_T \\ u_j = 0 & \text{on } \partial G \times [0, T] \\ u_j(\cdot, 0) = u_{j,0} & \text{on } G, \end{cases} \quad (2.47)$$

$j = 1, \dots, n_s$ . The weak formulation of (2.47) is completely analogous to the previous paragraph. To facilitate notation we set

$$\mathbf{H}_0^1(G) = (H_0^1(G))^{n_s}, \quad \mathbf{L}^2(G) = (L^2(G))^{n_s}, \quad \mathbf{H}^{-1}(G) = ((H_0^1(G))^{n_s})^*.$$

Note that again the spaces  $\mathbf{H}_0^1(G)$ ,  $\mathbf{L}^2(G)$  and  $\mathbf{H}^{-1}(G)$  form a Gelfand triple. We define the bilinear forms

$$a(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^{n_s} D_j (\nabla u_j, \nabla v_j)_{(L^2)^m} = \sum_{j=1}^{n_s} \sum_{k=1}^m D_j (\partial_{x_k} u_j, \partial_{x_k} v_j)_{L^2}, \quad \mathbf{u}, \mathbf{v} \in \mathbf{H}_0^1, \quad (2.48)$$

and

$$b(\mathbf{u}, \mathbf{v}) = a(\mathbf{u}, \mathbf{v}) - (\mathbf{f}(\mathbf{u}), \mathbf{v})_{\mathbf{L}^2}, \quad \mathbf{u}, \mathbf{v} \in \mathbf{H}_0^1, \quad (2.49)$$

where  $\mathbf{f} = (f_1, \dots, f_{n_s})$ . The generalised time derivative of a function  $\mathbf{u} \in L^2(0, T; \mathbf{H}_0^1)$  is, of course, the element  $\mathbf{u}'$  of  $L^2(0, T; \mathbf{H}^{-1})$  that satisfies

$$\int_0^T \langle \mathbf{u}'(t), \mathbf{v} \rangle_{\mathbf{H}_0^1} \varphi(t) dt = - \int_0^T (\mathbf{u}(t), \mathbf{v})_{\mathbf{L}^2} \varphi'(t) dt, \quad (2.50)$$

for arbitrary  $\varphi \in C_0^\infty(0, T)$  and  $\mathbf{v} \in \mathbf{H}_0^1$ , and the space  $H^1(0, T; \mathbf{H}_0^1, \mathbf{L}^2)$  is defined in analogy to the scalar case. In the weak formulation of the PDE a function  $\mathbf{u} \in H^1(0, T; \mathbf{H}_0^1, \mathbf{L}^2)$  is sought that satisfies

$$\frac{d}{dt} (\mathbf{u}(t), \mathbf{v})_{\mathbf{L}^2} + b(\mathbf{u}(t), \mathbf{v}) = 0 \quad (2.51a)$$

for all  $v \in \mathbf{H}_0^1$  and a.e.  $t \in (0, T)$ , and

$$\mathbf{u}(0) = \mathbf{u}_0 \in \mathbf{L}^2. \quad (2.51b)$$

In the same way as for the example model, the bilinear form  $b(\cdot, \cdot)$  is bounded and satisfies a Gårding inequality, which ensures the existence of a unique solution.

Again the bilinear form  $a(\cdot, \cdot)$  induces a bounded linear and (strongly) monotone operator  $A : \mathbf{H}_0^1 \rightarrow \mathbf{H}^{-1}$ . Since the reaction functions  $f_j$ ,  $j = 1, \dots, n_s$ , are still linear, the operator  $F : \mathbf{L}^2 \rightarrow \mathbf{L}^2$  given by  $(F\mathbf{u})(\cdot) = \mathbf{f}(\mathbf{u}(\cdot))$  is linear and bounded, and the equivalent operator equation in  $H^1(0, T; \mathbf{H}_0^1, \mathbf{L}^2)$  reads

$$\mathbf{u}' + A\mathbf{u} = F\mathbf{u}, \quad \mathbf{u}(0) = \mathbf{u}_0 \in \mathbf{L}^2. \quad (2.52)$$

### 2.2.2 A semi-discrete finite-difference approximation

In this section we discuss spatially semi-discretised approximating equations for the macroscopic PDEs (2.33) and (2.12) introduced in the previous section. In Sections 2.3 and 2.4 we shall show that their solutions are indeed approximations, i.e., that they converge to the solution of the respective PDE if the grid mesh tends to zero. The solutions of the approximating equations serve as auxiliary functions and correspond to the expected value of the particle density regarded as  $L^2$ -valued random variable. Their introduction allows, as we will see, to deal separately with the spatial discreteness and the randomness of the particle density.

We have already introduced discrete Lebesgue spaces in Section 2.1.2. The particle densities introduced there have been identified as elements of the discrete Lebesgue space  $\mathcal{L}^2(\mathcal{G}_l)$ , where  $l$  was the edge length of a cell. For an appropriate formulation of the approximating problems we need, in addition, a discrete version of the Sobolev space  $H_0^1(G)$ . Therefore we now introduce the discrete Sobolev space  $\mathcal{H}_0^1(\mathcal{G}_h)$  and discuss some of its properties (cf. Chapter 35 of Zeidler (1990c)).

#### Discrete Sobolev spaces

Let, as in Section 2.1.2,  $\mathcal{Z}_h(z_0)$  be a lattice in  $\mathbb{R}^m$  with grid mesh  $h \in I = (0, h_0] \subset \mathbb{R}^+$  and  $c_h(z)$  the open cube with midpoint  $z$  having edge length  $h$ . The lattice functions on  $\mathcal{Z}_h(z_0)$ , the interior lattice points  $\mathcal{G}_h$  of a bounded domain  $G \subset \mathbb{R}^m$ , and the discrete Lebesgue space  $\mathcal{L}^2(\mathcal{G}_h)$  have already been defined in Section 2.1.2 as well as the lattice points  $\mathcal{G}_h^1$ . We now define the discrete derivatives of a lattice function.

**Definition 2.2.1.** For a lattice function  $u_h$  the *discrete derivatives*  $\partial_k^+ u_h$  and  $\partial_k^- u_h$  are defined as the lattice functions given by

$$\partial_k^\pm u_h(z) = \frac{u_h(z \pm h e_k) - u_h(z)}{\pm h}, \quad k = 1, \dots, m.$$

Higher derivatives are obtained by repeated application of  $\partial_k^\pm$ .

We are now in a position to define the discrete Sobolev space  $\mathcal{H}_0^1(\mathcal{G}_h)$ .

**Definition 2.2.2.** By the *discrete Sobolev space*  $\mathcal{H}_0^1(\mathcal{G}_h)$  we understand the set of all lattice functions that vanish outside  $\mathcal{G}_h^1$  equipped with the scalar product

$$(u_h, v_h)_{\mathcal{H}_0^1(\mathcal{G}_h)} = (u_h, v_h)_{\mathcal{L}^2(\mathcal{G}_h)} + \sum_{k=1}^m (\partial_k^+ u_h, \partial_k^+ v_h)_{\mathcal{L}^2(\mathcal{G}_h)}.$$

The space  $\mathcal{H}_0^1(\mathcal{G}_h)$  has many properties in common with the Sobolev  $H_0^1(G)$  defined on a continuous domain  $G$ , e.g., we have a discrete integration by parts formula.

**Lemma 2.2.3.** For functions  $u_h, v_h \in \mathcal{H}_0^1(\mathcal{G}_h)$  we have

$$(\partial_k^+ u_h, v_h)_{\mathcal{L}^2} = -(u_h, \partial_k^- v_h)_{\mathcal{L}^2}, \quad k = 1, \dots, m.$$

**Proof.** The proof is (really!) straightforward. □

Moreover, there is a discrete analogue of Poincaré's inequality.

**Lemma 2.2.4.** For functions  $u_h, v_h \in \mathcal{H}_0^1(\mathcal{G}_h)$  we have

$$(u_h, v_h)_{\mathcal{H}_0^1} \leq C (\nabla^+ u_h, \nabla^+ v_h)_{(\mathcal{L}^2)^m},$$

and the constant  $C$  depends only on the domain  $G$ .

**Proof.** Cf. Proposition 3.3 in Chapter 1 of Temam (2001). □

Note that, in analogy to the spaces defined on a continuous domain,  $\mathcal{H}_0^1(\mathcal{G}_h)$ ,  $\mathcal{L}^2(\mathcal{G}_h)$  and  $\mathcal{H}^{-1}(\mathcal{G}_h) = (\mathcal{H}_0^1(\mathcal{G}_h))^*$  form a Gelfand triple:

$$\mathcal{H}_0^1(\mathcal{G}_h) \hookrightarrow \mathcal{L}^2(\mathcal{G}_h) \cong (\mathcal{L}^2(\mathcal{G}_h))^* \hookrightarrow \mathcal{H}^{-1}(\mathcal{G}_h). \quad (2.53)$$

It is easily checked that the embedding constants are bounded independent of  $h$ . Of course, since the domain  $G$  is supposed to be bounded and thus  $\mathcal{H}_0^1(\mathcal{G}_h)$  is finite-dimensional, the reverse inclusions hold as well. However, the embedding constants of the reverse inclusions are *not* uniformly bounded.

We are now going to formulate approximating problems for the macroscopic PDEs of the example model and the general linear model introduced in Section 2.2.1.

The approximating problem for the example model

Let  $\mathcal{G}_h$  be the interior lattice points of  $G = (0, L)$  as in Section 2.1.2. The discrete analogue of the initial-boundary value problem (2.33) is to find, for a given lattice function  $u_{h,0}$  and diffusion coefficient  $D_h > 0$ , the functions  $u_h(z, \cdot) : [0, T] \rightarrow \mathbb{R}$ ,  $z \in \mathcal{G}_h$ , such that

$$\begin{cases} u_h' - D_h \partial^- \partial^+ u_h = f(u_h) & \text{on } \mathcal{G}_h^1 \times (0, T) \\ u_h = 0 & \text{on } (\mathcal{G}_h \setminus \mathcal{G}_h^1) \times [0, T] \\ u_h(\cdot, 0) = u_{h,0} & \text{on } \mathcal{G}_h^1. \end{cases} \quad (2.54)$$

This is in fact an initial value problem for a linear finite-dimensional system of ODEs. Consequently, the existence of a unique solution for all times is ensured by the Picard-Lindelöf theorem. It will turn out to be useful to introduce a ‘discrete weak formulation’ in analogy to the weak formulation of the original PDE (2.40). Note that the expressions

$$a_h(u_h, v_h) = D_h(\partial^+ u_h, \partial^+ v_h)_{\mathcal{H}_0^1}, \quad u_h, v_h \in \mathcal{H}_0^1, \quad (2.55)$$

and

$$(f(u_h), v_h)_{\mathcal{L}^2} = (k_1 - k_2)(u_h, v_h)_{\mathcal{L}^2}, \quad u_h, v_h \in \mathcal{H}_0^1, \quad (2.56)$$

are bilinear forms on  $\mathcal{H}_0^1 \times \mathcal{H}_0^1$ . Furthermore, we define the bilinear form  $b_h(\cdot, \cdot)$  by

$$b_h(u_h, v_h) = a_h(u_h, v_h) - (f(u_h), v_h)_{\mathcal{L}^2}, \quad u_h, v_h \in \mathcal{H}_0^1. \quad (2.57)$$

The solution of the approximating problem (2.54) can be regarded as an element of the space  $C^1([0, T], \mathcal{H}_0^1)$ , and it also solves the following weak formulation:

$$\frac{d}{dt}(u_h(t), v_h)_{\mathcal{L}^2} + b_h(u_h(t), v_h) = 0 \quad (2.58)$$

for all  $v_h \in \mathcal{H}_0^1$  and  $t \in (0, T)$ . This follows from (2.54) by multiplying with  $h^m v_h(z)$ , summing over all  $z \in \mathcal{G}_h$  and a discrete integration by parts.

As a consequence of the discrete Poincaré inequality the bilinear form  $a_h(\cdot, \cdot)$  is bounded and coercive. Therefore, in analogy to the corresponding bilinear form  $a(\cdot, \cdot)$  of the PDE problem, it induces a bounded linear and (strongly) monotone operator  $A_h : \mathcal{H}_0^1 \rightarrow \mathcal{H}^{-1}$  by

$$\langle A_h u_h, v_h \rangle_{\mathcal{H}_0^1} = a_h(u_h, v_h), \quad u_h, v_h \in \mathcal{H}_0^1. \quad (2.59)$$

### The approximating problem for the general linear model

Let now  $\mathcal{G}_h$  be the interior lattice points generated by a lattice  $\mathcal{Z}_h(z_0)$  of a general bounded domain  $G \subset \mathbb{R}^m$  with Lipschitz boundary representing the chemical reactor (cf. Section 2.1.2). Recall that we are dealing with  $n_r$  equations for  $n_s$  species. To facilitate notation, we set

$$\mathcal{H}_0^1(\mathcal{G}_h) = (\mathcal{H}_0^1(\mathcal{G}_h))^{n_s}, \quad \mathcal{L}^2(\mathcal{G}_h) = (\mathcal{L}^2(\mathcal{G}_h))^{n_s}, \quad \mathcal{H}^{-1}(\mathcal{G}_h) = ((\mathcal{H}_0^1(\mathcal{G}_h))^{n_s})^*.$$

The discrete analogue of the PDE system (2.47) is the ODE system

$$\begin{cases} u'_{h,j} - D_{h,j} \nabla^- \cdot \nabla^+ u_{h,j} = f_j(\mathbf{u}_h) & \text{on } \mathcal{G}_h^1 \times (0, T) \\ u_{h,j} = 0 & \text{on } (\mathcal{G}_h \setminus \mathcal{G}_h^1) \times [0, T] \\ u_{h,j}(\cdot, 0) = u_{h,j,0} & \text{on } \mathcal{G}_h^1, \end{cases} \quad (2.60)$$

$j = 1, \dots, n_s$ , where  $\mathbf{u}_h = (u_{h,1}, \dots, u_{h,n_s})$ . This is still an initial-value problem for a linear finite-dimensional system of ODEs and thus solvability for all times and uniqueness of solutions is ensured by the Picard-Lindelöf theorem.

In analogy to the previous paragraph we have on  $\mathcal{H}_0^1 \times \mathcal{H}_0^1$  the bilinear forms

$$a_h(\mathbf{u}_h, \mathbf{v}_h) = \sum_{j=1}^{n_s} \sum_{k=1}^m D_{h,j} (\partial_k^+ u_{h,j}, \partial_k^+ v_{h,j})_{\mathcal{L}^2}, \quad \mathbf{u}_h, \mathbf{v}_h \in \mathcal{H}_0^1, \quad (2.61)$$

and

$$b_h(\mathbf{u}_h, \mathbf{v}_h) = a_h(\mathbf{u}_h, \mathbf{v}_h) - (\mathbf{f}(\mathbf{u}_h), \mathbf{v}_h)_{\mathcal{L}^2}, \quad \mathbf{u}_h, \mathbf{v}_h \in \mathcal{H}_0^1. \quad (2.62)$$

The solution  $\mathbf{u}_h$  of (2.60) can be regarded as a function in  $C^1([0, T], \mathcal{H}_0^1)$ , and it also solves the discrete weak problem

$$\frac{d}{dt} (\mathbf{u}_h(t), \mathbf{v}_h)_{\mathcal{L}^2} + b_h(\mathbf{u}_h(t), \mathbf{v}_h) = 0 \quad (2.63)$$

for all  $\mathbf{v}_h \in \mathcal{H}_0^1$  and  $t \in (0, T)$ . Note that once more the operator  $A_h : \mathcal{H}_0^1 \rightarrow \mathcal{H}^{-1}$  given by

$$\langle A_h \mathbf{u}_h, \mathbf{v}_h \rangle_{\mathcal{H}_0^1} = a_h(\mathbf{u}_h, \mathbf{v}_h), \quad \mathbf{u}_h, \mathbf{v}_h \in \mathcal{H}_0^1, \quad (2.64)$$

is bounded, linear and (strongly) monotone due to the discrete Poincaré inequality.

### 2.2.3 External approximation schemes

In this section we review some important techniques that we need for the first part of the proof of the law of large numbers, which consists in proving that the approximations discussed in the previous section actually converge. We introduce the general concept of *external approximation* of a normed vector space following Temam (1973, 2001) and Zeidler (1990c). External approximations are motivated, for instance, by the study of convergence of finite difference methods for weak solutions of PDEs as introduced in Sections 2.2.1 and 2.2.2. Recall that the solution  $u_h(\cdot, t)$  of the approximating problem for the example model and its discrete spatial derivatives  $\partial_k^+ u_h(\cdot, t)$  are step functions. The solution of the original equation, on the other hand, is at each moment in time a function in the Sobolev space  $H_0^1$ . Therefore it is not obvious how to define and estimate the error between the solution and its approximation because they do not belong to the same function space. (The same sort of problem can arise also with finite-element methods if one uses so-called non-conforming finite elements.) The concept of external approximation provides a framework to handle this kind of difficulty.

#### External approximations

In this paragraph we discuss the general concept of external approximation of a normed vector space denoted by  $W$ . Elements of  $W$  are approximated by elements from a family of normed vector spaces  $\mathcal{W}_h$ ,  $h \in I = (0, h_0] \subset \mathbb{R}^+$ , which are finite-dimensional in practice. The parameter  $h$  should be thought of as a grid mesh.

**Definition 2.2.5.** An *external approximation* of  $W$  consists of

- (i) a normed vector space  $X$  and an isometric embedding  $J : W \rightarrow X$ ,
- (ii) a family of triples  $(\mathcal{W}_h, R_h, J_h)_{h \in I}$ , where for each  $h$ 
  - $\mathcal{W}_h$  is a normed vector space,
  - $R_h$  is an (arbitrary) operator from  $W$  to  $\mathcal{W}_h$ ,
  - $J_h$  is a bounded linear operator from  $\mathcal{W}_h$  to  $X$ .

The operators  $R_h$  and  $J_h$  are called the *restriction operator* and the *extension operator*, respectively. This terminology is useful for solving problems of the kind discussed above: the object we want to approximate, e.g., the solution of a PDE, is an element of the space  $W$ , whereas the approximation belongs to the space  $\mathcal{W}_h \not\subset W$ . In order to compare solution and approximation both have to be mapped to a third space  $X$ . If  $X = W$  and  $J = \text{id}$ , we speak of an *internal approximation*.

**Definition 2.2.6.** For  $u \in W$  and  $u_h \in \mathcal{W}_h$  we define

- (i)  $\|Ju - J_h u_h\|_X$  as the *error* between  $u$  and  $u_h$ ,
- (ii)  $\|R_h u - u_h\|_{\mathcal{W}_h}$  as the *discrete error* between  $u$  and  $u_h$ ,
- (iii)  $\|Ju - J_h R_h u\|_X$  as the *truncation error* for  $u$ .

The following notions are useful for discussions of stability and consistency.

**Definition 2.2.7.** The family of extension operators  $(J_h)_{h \in I}$  is called *stable* if the norms of the operators  $J_h$  are uniformly bounded, i.e.,

$$\sup_{h \in I} \|J_h\|_{\mathcal{L}(\mathcal{W}_h, X)} < \infty.$$

The external approximation of  $W$  consisting of  $X$ ,  $J$ , and  $(\mathcal{W}_h, R_h, J_h)_{h \in I}$  is called *stable* if the family  $(J_h)_{h \in I}$  is stable.

Stability of the restriction operators is defined in the same way and may be useful, but is not always necessary and thus not part of the definition of a stable external approximation.

**Remark 2.2.8.** In the following we often sloppily speak of a sequence  $(u_h)$ ,  $h \searrow 0$ , of elements of  $\mathcal{W}_h$ , by which we mean a sequence  $(u_{h_i})_{i \in \mathbb{N}_0}$  of elements  $u_{h_i} \in \mathcal{W}_{h_i}$ , where  $(h_i)_{i \in \mathbb{N}_0}$  is a sequence in  $I$  that converges to zero for  $i \rightarrow \infty$ .

**Definition 2.2.9.** An external approximation of a normed vector space  $W$  is said to be *convergent* if the following conditions hold:

(i) For all  $u \in W$  the truncation error converges to zero for  $h \rightarrow 0$ , i.e.,

$$\|Ju - J_h R_h u\|_X \rightarrow 0 \quad (h \rightarrow 0).$$

(ii) If  $(u_h)$ ,  $h \searrow 0$ , is a sequence in  $\mathcal{W}_h$  such that  $J_h u_h$  converges weakly to  $\tilde{u}$  in  $X$ , then  $\tilde{u} \in J(W)$ , i.e., since  $J$  is one-to-one,  $\tilde{u} = Ju$  for a unique  $u \in W$ .

The second condition comes into play if the embedding  $J$  is not surjective. For internal approximations it is apparently superfluous.

**Remark 2.2.10.** For a stable and convergent external approximation, discrete convergence implies convergence, i.e., if  $(u_h)$ ,  $h \searrow 0$ , is a sequence in  $\mathcal{W}_h$ , and the discrete error between  $u_h$  and  $u \in W$  converges to zero, then the error vanishes as well.

**Proof.** The claim follows immediately from the estimate

$$\|Ju - J_h u_h\|_X \leq \|Ju - J_h R_h u\|_X + \|J_h\|_{\mathcal{L}(\mathcal{W}_h, X)} \|R_h u - u_h\|_{\mathcal{W}_h}$$

and the definitions. □

The next lemma is important in connection with evolution problems. We assume in addition that  $W$ ,  $\mathcal{W}_h$  and  $X$  are separable Banach spaces.

**Lemma 2.2.11.** *Let  $\{X, J, (\mathcal{W}_h, R_h, J_h)_{h \in I}\}$  be a stable and convergent external approximation of  $W$ , and assume that the sequence of functions  $u_h : [0, T] \rightarrow \mathcal{W}_h$  satisfies*

$$J_h u_h \rightharpoonup \tilde{u} \quad (h \searrow 0)$$

*in  $L^2(0, T; X)$ . Then there is a function  $u \in L^2(0, T; W)$  so that  $\tilde{u}(t) = Ju(t)$  for a.e.  $t$ .*

**Proof.** See Temam (2001), p. 238. □

Approximation of Lebesgue and Sobolev spaces by their discrete analogues

Let  $G \subset \mathbb{R}^m$  be a bounded domain with Lipschitz boundary. In this paragraph we discuss an internal approximation of  $L^2(G)$  and an external approximation of  $H_0^1(G)$  in terms of the discrete spaces  $\mathcal{L}^2(\mathcal{G}_h)$  and  $\mathcal{H}_0^1(\mathcal{G}_h)$ , where  $h \in I = (0, h_0] \subset \mathbb{R}^+$ . For functions  $u \in L^2(G)$  we define the lattice function  $\bar{u}_h$  as follows. We first extend  $u$  to a function  $\tilde{u} \in L^2(\mathbb{R}^m)$  by setting  $\tilde{u} = u$  on  $G$  and  $\tilde{u} \equiv 0$  on  $\mathbb{R}^m \setminus G$ . The lattice function  $\bar{u}_h$  is then defined as

$$\bar{u}_h(z) = h^{-m} \int_{c_h(z)} u(x) dx, \quad z \in \mathcal{Z}_h(z_0). \quad (2.65)$$

The restriction operator  $Q_h : L^2(G) \rightarrow \mathcal{L}^2(\mathcal{G}_h)$  is given by

$$Q_h u(z) = \begin{cases} \bar{u}_h(z) & \text{if } z \in \mathcal{G}_h \\ 0 & \text{otherwise,} \end{cases} \quad (2.66)$$

and the restriction operator  $R_h : H_0^1(G) \rightarrow \mathcal{H}_0^1(\mathcal{G}_h)$  is defined as

$$R_h u(z) = \begin{cases} \bar{u}_h(z) & \text{if } z \in \mathcal{G}_h^1 \\ 0 & \text{otherwise.} \end{cases} \quad (2.67)$$

They have the following properties.

**Lemma 2.2.12.** *The restriction operators  $Q_h$  and  $R_h$  are bounded linear operators, i.e., they are elements of  $\mathcal{L}(L^2(G), \mathcal{L}^2(\mathcal{G}_h))$  and  $\mathcal{L}(H_0^1(G), \mathcal{H}_0^1(\mathcal{G}_h))$ , respectively. Moreover, the families  $(Q_h)_{h \in I}$  and  $(R_h)_{h \in I}$  are stable, and  $\sup_h \|R_h\|_{\mathcal{L}(H_0^1, \mathcal{H}_0^1)} \leq 1$ .*

**Proof.** Let  $u_h = Q_h u$  for an arbitrary function  $u \in L^2$ .

$$\begin{aligned} \|u_h\|_{\mathcal{L}^2}^2 &= h^m \sum_{z \in \mathcal{G}_h} u_h(z)^2 = h^m \sum_{z \in \mathcal{G}_h} \left( h^{-m} \int_{c_h(z)} u(x) dx \right)^2 \\ &\leq h^m \sum_{z \in \mathcal{G}_h} \left( h^{-m} \left( \int_{c_h(z)} u(x)^2 dx \right)^{1/2} h^{m/2} \right)^2 \\ &= \sum_{z \in \mathcal{G}_h} \int_{c_h(z)} u(x)^2 dx \leq \|u\|_{L^2}^2. \end{aligned}$$

To finish the proof we may assume that  $u \in C_0^\infty(G)$  (because  $C_0^\infty(G)$  is dense in  $H_0^1(G)$ ), and we set  $u_h = R_h u$ . For  $k = 1, \dots, m$ ,

$$\begin{aligned} \|\partial_k^+ u_h\|_{\mathcal{L}^2}^2 &= h^m \sum_{z \in \mathcal{G}_h} \frac{1}{h^2} (u_h(z + h e_k) - u_h(z))^2 \\ &= h^m h^{-2} \sum_{z \in \mathcal{G}_h} \left| h^{-m} \int_{c_h(z)} (u(x + h e_k) - u(x)) dx \right|^2 \\ &= h^{-2} h^{-m} \sum_{z \in \mathcal{G}_h} \left| \int_{c_h(z)} \int_0^1 \frac{d}{dt} u(x + t h e_k) dt dx \right|^2 \\ &= h^{-m} \sum_{z \in \mathcal{G}_h} \left| \int_{c_h(z)} \int_0^1 \partial_{x_k} u(x + t h e_k) dt dx \right|^2 \\ &\leq \sum_{z \in \mathcal{G}_h} \int_0^1 \int_{c_h(z)} |\partial_{x_k} u(x + t h e_k)|^2 dx dt \leq \|\partial_{x_k} u\|_{L^2}^2. \end{aligned}$$

□

The proofs of the next two theorems, which are rather technical, are omitted. We refer the reader to Raviart (1967) or Temam (1973, 2001), where similar results are derived for symmetric difference operators. Note that the restriction to  $G$  of the extended version of a lattice function  $u_h$  is in  $L^2(G)$  and

$$\|u_h|_G\|_{L^2(G)} = \|u_h\|_{\mathcal{L}^2(\mathcal{G}_h)}. \quad (2.68)$$

In this sense  $\mathcal{L}^2(\mathcal{G}_h)$  can be regarded as a linear subspace of  $L^2(G)$ , and we define the extension operators  $P_h : \mathcal{L}^2(\mathcal{G}_h) \rightarrow L^2(G)$  by

$$P_h u_h = u_h|_G. \quad (2.69)$$

**Theorem 2.2.13.** *The internal approximation of  $L^2(G)$  by  $(\mathcal{L}^2(\mathcal{G}_h), Q_h, P_h)_{h \in I}$  is stable and convergent.*

**Proof.** Cf. Raviart (1967), Chapter 0, or (Temam, 1973), Chapter 9.  $\square$

We now turn to the announced external approximation of the Sobolev space  $H_0^1(G)$ . To this end, we set  $X = (L^2(G))^{m+1}$  and endow it with the canonical scalar product. The embedding  $J : H_0^1(G) \rightarrow X$  is given by

$$u \mapsto Ju = (u, \partial_{x_1} u, \dots, \partial_{x_m} u), \quad (2.70)$$

where  $\partial_{x_k} u$  denotes the  $k$ th weak derivative. Apparently,  $J$  is an isometric embedding. The extension operators  $J_h : \mathcal{H}_0^1(\mathcal{G}_h) \rightarrow X$  are given by

$$u_h \mapsto J_h u_h = (u_h|_G, \partial_1^+ u_h|_G, \dots, \partial_m^+ u_h|_G). \quad (2.71)$$

**Theorem 2.2.14.** *The external approximation of  $H_0^1(G)$  by  $X$ ,  $J$ , and  $(\mathcal{H}_0^1(\mathcal{G}_h), R_h, J_h)_{h \in I}$  is stable and convergent.*

**Proof.** Cf. Raviart (1967), Chapter 0, or (Temam, 1973), Chapter 9.  $\square$

## 2.3 Law of large numbers for the example model

In the present section we shall eventually put to work all the machinery that has been introduced so far. As already announced in the introduction to this chapter, the proof of the law of large numbers proceeds in two steps. We first show convergence of the semi-discrete approximation to the solution of the PDE and then estimate the difference between the approximation and the stochastic particle density. This procedure allows to deal separately with the spatial discreteness and the randomness of the particle density. The first part of the proof is thus purely analytical; it is only in the second part that probabilistic arguments come into play.

### 2.3.1 Convergence of the approximation

We first prove weak convergence of the solutions of the approximating problem (2.54) introduced in Section 2.2.2 to the solution of the weak PDE problem (2.40) from Section 2.2.1. We shall reuse the notation from Sections 2.1 and 2.2 concerning the example model. (Recall, in particular, Remark 2.1.4 on lattice functions.) Here  $G$  is the interval  $(0, L)$  and the space  $X$  is given by  $(L^2(G))^2$ . The following lemma states two consistency conditions that are necessary to show convergence of the approximations.

**Lemma 2.3.1.** *If  $(u_h)$  and  $(v_h)$ ,  $h \searrow 0$ , are sequences in  $L^2(0, T; \mathcal{H}_0^1)$  such that  $J_h u_h$  converges weakly to  $Ju$ , and  $J_h v_h \rightarrow Jv$  strongly in  $L^2(0, T; X)$  for functions  $u, v \in L^2(0, T; H_0^1)$ , then*

$$\int_0^T (u_h(t), v_h(t))_{L^2} dt \rightarrow \int_0^T (u(t), v(t))_{L^2} dt,$$

and

$$\int_0^T b_h(u_h(t), v_h(t)) dt \rightarrow \int_0^T b(u(t), v(t)) dt.$$

**Proof.** By the definition of the external approximation of the Sobolev space  $H_0^1$  in terms of the discrete spaces  $\mathcal{H}_0^1$ ,  $J_h u_h \rightharpoonup Ju$  in  $L^2(0, T; X)$  implies that  $u_h \rightharpoonup u$  and  $\partial^+ u_h \rightharpoonup \partial_x u$  in  $L^2(0, T; L^2)$ . Moreover,  $v_h$  converges strongly to  $v$ , and  $\partial^+ v_h$  converges strongly to  $\partial_x v$  in  $L^2(0, T; L^2)$ . The claim thus follows from a well-known limit relation for time-integrals of functions in  $L^2(0, T; L^2)$  (see, e.g., Zeidler (1990b), Proposition 23.9).  $\square$

We are now going to prove a first convergence theorem.

**Theorem 2.3.2.** *Let  $(u_h)$ ,  $h \searrow 0$ , be a sequence of solutions of (2.54) to the initial values  $u_{h,0}$ . Assume that  $u_{h,0}$  converges strongly to  $u_0$  in  $L^2$  and that  $D_h \rightarrow D$ . Then  $J_h u_h$  converges weakly to  $Ju$  in  $L^2(0, T; X)$ , where  $u$  is the solution of the weak PDE problem (2.40) to the initial value  $u_0$ .*

**Proof.** The proof is similar to the corresponding convergence proof for the Faedo-Galerkin method.

1. We first derive the a-priori estimates

$$\sup_h \max_{0 \leq t \leq T} \|u_h\|_{\mathcal{L}^2} < \infty, \quad (2.72)$$

$$\sup_h \|u_h\|_{L^2(0, T; \mathcal{H}_0^1)} < \infty, \quad (2.73)$$

$$\sup_h \|u_h'\|_{L^2(0, T; \mathcal{H}^{-1})} < \infty. \quad (2.74)$$

for the solution of the approximating problem. From the discrete weak formulation (2.58) we get by substituting  $u_h(t)$  for  $v_h$  that

$$\frac{1}{2} \frac{d}{dt} \|u_h(t)\|_{\mathcal{L}^2}^2 + a_h(u_h(t), u_h(t)) = (f(u_h(t)), u_h(t))_{\mathcal{L}^2}. \quad (2.75)$$

By integrating over time and making use of the coerciveness of the bilinear form  $a_h(\cdot, \cdot)$  and the linearity of the reaction function  $f$ , we deduce the estimate

$$\|u_h(t)\|_{\mathcal{L}^2}^2 + 2\alpha \int_0^t \|u_h(s)\|_{\mathcal{H}_0^1}^2 ds \leq \|u_h(0)\|_{\mathcal{L}^2}^2 + C \int_0^t \|u_h(s)\|_{\mathcal{L}^2}^2 ds. \quad (2.76)$$

Since  $u_{h,0}$  converges strongly to  $u_0$  in  $L^2$ , we have

$$\sup_h \|u_h(0)\|_{\mathcal{L}^2} \leq C. \quad (2.77)$$

Consequently, by invoking Gronwall's inequality,

$$\sup_h \max_{0 \leq t \leq T} \|u_h\|_{\mathcal{L}^2} \leq C. \quad (2.78)$$

Thus it follows from (2.76) that

$$\sup_h \|u_h\|_{L^2(0,T;\mathcal{H}_0^1)} \leq C. \quad (2.79)$$

The third a-priori estimate (2.74) follows again from the discrete weak formulation by using the boundedness of the bilinear form  $a_h(\cdot, \cdot)$  and the linearity of the reaction function  $f$ .

2. From the stability of the extension operators  $J_h$  it follows that

$$\sup_h \|J_h u_h\|_{L^2(0,T;X)} < \infty. \quad (2.80)$$

Owing to the latter estimate we can conclude that there exists a subsequence of  $(u_h)$  (still denoted by  $(u_h)$ ) such that  $J_h u_h$  converges weakly to a limit  $\tilde{u}$  in  $X$ . By Lemma 2.2.11 there is a function  $u$  in  $L^2(0,T;H_0^1)$  with  $\tilde{u}(t) = Ju(t)$  for a.e.  $t$ .

3. It remains to show that  $u$  solves the weak PDE problem (2.40) with initial value  $u_0$ . We first show that the equation

$$-(u_0, v)_{L^2} \varphi(0) - \int_0^T (u(t), v)_{L^2} \varphi'(t) dt + \int_0^T b(u(t), v) \varphi(t) dt = 0 \quad (2.81)$$

holds for arbitrary  $v \in H_0^1$  and  $\varphi \in C^1[0,T]$  with  $\varphi(T) = 0$ . To this end, let  $v_h = R_h v$ . From the discrete weak formulation we get by multiplying with  $\varphi(t)$  and integrating over time that

$$-(u_h(0), v_h)_{L^2} \varphi(0) - \int_0^T (u_h(t), v_h)_{L^2} \varphi'(t) dt + \int_0^T b_h(u_h(t), v_h) \varphi(t) dt = 0. \quad (2.82)$$

Because the external approximation of  $H_0^1(G)$  by the discrete Sobolev spaces  $\mathcal{H}_0^1(\mathcal{G}_h)$  is stable and convergent,  $J_h v_h$  converges strongly to  $Jv$  in  $X$ . In particular,  $v_h$  converges strongly to  $v$  in  $L^2$ . Moreover,  $J_h(v_h \varphi')$  converges strongly to  $J(v \varphi')$  in  $L^2(0,T;X)$ . Since  $u_h(0) \rightarrow u_0$  strongly in  $L^2$  by hypothesis, Eq. (2.81) follows from Lemma 2.3.1 by passing to the limit in Eq. (2.82).

Obviously the limit  $u$  satisfies Eq. (2.40a). (Take  $\varphi \in C_0^\infty(0,T)$ .) Furthermore, since the operators  $A : H_0^1 \rightarrow H^{-1}$  and  $F : L^2 \rightarrow L^2$  induced by the bilinear form  $a(\cdot, \cdot)$  and the reaction function  $f$  are linear and bounded, it follows that  $u$  has indeed a generalised time derivative  $u' \in L^2(0,T;H^{-1})$  given by  $u' = Fu - Au$ . It still remains to show that  $u(0) = u_0$ . An integration by parts formula for functions in  $H^1(0,T;H_0^1, L^2)$  (Zeidler (1990b), Proposition 23.23) yields

$$\begin{aligned} & (u(T), v)_{L^2} \varphi(T) - (u(0), v)_{L^2} \varphi(0) \\ &= \int_0^T \langle u'(t), \varphi(t)v \rangle_{H_0^1} dt + \int_0^T \langle (\varphi(t)v)', u(t) \rangle_{H_0^1} dt \\ &= - \int_0^T b(u(t), v) \varphi(t) dt + \int_0^T (u(t), v)_{L^2} \varphi'(t) dt \end{aligned} \quad (2.83)$$

for arbitrary  $v \in H_0^1$  and  $\varphi \in C^1[0, T]$ . We choose a  $\varphi \in C^1[0, T]$  with  $\varphi(0) = 1$  and  $\varphi(T) = 0$  and subtract the above identity from Eq. (2.81) to get

$$(u_0 - u(0), v)_{L^2} = 0 \quad \text{for all } v \in H_0^1. \quad (2.84)$$

Hence  $u_0 = u(0)$  in  $L^2$ , since  $H_0^1$  is dense in  $L^2$ .

4. It follows from the uniqueness of the solution of (2.40) that in fact the whole sequence  $J_h u_h$  converges weakly to  $Ju$  in  $L^2(0, T; X)$ .  $\square$

After having established the weak convergence of the solutions  $u_h$  of the approximating problem (2.54) to the solution  $u$  of the weak PDE problem (2.40), we now turn to the question whether the sequence  $(u_h)$  also converges in a stronger sense. We should emphasize here that for the linear problem under consideration convergence of the semi-discrete approximation can rather easily be proved in  $L^\infty(Q_T)$  (even with explicit convergence rates) if one has sufficient regularity of the solution of (2.40). But, in view of nonlinear problems, especially problems involving a nonlinear diffusion operator, the required high regularity is not always available. The ‘porous medium equation’

$$\partial_t u = \Delta u^\gamma, \quad \gamma > 1,$$

for instance, admits (explicit) weak solutions that are not differentiable in the classical sense (see, e.g., Evans (1998) or Friedman (1982)).

Thus we content ourselves (for the moment) with strong convergence in the space  $L^2(0, T; L^2) \cong L^2(Q_T)$ . A refined result will be presented in the next chapter in Section 3.3. As for fully discrete finite-difference approximations (Raviart, 1967; Temam, 2001), strong convergence will follow from a compactness theorem. Since in our case time is a continuous variable, we can rather easily obtain an appropriate result by adapting the proof from the PDE literature. Theorem 2.3.4 below, which is crucial also for the treatment of nonlinear problems, is sufficient for our purposes. For the next two theorems we suppose that  $G$  is a general bounded domain in  $\mathbb{R}^m$  with Lipschitz boundary, and we adopt the notation of Section 2.2.2. The following result is a discrete analogue of Rellich’s theorem.

**Theorem 2.3.3.** *Let  $(u_h)$ ,  $h \searrow 0$ , be a sequence of functions in  $\mathcal{H}_0^1(\mathcal{G}_h)$  that satisfies  $\sup_h \|u_h\|_{\mathcal{H}_0^1} < \infty$ . Then there is a subsequence  $(u_{h'})$ ,  $h' \searrow 0$ , that converges strongly to a limit in  $L^2(G)$ .*

**Proof.** Cf. Temam (2001), Chapter 2.  $\square$

With the aid of the previous result one can establish the next theorem.

**Theorem 2.3.4.** *Let  $(u_h)$ ,  $h \searrow 0$ , be a sequence of functions in  $L^2(0, T; \mathcal{H}_0^1(\mathcal{G}_h))$  that have time derivatives  $u'_h$  in  $L^2(0, T; \mathcal{H}^{-1}(\mathcal{G}_h))$  and satisfy*

$$\sup_h \int_0^T \left( \|u_h(t)\|_{\mathcal{H}_0^1}^2 + \|u'_h(t)\|_{\mathcal{H}^{-1}}^2 \right) dt < \infty.$$

*Then there is a subsequence  $(u_{h'})$ ,  $h' \searrow 0$ , that converges strongly in  $L^2(0, T; L^2(G))$ .*

**Proof:** The proof can be carried out along the lines of the proof of Theorem 12.1 in Chapter 1 of Lions (1969).  $\square$

The strong convergence of the solutions of the approximating problem for the example model is an immediate consequence of the above theorem. The required estimates have already been derived in the proof of Theorem 2.3.2.

**Corollary 2.3.5.** *Under the same hypotheses as in Theorem 2.3.2 the sequence  $(u_h)$ ,  $h \searrow 0$ , of solutions of the approximating problem (2.54) converges strongly to the solution  $u$  of the weak PDE problem (2.40) in  $L^2(0, T; L^2(G))$ .*

### 2.3.2 Convergence of the particle density

We shall now complete the second (and certainly more interesting) step of the program and estimate the distance between the particle density and the finite-difference approximation in  $L^2(0, T; L^2) \cong L^2(Q_T)$ . Henceforth  $u_l$  denotes the stochastic particle density of the example model introduced in Section 2.1.2. (For additional notation see also Section 2.1.2.)

In Section 2.1.2 we have defined the particle density  $u_l$  as an element of the discrete Lebesgue space  $\mathcal{L}^2(\mathcal{G}_l)$ . As state space  $S_l$  of the particle density process  $(u_l(t))_{t \geq 0}$  we have chosen the discrete subset of functions in  $\mathcal{L}^2(\mathcal{G}_l)$  that take values in  $\frac{1}{n}\mathbb{N}_0$ . Recall that  $\mathcal{G}_l = \{(i-1/2)l : i = 1, \dots, N\}$  is the set of midpoints  $z_i$  of the intervals  $c_l(z_i) = (z_i - l/2, z_i + l/2)$  of length  $l$ . The union of these interval is, up to a finite number of points, the interval  $(0, L)$  which represents the chemical reactor. The particle density process  $(u_l(t))_{t \geq 0}$  has the Markov property with respect to the filtration  $(\mathcal{F}_{l,t})_{t \geq 0}$  induced by itself. It corresponds to a Feller semigroup of bounded linear operators on  $\hat{C}(S_l)$  (the space of continuous real-valued functions vanishing at infinity) that is generated by

$$L_l g(u_l) = \sum_{\tilde{u}_l \neq u_l} q_l(u_l, \tilde{u}_l) (g(\tilde{u}_l) - g(u_l)), \quad g \in \hat{C}(S_l). \quad (2.85)$$

The transition intensities  $q_l(\cdot, \cdot)$  have been specified in Section 2.1.2.

In order to prove a limit theorem, we have to consider a sequence of particle density processes  $(u_{l_i})_{i \in \mathbb{N}_0}$ , where  $l_i \searrow 0$  for  $i \rightarrow \infty$ , but for notational simplicity the index  $i$  is usually omitted (cf. Remark 2.2.8). We may assume that the whole sequence of density processes is defined on the same probability space  $(\Omega, \mathcal{A}, P)$ . Furthermore, let us suppose that  $u_l(0)$  is non-random and recall that the three parameters  $l = L/N$ ,  $n$  and  $d$  are varied (cf. the discussion in Section 2.1.2). We make the following hypotheses for the scale parameters.

$$l \rightarrow 0, \quad n \rightarrow \infty, \quad (S1)$$

$$\frac{1}{2} d l^2 \rightarrow D, \quad (S2)$$

$$\frac{d}{n} \rightarrow 0, \quad (S3)$$

where  $D$  is the macroscopic diffusion coefficient.

**Theorem 2.3.6** (Law of large numbers). *Let  $u$  be the solution of the weak PDE problem (2.40) to the initial value  $u_0$ . Assume that (S1)–(S3) hold and that  $u_l(0)$  converges strongly to  $u_0$  in  $L^2$ . Then*

$$E \left[ \|u_l - u\|_{L^2(0,T;L^2)}^2 \right] = E \left[ \|u_l - u\|_{L^2(Q_T)}^2 \right] \rightarrow 0.$$

In other words, the particle density converges in  $L^2(\Omega; L^2(Q_T))$ .

In the previous paragraph we have already shown the convergence of the solutions of the auxiliary approximating problem (2.54) to the solution of the weak PDE problem (2.40). Therefore the law of large numbers follows immediately from the next result (cf. Remark 2.2.10).

**Theorem 2.3.7.** *Assume (S1)–(S3), and denote by  $v_l$  the solutions of the approximating problem (2.54) with  $D_l = \frac{1}{2}dl^2$  to the initial values  $v_{l,0}$ . Moreover, assume that*

$$\|u_l(0) - v_{l,0}\|_{\mathcal{L}^2} \rightarrow 0.$$

*Then*

$$\sup_{t \leq T} E \left[ \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 \right] \rightarrow 0.$$

The proof of Theorem 2.3.7 is based on the following lemma which identifies a (local) martingale associated to the process  $\|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2$ ,  $t \leq T$ . Let, for  $p \in \mathbb{N}$ ,  $\tau_p$  be the stopping time

$$\tau_p = \inf \left\{ t : \|u_l(t)\|_{\mathcal{L}^2} > p \right\} \wedge T, \quad (2.86)$$

where  $a \wedge b$  means  $\min(a, b)$  for  $a, b \in \mathbb{R}$ . Here, as usual,  $\inf \emptyset = \infty$ . Note that it follows immediately from the fact that the particle density process  $(u_l(t))$  has right-continuous paths with left limits that  $\tau_p \nearrow T$  almost surely for  $p \rightarrow \infty$ .

**Lemma 2.3.8.** *Let  $(M_l(t))_{t \leq T}$  be the process given by*

$$\begin{aligned} M_l(t) &= \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 - \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 \\ &\quad + 2 \int_0^t a_l(u_l(s) - v_l(s), u_l(s) - v_l(s)) ds \\ &\quad - 2 \int_0^t (f(u_l(s)) - f(v_l(s)), u_l(s) - v_l(s))_{\mathcal{L}^2} ds - R_l(t), \end{aligned} \quad (2.87)$$

*where*

$$R_l(t) = \frac{1}{n} \int_0^t \left( 2d(u_l(s), 1)_{\mathcal{L}^2} + (k_1 u_l(s), 1)_{\mathcal{L}^2} + (k_2 u_l(s), 1)_{\mathcal{L}^2} \right) ds. \quad (2.88)$$

*Then the stopped process  $(M_l(t \wedge \tau_p))_{t \leq T}$  is a martingale for each  $p \in \mathbb{N}$ .*

**Proof.** The martingale relationship follows, roughly speaking, from an application of Dynkin's formula (2.24) for the process  $(u_l(t), t)$  with the function  $(u_l, t) \mapsto \|u_l - v_l(t)\|_{\mathcal{L}^2}^2$  and explicit calculations. We now give the details of the proof.

1. Let, for  $i = 2, \dots, N-1$ ,  $\chi_i$  be the lattice function that is equal to one on the interval  $c_l(z_i)$  and zero elsewhere, and let  $\chi_1 = \chi_N \equiv 0$  (cf. Section 2.1.2). For any function  $g \in \hat{C}(S_l)$  we have

$$\begin{aligned} L_l g(u_l) &= \sum_{\tilde{u}_l \neq u_l} q_l(u_l, \tilde{u}_l) (g(\tilde{u}_l) - g(u_l)) \\ &= \sum_{i=2}^{N-1} n \frac{d}{2} u_l(z_i) \left( g(u_l - \frac{1}{n} \chi_i + \frac{1}{n} \chi_{i-1}) - 2g(u_l) + g(u_l - \frac{1}{n} \chi_i + \frac{1}{n} \chi_{i+1}) \right) \\ &\quad + \sum_{i=2}^{N-1} n k_1 u_l(z_i) \left( g(u_l + \frac{1}{n} \chi_i) - g(u_l) \right) \\ &\quad + \sum_{i=2}^{N-1} n k_2 u_l(z_i) \left( g(u_l - \frac{1}{n} \chi_i) - g(u_l) \right). \end{aligned} \tag{2.89}$$

Here the first sum consists of the terms that stem from 'diffusive' jumps and the second and third sum are the contributions coming from the chemical reactions. Thus the generator can be written as a sum  $L_l = L_{d,l} + L_{r,l}$ , where

$$L_{d,l} g(u_l) = \sum_{i=2}^{N-1} n \frac{d}{2} u_l(z_i) \left( g(u_l - \frac{1}{n} \chi_i + \frac{1}{n} \chi_{i-1}) - 2g(u_l) + g(u_l - \frac{1}{n} \chi_i + \frac{1}{n} \chi_{i+1}) \right), \tag{2.90}$$

and

$$\begin{aligned} L_{r,l} g(u_l) &= \sum_{i=2}^{N-1} n k_1 u_l(z_i) \left( g(u_l + \frac{1}{n} \chi_i) - g(u_l) \right) \\ &\quad + \sum_{i=2}^{N-1} n k_2 u_l(z_i) \left( g(u_l - \frac{1}{n} \chi_i) - g(u_l) \right). \end{aligned} \tag{2.91}$$

2. Next, we consider for arbitrary but fixed  $w_l \in \mathcal{H}_0^1$  the function  $g(\cdot, w_l) : S_l \rightarrow \mathbb{R}$  given by  $u_l \mapsto \|u_l - w_l\|_{\mathcal{L}^2}^2$ . The function  $g(\cdot, w_l)$  is unbounded and thus not in  $\hat{C}(S_l)$ . However, it makes sense to compute  $L_l g$  for any real-valued function  $g$  on  $S_l$ . We shall later handle this difficulty by a truncation argument. The diffusion part yields

$$\begin{aligned} L_{d,l} g(u_l, w_l) &= \sum_{i=2}^{N-1} n \frac{d}{2} u_l(z_i) \left( g(u_l - \frac{1}{n} \chi_i + \frac{1}{n} \chi_{i-1}, w_l) - 2g(u_l, w_l) \right. \\ &\quad \left. + g(u_l - \frac{1}{n} \chi_i + \frac{1}{n} \chi_{i+1}, w_l) \right) \\ &= \sum_{i=2}^{N-1} n \frac{d}{2} u_l(z_i) \left( \|u_l - \frac{1}{n} \chi_i + \frac{1}{n} \chi_{i-1} - w_l\|_{\mathcal{L}^2}^2 - 2\|u_l - w_l\|_{\mathcal{L}^2}^2 \right. \\ &\quad \left. + \|u_l - \frac{1}{n} \chi_i + \frac{1}{n} \chi_{i+1} - w_l\|_{\mathcal{L}^2}^2 \right). \end{aligned} \tag{2.92}$$

Hence, by explicit calculation, it follows that

$$\begin{aligned}
L_{d,l}g(u_l, w_l) &= \sum_{i=2}^{N-1} n \frac{d}{2} u_l(z_i) \left( \frac{2}{n} (u_l - w_l, \chi_{i-1} - 2\chi_i + \chi_{i+1})_{\mathcal{L}^2} \right. \\
&\quad \left. + \frac{1}{n^2} \|\chi_{i-1} - \chi_i\|_{\mathcal{L}^2}^2 + \frac{1}{n^2} \|\chi_{i+1} - \chi_i\|_{\mathcal{L}^2}^2 \right) \\
&= \sum_{i=2}^{N-1} n \frac{d}{2} u_l(z_i) \left( 2 \frac{l}{n} (u_l(z_{i-1}) - w_l(z_{i-1}) - 2(u_l(z_i) - w_l(z_i)) \right. \\
&\quad \left. + u_l(z_{i+1}) - w_l(z_{i+1})) + 4 \frac{l}{n^2} \right) \\
&= 2 \frac{1}{2} dl^2 (u_l, \partial^- \partial^+ (u_l - w_l))_{\mathcal{L}^2} + 2 \frac{d}{n} (u_l, 1)_{\mathcal{L}^2} \\
&= -2 a_l(u_l, u_l - w_l) + 2 \frac{d}{n} (u_l, 1)_{\mathcal{L}^2}.
\end{aligned} \tag{2.93}$$

Here the bilinear form  $a_l(\cdot, \cdot)$  is given by formula (2.55) with  $D_l = \frac{1}{2} dl^2$ . Computing the reaction terms yields

$$\begin{aligned}
L_{r,l}g(u_l, w_l) &= \sum_{i=2}^{N-1} n k_1 u_l(z_i) (g(u_l + \frac{1}{n} \chi_i, w_l) - g(u_l, w_l)) \\
&\quad + \sum_{i=2}^{N-1} n k_2 u_l(z_i) (g(u_l - \frac{1}{n} \chi_i, w_l) - g(u_l, w_l)) \\
&= \sum_{i=2}^{N-1} n k_1 u_l(z_i) \left( \|u_l + \frac{1}{n} \chi_i - w_l\|_{\mathcal{L}^2}^2 - \|u_l - w_l\|_{\mathcal{L}^2}^2 \right) \\
&\quad + \sum_{i=2}^{N-1} n k_2 u_l(z_i) \left( \|u_l - \frac{1}{n} \chi_i - w_l\|_{\mathcal{L}^2}^2 - \|u_l - w_l\|_{\mathcal{L}^2}^2 \right) \\
&= \sum_{i=2}^{N-1} n k_1 u_l(z_i) \left( \frac{2}{n} (u_l - w_l, \chi_i)_{\mathcal{L}^2} + \frac{l}{n^2} \right) \\
&\quad + \sum_{i=2}^{N-1} n k_2 u_l(z_i) \left( -\frac{2}{n} (u_l - w_l, \chi_i)_{\mathcal{L}^2} + \frac{l}{n^2} \right) \\
&= \sum_{i=2}^{N-1} n k_1 u_l(z_i) \left( 2 \frac{l}{n} (u_l(z_i) - w_l(z_i)) + \frac{l}{n^2} \right) \\
&\quad + \sum_{i=2}^{N-1} n k_2 u_l(z_i) \left( -2 \frac{l}{n} (u_l(z_i) - w_l(z_i)) + \frac{l}{n^2} \right) \\
&= 2 (k_1 u_l - k_2 u_l, u_l - w_l)_{\mathcal{L}^2} + \frac{1}{n} (k_1 u_l, 1)_{\mathcal{L}^2} + \frac{1}{n} (k_2 u_l, 1)_{\mathcal{L}^2} \\
&= 2 (f(u_l), u_l - w_l)_{\mathcal{L}^2} + \frac{1}{n} (k_1 u_l, 1)_{\mathcal{L}^2} + \frac{1}{n} (k_2 u_l, 1)_{\mathcal{L}^2}.
\end{aligned} \tag{2.94}$$

By gathering together the different contributions we finally find that

$$\begin{aligned} L_l g(u_l, w_l) &= -2 a_l(u_l, u_l - w_l) + 2 (f(u_l), u_l - w_l)_{\mathcal{L}^2} \\ &\quad + 2 \frac{d}{n} (u_l, 1)_{\mathcal{L}^2} + \frac{1}{n} (k_1 u_l, 1)_{\mathcal{L}^2} + \frac{1}{n} (k_2 u_l, 1)_{\mathcal{L}^2}. \end{aligned} \quad (2.95)$$

3. Consider now the process  $(u_l(t), t)$ ,  $t \leq T$ , with state space  $S_l \times [0, T]$ . It follows from Dynkin's formula (2.24) and a well-known argument (see, e.g., the proof of Theorem 4.2.1 in Stroock & Varadhan (1979)) that, for functions  $\Phi : S_l \times [0, T] \rightarrow \mathbb{R}$ ,  $(u_l, t) \mapsto \Phi(u_l, t)$ , which satisfy  $\Phi(\cdot, t) \in \hat{C}(S_l)$  and  $\Phi(u_l, \cdot) \in C^1[0, T]$ , the process

$$\Phi(u_l(t), t) - \Phi(u_l(0), 0) - \int_0^t (L_l \Phi(u_l(s), s) + \partial_s \Phi(u_l(s), s)) ds, \quad t \leq T, \quad (2.96)$$

is a martingale (with respect to  $(\mathcal{F}_{l,t})$ ). Recall that the solutions of the approximating problem (2.33) are denoted by  $v_l$  and consider for arbitrary but fixed  $w_l \in \mathcal{H}_0^1$  the function  $h(\cdot, w_l) : [0, T] \rightarrow \mathbb{R}$ ,  $t \mapsto \|w_l - v_l(t)\|_{\mathcal{L}^2}^2$ . Observe that

$$\begin{aligned} h'(t, w_l) &= -2 (v_l'(t), w_l - v_l(t))_{\mathcal{L}^2} \\ &= 2 a_l(v_l(t), w_l - v_l(t)) - 2 (f(v_l), w_l - v_l(t))_{\mathcal{L}^2}. \end{aligned}$$

Now let  $\Phi(u_l, t) = \|u_l - v_l(t)\|_{\mathcal{L}^2}^2$ , and

$$\tilde{\Phi}(u_l, t) = \begin{cases} \|u_l - v_l(t)\|_{\mathcal{L}^2}^2 & \text{if } \|u_l\|_{\mathcal{L}^2} \leq \tilde{p} \\ 0 & \text{if } \|u_l\|_{\mathcal{L}^2} > \tilde{p} \end{cases} \quad (2.97)$$

for  $\tilde{p} \in \mathbb{N}$ . The martingale formula (2.96) holds for  $\tilde{\Phi}$ . If we denote the resulting martingale by  $\tilde{M}_l$ , then the stopped process  $\tilde{M}_l(\cdot \wedge \tau_p)$  is a martingale as well. That is,

$$\tilde{\Phi}(u_l(t \wedge \tau_p), t \wedge \tau_p) - \tilde{\Phi}(u_l(0), 0) - \int_0^{t \wedge \tau_p} (L_l \tilde{\Phi}(u_l(s), s) + \partial_s \tilde{\Phi}(u_l(s), s)) ds, \quad (2.98)$$

$t \leq T$ , is a martingale for each  $p \in \mathbb{N}$ . Recall that  $q_l(u_l, \tilde{u}_l)$  is nonzero only for transitions of the form  $u_l \rightarrow \tilde{u}_l = u_l + \delta$  for  $\delta$  from a finite set  $\mathcal{T} \subset \mathcal{L}^2$ . Hence, if we choose  $\tilde{p} > p$  sufficiently large, we are allowed to replace  $\tilde{\Phi}$  by  $\Phi$ ,  $L_l \tilde{\Phi}$  by  $L_l \Phi$  and  $\partial_s \tilde{\Phi}$  by  $\partial_s \Phi$  in the formula above. Note that  $L_l \Phi(u_l(s), s) = L_l g(u_l(s), v_l(s))$ , and  $\partial_s \Phi(u_l(s), s) = h'(s, u_l(s))$ . Consequently, by substituting the computations above, it follows that

$$\begin{aligned} \|u_l(t \wedge \tau_p) - v_l(t \wedge \tau_p)\|_{\mathcal{L}^2}^2 &= \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 \\ &\quad - 2 \int_0^{t \wedge \tau_p} a_l(u_l(s) - v_l(s), u_l(s) - v_l(s)) ds \\ &\quad + 2 \int_0^{t \wedge \tau_p} (f(u_l(s)) - f(v_l(s)), u_l(s) - v_l(s))_{\mathcal{L}^2} ds \\ &\quad + \frac{1}{n} \int_0^{t \wedge \tau_p} \left( 2d(u_l(s), 1)_{\mathcal{L}^2} + (k_1 u_l(s), 1)_{\mathcal{L}^2} + (k_2 u_l(s), 1)_{\mathcal{L}^2} \right) ds, \end{aligned} \quad (2.99)$$

$t \leq T$ , is a martingale for each  $p \in \mathbb{N}$ , and the claim is proved.  $\square$

**Proof of Theorem 2.3.7.** The idea is to use the previous lemma to obtain an estimate for  $\sup_{t \leq T} E \left[ \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 \right]$ . To this end, we stop the local martingale (2.87) at  $\tau_p$  and take expectations. Owing to the monotonicity of the operator  $A_l$  associated to the bilinear form  $a_l(\cdot, \cdot)$  and the linearity of the reaction function  $f$ , we can deduce the estimate ( $\hat{k}$  denoting  $\max_{i=1,2} k_i$ )

$$\begin{aligned}
& E \left[ \|u_l(t \wedge \tau_p) - v_l(t \wedge \tau_p)\|_{\mathcal{L}^2}^2 \right] \\
& \leq \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 + 4\hat{k} E \int_0^{t \wedge \tau_p} \|u_l(s) - v_l(s)\|_{\mathcal{L}^2}^2 ds \\
& \quad + E \int_0^{t \wedge \tau_p} 2\sqrt{L}(d/n + \hat{k}/n) \|u_l(s)\|_{\mathcal{L}^2} ds \\
& \leq \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 + 4\hat{k} E \int_0^{t \wedge \tau_p} \|u_l(s) - v_l(s)\|_{\mathcal{L}^2}^2 ds \\
& \quad + 2\sqrt{L}((d + \hat{k})/n) E \int_0^{t \wedge \tau_p} \left( \frac{1}{2} \|v_l(s)\|_{\mathcal{L}^2}^2 + \frac{1}{2} \right) ds \\
& \quad + 2\sqrt{L}((d + \hat{k})/n) E \int_0^{t \wedge \tau_p} \left( \frac{1}{2} \|u_l(s) - v_l(s)\|_{\mathcal{L}^2}^2 + \frac{1}{2} \right) ds.
\end{aligned} \tag{2.100}$$

Consequently,

$$\begin{aligned}
& E \left[ \|u_l(t \wedge \tau_p) - v_l(t \wedge \tau_p)\|_{\mathcal{L}^2}^2 \right] \leq \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 \\
& \quad + \sqrt{L}((d + \hat{k})/n) E \int_0^{t \wedge \tau_p} (\|v_l(s)\|_{\mathcal{L}^2}^2 + 2) ds \\
& \quad + \left( \sqrt{L}((d + \hat{k})/n) + 4\hat{k} \right) E \int_0^{t \wedge \tau_p} \|u_l(s) - v_l(s)\|_{\mathcal{L}^2}^2 ds \\
& \leq \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 + \sqrt{L}((d + \hat{k})/n) (\|v_l\|_{\mathcal{L}^2(0,T;\mathcal{L}^2)}^2 + 2T) \\
& \quad + \left( \sqrt{L}((d + \hat{k})/n) + 4\hat{k} \right) E \int_0^{t \wedge \tau_p} \|u_l(s) - v_l(s)\|_{\mathcal{L}^2}^2 ds.
\end{aligned} \tag{2.101}$$

The above inequality is valid even with  $t \wedge \tau_p$  replaced by  $t$ . To see this, note first that

$$\int_0^t E \left[ \|u_l(s)\|_{\mathcal{L}^2}^2 \right] ds < \infty \quad \text{for } t \leq T. \tag{2.102}$$

This follows by observing that the growth of the process  $\|u_l(t)\|_{\mathcal{L}^2}$ ,  $t \geq 0$ , can be estimated by an appropriate pure-birth (Yule) process for which Kolmogorov's forward system can be solved explicitly (see, e.g., Karlin & Taylor (1975), p. 122). Hence,

$$\begin{aligned}
E \int_0^{t \wedge \tau_p} \|u_l(s) - v_l(s)\|_{\mathcal{L}^2}^2 ds & \rightarrow E \int_0^t \|u_l(s) - v_l(s)\|_{\mathcal{L}^2}^2 ds \\
& = \int_0^t E \left[ \|u_l(s) - v_l(s)\|_{\mathcal{L}^2}^2 \right] ds
\end{aligned} \tag{2.103}$$

for  $p \rightarrow \infty$  by monotone convergence and Fubini's theorem. The inequality follows from Fatou's lemma. Finally, by invoking Gronwall's inequality (see, e.g., Theorem 5.1 in the

appendix of Ethier & Kurtz (1986)),

$$\begin{aligned} \sup_{t \leq T} E \left[ \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 \right] &\leq \left( \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 \right. \\ &\quad \left. + \sqrt{L}((d + \hat{k})/n) (\|v_l\|_{L^2(0,T;\mathcal{L}^2)}^2 + 2T) \right) \times \\ &\quad \times \exp \left( (\sqrt{L}((d + \hat{k})/n) + 4\hat{k}) T \right). \end{aligned} \quad (2.104)$$

Because  $\|v_l(0) - u_l(0)\|_{\mathcal{L}^2} \rightarrow 0$  for  $l \rightarrow 0$  by hypothesis and  $\sup_l \|v_l\|_{L^2(0,T;\mathcal{L}^2)} < \infty$  due to the a-priori estimate (2.72), it follows from the scaling hypotheses (S1)–(S3) that

$$\sup_{t \leq T} E \left[ \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 \right] \rightarrow 0. \quad \square$$

## 2.4 Law of large numbers for the general linear model

For the derivation of the law of large numbers for the general linear model we shall follow exactly the same strategy as for the example model.

### 2.4.1 Convergence of the approximation

We use the definitions and notation of Sections 2.1.2 and 2.2.2. The space  $X$  that is introduced to compare solution and approximation is defined as  $X = \prod_{j=1}^{n_s} \prod_{k=1}^{m+1} L^2(G)$ . The restriction operators  $Q_h : L^2(G) \rightarrow \mathcal{L}^2(\mathcal{G}_h)$  and  $R_h : \mathbf{H}_0^1(G) \rightarrow \mathcal{H}_0^1(\mathcal{G}_h)$ , the extension operators  $P_h : \mathcal{L}^2(\mathcal{G}_h) \rightarrow L^2(G)$  and  $J_h : \mathcal{H}_0^1(\mathcal{G}_h) \rightarrow X$ , and the embedding  $J : \mathbf{H}_0^1(G) \rightarrow X$  have to be redefined in the obvious way. Clearly,  $(\mathcal{L}^2(\mathcal{G}_h), Q_h, P_h)_{h \in I}$ , where  $I = (0, h_0] \subset \mathbb{R}^+$ , is a stable and convergent internal approximation of  $L^2(G)$ , and  $\{X, J, (\mathcal{H}_0^1(\mathcal{G}_h), R_h, J_h)_{h \in I}\}$  is a stable and convergent external approximation of  $\mathbf{H}_0^1(G)$ . The following lemma can be proved in the same way as Lemma 2.3.1.

**Lemma 2.4.1.** *If  $(\mathbf{u}_h)$ ,  $(\mathbf{v}_h)$ ,  $h \searrow 0$ , are two sequences in  $L^2(0, T; \mathcal{H}_0^1)$  such that  $J_h \mathbf{u}_h$  converges weakly to  $J\mathbf{u}$  and  $J_h \mathbf{v}_h$  converges strongly to  $J\mathbf{v}$  in  $L^2(0, T; X)$  for functions  $u, v \in L^2(0, T; \mathbf{H}_0^1)$ , then*

$$\int_0^T (\mathbf{u}_h(t), \mathbf{v}_h(t))_{\mathcal{L}^2} dt \rightarrow \int_0^T (\mathbf{u}(t), \mathbf{v}(t))_{L^2} dt,$$

and

$$\int_0^T b_h(\mathbf{u}_h(t), \mathbf{v}_h(t)) dt \rightarrow \int_0^T b(\mathbf{u}(t), \mathbf{v}(t)) dt.$$

In the next theorem we show convergence of the solutions of the approximating problem (2.60) for the general linear model. As compared to the example model, there are only a few minor changes in the proof and therefore we shall only sketch it.

**Theorem 2.4.2.** *Let  $\mathbf{u}$  be the solution of the weak PDE problem (2.51) for the general linear model to the initial value  $\mathbf{u}_0$ , and let  $(\mathbf{u}_h)$ ,  $h \searrow 0$ , be a sequence of solutions of the approximating problem (2.60) to the initial values  $\mathbf{u}_{h,0}$ . Assume  $\mathbf{u}_{h,0} \rightarrow \mathbf{u}_0$  strongly in  $\mathbf{L}^2$ , and  $D_{h,j} \rightarrow D_j$ ,  $j = 1, \dots, n_s$ . Then  $J_h \mathbf{u}_h$  converges weakly to  $J\mathbf{u}$  in  $L^2(0, T; X)$ , and  $\mathbf{u}_h$  converges strongly to  $\mathbf{u}$  in  $L^2(0, T; \mathbf{L}^2)$ .*

**Proof:** 1. We first derive the a-priori estimates

$$\sup_h \max_{0 \leq t \leq T} \|\mathbf{u}_h\|_{\mathcal{L}^2} < \infty, \quad (2.105)$$

$$\sup_h \|\mathbf{u}_h\|_{L^2(0, T; \mathcal{H}_0^1)} < \infty, \quad (2.106)$$

$$\sup_h \|\mathbf{u}'_h\|_{L^2(0, T; \mathcal{H}^{-1})} < \infty. \quad (2.107)$$

By inserting  $u_h(t)$  for  $v_h$  in the discrete weak formulation and making use of the coerciveness of the bilinear form  $a_h(\cdot, \cdot)$  and the linearity of the reaction functions we get the estimate

$$\begin{aligned} \|\mathbf{u}_h(t)\|_{\mathcal{L}^2}^2 + 2\alpha \int_0^t \|\mathbf{u}_h(s)\|_{\mathcal{H}_0^1}^2 ds &\leq \|\mathbf{u}_h(0)\|_{\mathcal{L}^2}^2 + 2 \int_0^t (\mathbf{f}(\mathbf{u}_h), \mathbf{u}_h)_{\mathcal{L}^2} \\ &\leq \|\mathbf{u}_h(0)\|_{\mathcal{L}^2}^2 + C \int_0^t \|\mathbf{u}_h(s)\|_{\mathcal{L}^2}^2 ds. \end{aligned}$$

Here the constant  $C$  is independent of  $h$ , and thus Gronwall's inequality yields

$$\sup_h \max_{0 \leq t \leq T} \|\mathbf{u}_h\|_{\mathcal{L}^2} < \infty. \quad (2.108)$$

Consequently, we also have

$$\sup_h \|\mathbf{u}_h\|_{L^2(0, T; \mathcal{H}_0^1)} < \infty. \quad (2.109)$$

The third a-priori estimate follows again from the boundedness of the bilinear form  $a_h(\cdot, \cdot)$  and the linearity of  $\mathbf{f}$ .

2. Because the extension operators  $J_h : \mathcal{H}_0^1 \rightarrow X$  are stable, we can choose a subsequence of  $(\mathbf{u}_h)$  (still denoted by  $(\mathbf{u}_h)$ ) such that  $J_h \mathbf{u}_h$  converges weakly to  $J\mathbf{u}$  in  $L^2(0, T; X)$  for a  $\mathbf{u} \in L^2(0, T; \mathbf{H}_0^1)$ . Furthermore, we may assume that  $\mathbf{u}_h \rightarrow \mathbf{u}$  strongly in  $L^2(0, T; \mathbf{L}^2)$  due to the compactness theorem 2.3.4.

3. From the discrete weak formulation and the previous lemma we deduce that

$$-(\mathbf{u}_0, \mathbf{v})_{\mathbf{L}^2} \varphi(0) - \int_0^T (\mathbf{u}(t), \mathbf{v})_{\mathbf{L}^2} \varphi'(t) dt + \int_0^T b(\mathbf{u}(t), \mathbf{v}) \varphi(t) dt = 0 \quad (2.110)$$

for arbitrary  $\mathbf{v} \in \mathbf{H}_0^1$  and  $\varphi \in C^1[0, T]$  with  $\varphi(T) = 0$ . It follows in the same way as in the proof for the example model that  $\mathbf{u}$  is indeed the unique solution of the weak PDE problem (2.51) with initial value  $\mathbf{u}(0) = \mathbf{u}_0$ .

4. Since the limit is unique, the entire sequence converges strongly to  $\mathbf{u}$  in  $L^2(0, T; \mathbf{L}^2)$ .  $\square$

### 2.4.2 Convergence of the particle density

Henceforth we denote the stochastic particle density by  $\mathbf{u}_l$  and the solutions of the approximating problem (2.60) by  $\mathbf{v}_l$ , where  $l$  is the edge length of a cell. We assume that  $\mathbf{u}_l(0)$  is non-random, and the hypotheses for the scale parameters are

$$l \rightarrow 0, \quad n \rightarrow \infty, \quad (\text{S1})$$

$$\frac{d_j}{2m} l^2 \rightarrow D_j, \quad (\text{S2})$$

$$\frac{d_j}{n} \rightarrow 0, \quad (\text{S3})$$

$j = 1, \dots, n_s$ , where the  $D_j$  are the macroscopic diffusion coefficients.

**Theorem 2.4.3** (Law of large numbers). *Let  $\mathbf{u}$  be the solution of the weak PDE problem (2.51) to the initial value  $\mathbf{u}_0$ . Assume that (S1)–(S3) hold and that  $\mathbf{u}_l(0)$  converges strongly to  $\mathbf{u}_0$  in  $\mathbf{L}^2$ . Then*

$$E \left[ \|\mathbf{u}_l - \mathbf{u}\|_{L^2(0,T;\mathbf{L}^2)}^2 \right] = E \left[ \|\mathbf{u}_l - \mathbf{u}\|_{(L^2(Q_T))^{n_s}}^2 \right] \rightarrow 0.$$

In view of Theorem 2.4.2, the law of large numbers follows directly from the next result.

**Theorem 2.4.4.** *Assume (S1)–(S3) and denote by  $\mathbf{v}_l$  the solutions of the approximating problem (2.60) with  $D_{l,j} = \frac{d_j}{2m} l^2$ ,  $j = 1, \dots, n_s$ , to the initial values  $\mathbf{v}_{l,0}$ . Moreover, assume that*

$$\|\mathbf{u}_l(0) - \mathbf{v}_{l,0}\|_{\mathcal{L}^2} \rightarrow 0.$$

Then

$$\sup_{t \leq T} E \left[ \|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2 \right] \rightarrow 0.$$

The proof of Theorem 2.4.4 is again based on a lemma that identifies a martingale associated to the process  $\|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2$ ,  $t \leq T$ . We define for  $p \in \mathbb{N}$  the stopping time  $\tau_p$  by

$$\tau_p = \inf \left\{ t : \|\mathbf{u}_l(t)\|_{\mathcal{L}^2} > p \right\} \wedge T. \quad (2.111)$$

As compared to the example model, we have to deal in the general case with the slight technical difficulty that reactive jumps that would lead out of the state space, i.e., to negative concentrations, are not ‘automatically’ excluded. It might happen that a  $K_i(\mathbf{w})$  is positive for a certain concentration vector  $\mathbf{w} \in \frac{1}{n} \mathbb{N}_0^{n_s}$  although the transition from a state  $\mathbf{u}_l \in S_l$  with  $\mathbf{u}_l(z) = \mathbf{w}$  for a  $z \in \mathcal{G}_l$  to  $\tilde{\mathbf{u}}_l = \mathbf{u}_l + \frac{1}{n} \sum_{j=1}^{n_s} \nu_{ij} \chi_{j,z}$  (i.e.,  $\tilde{u}_{l,j}(z) = w_j + \frac{1}{n} \nu_{ij}$  for  $j = 1, \dots, n_s$ ) is not allowed because it would lead to negative concentrations. This might be the case if  $w_j$  is close to zero for a certain  $j$  and  $\nu_{ij}$  is negative, say,  $w_j = 1/n$  and  $\nu_{ij} = -2$ . (For the example model this situation could not arise because the particle density could only increase

or decrease by  $1/n$ .) However, we may always assume (by possibly modifying the original  $K_i$ ) that there are measurable functions  $K_{l,i} : \mathbb{R}^{n_s} \rightarrow \mathbb{R}$ ,  $i = 1, \dots, n_r$ , that converge uniformly to  $K_i$  for  $l \rightarrow 0$  such that the transition intensities are left unchanged and intensity zero is automatically assigned to jumps that would leave the state space. That is,

$$K_{l,i}(\mathbf{w}) = \begin{cases} K_i(\mathbf{w}) & \text{if } w_j + \frac{1}{n}\nu_{ij} \geq 0 \text{ for all } j = 1, \dots, n_s \\ 0 & \text{otherwise} \end{cases} \quad (2.112)$$

for all  $\mathbf{w} \in \frac{1}{n}\mathbb{N}_0^{n_s}$ , and

$$\sup_{R^{n_s}} |K_i - K_{l,i}| \rightarrow 0 \quad (l \rightarrow 0). \quad (2.113)$$

The vector of reaction functions corresponding to the  $K_{l,i}$ , which is defined as in Eq. (2.6), is denoted by  $\mathbf{f}_l$ .

**Lemma 2.4.5.** *Let  $(M_l(t))_{t \leq T}$  be the process given by*

$$\begin{aligned} M_l(t) &= \|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2 - \|\mathbf{u}_l(0) - \mathbf{v}_l(0)\|_{\mathcal{L}^2}^2 \\ &\quad + 2 \int_0^t a_l(\mathbf{u}_l(s) - \mathbf{v}_l(s), \mathbf{u}_l(s) - \mathbf{v}_l(s)) ds \\ &\quad - 2 \int_0^t \left( \mathbf{f}_l(\mathbf{u}_l(s)) - \mathbf{f}_l(\mathbf{v}_l(s)), \mathbf{u}_l(s) - \mathbf{v}_l(s) \right)_{\mathcal{L}^2} ds - R_l(t), \end{aligned} \quad (2.114)$$

where

$$R_l(t) = \frac{1}{n} \sum_{j=1}^{n_s} 2d_j \int_0^t (u_{l,j}(s), 1)_{\mathcal{L}^2} ds + \frac{1}{n} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} \nu_{ij}^2 \int_0^t (K_{l,i}(\mathbf{u}_l(s)), 1)_{\mathcal{L}^2} ds. \quad (2.115)$$

Then the stopped process  $M_l(t \wedge \tau_p)_{t \leq T}$  is a martingale for each  $p \in \mathbb{N}$ .

**Proof.** We follow the same strategy as in the proof for the example model.

1. For arbitrary but fixed  $\mathbf{w}_l \in \mathcal{H}_0^1$  we define the function  $g(\cdot, \mathbf{w}_l) : S_l \rightarrow \mathbb{R}$ ,  $\mathbf{u}_l \mapsto g(\mathbf{u}_l, \mathbf{w}_l) = \|\mathbf{u}_l - \mathbf{w}_l\|_{\mathcal{L}^2}^2$ , and we are going to compute  $L_l g(\mathbf{u}_l, \mathbf{w}_l)$ . The generator  $L_l$  can again be written as  $L_l = L_{d,l} + L_{r,l}$  if we separate reactive and diffusive jumps. We start with the computation

of  $L_{d,l}g(\mathbf{u}_l, \mathbf{w}_l)$ . (Cf. the corresponding calculation (2.93) for the example model.)

$$\begin{aligned}
L_{d,l}g(\mathbf{u}_l, \mathbf{w}_l) &= \sum_{j=1}^{n_s} \sum_{z \in \mathcal{G}_l^1} \sum_{k=1}^m n \frac{d_j}{2m} u_{l,j}(z) \left( \|\mathbf{u}_l - \frac{1}{n} \boldsymbol{\chi}_{j,z} + \frac{1}{n} \boldsymbol{\chi}_{j,(z-le_k)} - \mathbf{w}_l\|_{\mathcal{L}^2}^2 \right. \\
&\quad \left. - 2 \|\mathbf{u}_l - \mathbf{w}_l\|_{\mathcal{L}^2}^2 + \|\mathbf{u}_l - \frac{1}{n} \boldsymbol{\chi}_{j,z} + \frac{1}{n} \boldsymbol{\chi}_{j,(z+le_k)} - \mathbf{w}_l\|_{\mathcal{L}^2}^2 \right) \\
&= \sum_{j=1}^{n_s} \sum_{z \in \mathcal{G}_l^1} \sum_{k=1}^m n \frac{d_j}{2m} u_{l,j}(z) \left( 2 \frac{l^{m+2}}{n} \left( \partial_k^- \partial_k^+ u_{l,j}(z) - \partial_k^- \partial_k^+ w_{l,j}(z) \right) + 4 \frac{l^m}{n^2} \right) \\
&= \sum_{j=1}^{n_s} \sum_{k=1}^m \left( 2 \frac{d_j}{2m} l^2 (u_{l,j}, \partial_k^- \partial_k^+ (u_{l,j} - w_{l,j}))_{\mathcal{L}^2} + \frac{2}{m} \frac{d_j}{n} (u_{l,j}, 1)_{\mathcal{L}^2} \right) \\
&= -2 \sum_{j=1}^{n_s} \sum_{k=1}^m \frac{d_j}{2m} l^2 (\partial_k^+ u_{l,j}, \partial_k^+ (u_{l,j} - w_{l,j}))_{\mathcal{L}^2} + 2 \sum_{j=1}^{n_s} \frac{d_j}{n} (u_{l,j}, 1)_{\mathcal{L}^2} \\
&= -2 a_l(\mathbf{u}_l, \mathbf{u}_l - \mathbf{w}_l) + \frac{1}{n} \sum_{j=1}^{n_s} 2d_j (u_{l,j}, 1)_{\mathcal{L}^2}.
\end{aligned} \tag{2.116}$$

Here  $a_l(\cdot, \cdot)$  is given by (2.61) with  $D_{l,j} = \frac{d_j}{2m} l^2$ . Computing the reaction part yields

$$\begin{aligned}
L_{r,l}g(\mathbf{u}_l, \mathbf{w}_l) &= \sum_{z \in \mathcal{G}_l^1} \sum_{i=1}^{n_r} n K_{l,i}(\mathbf{u}_l(z)) \left( \|\mathbf{u}_l + \frac{1}{n} \sum_{j=1}^{n_s} \nu_{ij} \boldsymbol{\chi}_{j,z} - \mathbf{w}_l\|_{\mathcal{L}^2}^2 - \|\mathbf{u}_l - \mathbf{w}_l\|_{\mathcal{L}^2}^2 \right) \\
&= \sum_{z \in \mathcal{G}_l^1} \sum_{i=1}^{n_r} n K_{l,i}(\mathbf{u}_l(z)) \left( 2 \frac{l^m}{n} \sum_{j=1}^{n_s} \nu_{ij} (u_{l,j}(z) - w_{l,j}(z)) + \frac{l^m}{n^2} \sum_{j=1}^{n_s} \nu_{ij}^2 \right).
\end{aligned} \tag{2.117}$$

Hence,

$$\begin{aligned}
L_{r,l}g(\mathbf{u}_l, \mathbf{w}_l) &= \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} \sum_{z \in \mathcal{G}_l^1} \left( 2 n \frac{l^m}{n} \nu_{ij} K_{l,i}(\mathbf{u}_l(z)) (u_{l,j}(z) - w_{l,j}(z)) + n \frac{l^m}{n^2} \nu_{ij}^2 K_{l,i}(\mathbf{u}_l(z)) \right) \\
&= 2 (\mathbf{f}_l(\mathbf{u}_l), \mathbf{u}_l - \mathbf{w}_l)_{\mathcal{L}^2} + \frac{1}{n} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} \nu_{ij}^2 (K_{l,i}(\mathbf{u}_l), 1)_{\mathcal{L}^2}.
\end{aligned} \tag{2.118}$$

By gathering together the different contributions we finally get

$$\begin{aligned}
L_l g(\mathbf{u}_l, \mathbf{w}_l) &= -2 a_l(\mathbf{u}_l, \mathbf{u}_l - \mathbf{w}_l) + 2 (\mathbf{f}_l(\mathbf{u}_l), \mathbf{u}_l - \mathbf{w}_l)_{\mathcal{L}^2} \\
&\quad + \frac{1}{n} \sum_{j=1}^{n_s} 2d_j (u_{l,j}, 1)_{\mathcal{L}^2} + \frac{1}{n} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} \nu_{ij}^2 (K_{l,i}(\mathbf{u}_l), 1)_{\mathcal{L}^2}.
\end{aligned} \tag{2.119}$$

Recall that we denote by  $\mathbf{v}_l(t)$  the solutions of the approximating problem (2.60), and consider the function  $h(\cdot, \mathbf{w}_l) : [0, T] \rightarrow \mathbb{R}$  with  $\mathbf{w}_l \in \mathcal{H}_0^1$  as parameter given by

$$t \mapsto h(t, \mathbf{w}_l) = \|\mathbf{w}_l - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2. \tag{2.120}$$

Note that

$$\begin{aligned} h'(t, \mathbf{w}_l) &= -2 \left( \mathbf{v}'_l(t), \mathbf{w}_l - \mathbf{v}_l(t) \right)_{\mathcal{L}^2} \\ &= 2 a_l(\mathbf{v}_l(t), \mathbf{w}_l - \mathbf{v}_l(t)) - 2 \left( \mathbf{f}(\mathbf{v}_l(t)), \mathbf{w}_l - \mathbf{v}_l(t) \right)_{\mathcal{L}^2}. \end{aligned} \quad (2.121)$$

2. Consider now the function  $\Phi : S_l \times [0, T] \rightarrow \mathbb{R}$  given by

$$(\mathbf{u}_l, t) \mapsto \Phi(\mathbf{u}_l, t) = \|\mathbf{u}_l - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2. \quad (2.122)$$

It follows by the same arguments as in the proof of Lemma 2.3.8 that the process

$$\begin{aligned} M_l(t) &= \|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2 - \|\mathbf{u}_l(0) - \mathbf{v}_l(0)\|_{\mathcal{L}^2}^2 \\ &\quad - \int_0^t (L_l \Phi(\mathbf{u}_l(s), s) + \partial_s \Phi(\mathbf{u}_l(s), s)) ds \\ &= \|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2 - \|\mathbf{u}_l(0) - \mathbf{v}_l(0)\|_{\mathcal{L}^2}^2 \\ &\quad - \int_0^t (L_l g(\mathbf{u}_l(s), \mathbf{v}_l(s)) + h'(s, \mathbf{u}_l(s))) ds, \end{aligned} \quad (2.123)$$

$t \leq T$ , stopped at  $\tau_p$  is a martingale for each  $p \in \mathbb{N}$ . Plugging the explicit computations above into Eq. (2.123) yields Eq. (2.114).  $\square$

We are now ready to finish also the proof of the law of large numbers for the general linear model.

**Proof of Theorem 2.4.4.** Let  $\hat{d} = \max_{j=1, \dots, n_s} d_j$ . By stopping the local martingale (2.114) and taking expectations we get the estimate

$$\begin{aligned} &E \left[ \|\mathbf{u}_l(t \wedge \tau_p) - \mathbf{v}_l(t \wedge \tau_p)\|_{\mathcal{L}^2}^2 \right] + E \int_0^{t \wedge \tau_p} a_l(\mathbf{u}_l(s) - \mathbf{v}_l(s), \mathbf{u}_l(s) - \mathbf{v}_l(s)) ds \\ &\leq \|\mathbf{u}_l(0) - \mathbf{v}_l(0)\|_{\mathcal{L}^2}^2 + E \int_0^{t \wedge \tau_p} \left| \left( \mathbf{f}_l(\mathbf{u}_l(s)) - \mathbf{f}(\mathbf{v}_l(s)), \mathbf{u}_l(s) - \mathbf{v}_l(s) \right)_{\mathcal{L}^2} \right| ds \\ &\quad + \frac{1}{n} \sum_{j=1}^{n_s} 2d_j E \int_0^{t \wedge \tau_p} (u_{l,j}(s), 1)_{\mathcal{L}^2} ds \\ &\quad + \frac{1}{n} \sum_{j=1}^{n_s} \sum_{i=1}^{n_r} \nu_{ij}^2 E \int_0^{t \wedge \tau_p} (K_i(\mathbf{u}_l(s)), 1)_{\mathcal{L}^2} ds. \end{aligned} \quad (2.124)$$

From the monotonicity of the operator  $A_l$  associated to the bilinear form  $a_l(\cdot, \cdot)$  and a few elementary estimates it follows that

$$\begin{aligned} &E \left[ \|\mathbf{u}_l(t \wedge \tau_p) - \mathbf{v}_l(t \wedge \tau_p)\|_{\mathcal{L}^2}^2 \right] \leq \|\mathbf{u}_l(0) - \mathbf{v}_l(0)\|_{\mathcal{L}^2}^2 + C \sup_{\mathbb{R}^{n_s}} |\mathbf{f}_l - \mathbf{f}| \\ &\quad + C ((\hat{d} + C)/n) \int_0^T \left( \|\mathbf{v}_l(s)\|_{\mathcal{L}^2}^2 + 1 \right) ds \\ &\quad + C \left( (\hat{d} + C)/n + \sup_{\mathbb{R}^{n_s}} |\mathbf{f}_l - \mathbf{f}| + 1 \right) E \int_0^{t \wedge \tau_p} \|\mathbf{u}_l(s) - \mathbf{v}_l(s)\|_{\mathcal{L}^2}^2 ds, \end{aligned} \quad (2.125)$$

where the constant  $C$  does not depend on  $l$ . Note that again

$$\int_0^T E \left[ \|\mathbf{u}_l(s)\|_{\mathcal{L}^2}^2 \right] < \infty. \quad (2.126)$$

By letting  $p \rightarrow \infty$  we deduce from the monotone convergence theorem and Fatou's lemma that the above estimate (2.125) is valid even with  $t \wedge \tau_p$  replaced by  $t$ . Gronwall's inequality yields the estimate

$$\begin{aligned} E \left[ \|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2 \right] &\leq \left( \|\mathbf{u}_l(0) - \mathbf{v}_l(0)\|_{\mathcal{L}^2}^2 + C \sup_{\mathbb{R}^{n_S}} |\mathbf{f}_l - \mathbf{f}| \right. \\ &\quad \left. + C ((\hat{d} + C)/n) \left( \|v_l(s)\|_{L^2(0,T;\mathcal{L}^2)}^2 + T \right) \right) \times \\ &\quad \times \exp \left( C ((\hat{d} + C)/n + \sup_{\mathbb{R}^{n_S}} |\mathbf{f}_l - \mathbf{f}| + 1) T \right). \end{aligned} \quad (2.127)$$

Finally, it follows from the assumptions (S1)–(S3) for the scale parameters and (2.113) that

$$\sup_{t \leq T} E \left[ \|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2 \right] \rightarrow 0.$$

□

## Discussion

Nonlinear mesoscopic stochastic particle models that approximate the solutions of certain semi- and quasilinear parabolic equations will be treated in the next chapter in a similar fashion.

While the scaling relations (S1) and (S2) for the example model, for instance, appear in a natural way, condition (S3) is hard to justify in physical terms. It serves to damp out the fluctuating term in the remainder  $R_l(t)$  (Eq. (2.88)) that stems from diffusive jumps. Stated in terms of  $n$  and  $l$ , condition (S3) reads  $(1/n)/l^2 \rightarrow 0$ . Heuristically,  $1/\sqrt{n}$  is a measure for the size of fluctuations in a single cell. Therefore  $(1/\sqrt{n})/l$  may be interpreted as a measure for the concentration gradients caused by fluctuations. Condition (S3) forces these gradients to vanish asymptotically. The same discussion applies, of course, to the general linear model.

The scaling relation (S3) also appears in Arnold & Theodosopulu (1980) and Kotelenz (1986) in their treatment of single-species models with linear reaction kinetics. In addition, Kotelenz (1986, 1988) is able to prove a law of large numbers in a weaker norm for a single-species model with linear or polynomial kinetics using only (S1) and (S2). Under the same hypotheses Blount (1994) has a stronger result for the model with polynomial kinetics and a law of large numbers for a particular model with constant  $n$ . However, all authors mentioned so far work with particle densities defined on the unit cube in  $\mathbb{R}^m$ , which has the advantage that the eigenvalues and eigenfunctions of the Laplacian are explicitly known. This knowledge is exploited in Blount (1994) to get rid of condition (S3). Guias (2002) assumes

$(1/n)/l^{2+m} \rightarrow 0$  for a law of large numbers for a particular nonlinear two-species model on a general domain.

We have already mentioned in Section 2.3.1 that the main motivation for considering weak solutions of the limit equation is that certain nonlinear PDEs that appear as limit of a stochastic particle model admit solutions that are not differentiable in the classical sense. If, on the other hand, the solution of the limit equations is sufficiently smooth, as is the case for the linear equations of the present chapter provided the data are sufficiently regular, one can obtain explicit rates for the convergence of the solution of the approximating problem  $v_l$  to the solution of the limit equation  $u$ . Let us assume, for instance, that this rate of convergence is  $O(l)$ , and suppose that  $n = O(l^{-\alpha})$  and  $d = O(l^{-\beta})$ , where  $\alpha > \beta > 0$ , in order to satisfy condition (S3). Moreover, assume that  $\|u_l(0) - v_{l,0}\|_{L^2} = O(l)$ . Then it can easily be seen from the estimate (2.104) that the rate of convergence in the law of large numbers for the example model (Theorem 2.3.6) is  $O(l^{\alpha-\beta})$  if  $\beta + 1 > \alpha > \beta$ , and  $O(l)$  if  $\alpha \geq \beta + 1$ . In other words, the rate of convergence is determined by the ratio  $d/n$  if  $\alpha - \beta < 1$ .

## Chapter 3

# Nonlinearities and a refined law of large numbers

### Overview

In the present chapter we study the convergence of certain nonlinear mesoscopic stochastic particle models. We first consider a model with reaction rates that are nonlinear but Lipschitz continuous. It turns out that this constitutes only a slight extension of the general linear case treated in the previous chapter. In Section 3.1.2 we discuss the more realistic case of a vector of reaction functions  $\mathbf{f} = (f_1, \dots, f_{n_s})$  admitting an *invariant region*. More precisely, we shall assume that the vector field  $\mathbf{f}$  points inwards on the boundary of the cube  $[0, 1]^{n_s}$ . This generalises the single-species models with polynomial reaction kinetics treated in Kotelenz (1988) and Blount (1994), and the two-species model discussed in Guiaş (2002). In Section 3.2 we consider two models with a nonlinear diffusion mechanism. For the sake of simplicity, we restrict the discussion to single-species models in one space dimension without chemical reactions. In Section 3.2.1 we investigate what happens when the intensity for a jump of a particle to a neighboring cell depends on the local concentration, i.e.,  $d = d(u_l(z))$ , where  $d(\cdot)$  is monotonously increasing on  $\mathbb{R}_0^+$ . Thereafter, in Section 3.2.2, we have a look at the case where the intensity for a jump to a neighboring cell depends on the absolute value of the (discrete) concentration gradient (i.e.,  $d = d(|\partial^+ u_l(z)|)$  for a jump to the right and  $d = d(|-\partial^- u_l(z)|)$  for a jump to the left, respectively). The limit behaviour of these models has (to the best of our knowledge) not been investigated yet. The first one ( $d = d(u_l(z))$ ) resembles the so-called zero-range process that is extensively investigated in the literature on interacting particle systems (Kipnis & Landim, 1999). Finally, in Section 3.3, we demonstrate for the linear example model from the previous chapter how the law of large numbers obtained there can be refined.

### 3.1 Nonlinear reaction kinetics

In this section we focus on nonlinear reaction kinetics; diffusion is still assumed to be linear.

#### 3.1.1 Lipschitz conditions

The macroscopic PDE and its weak formulation

We use the same notation as in Sections 2.1 and 2.2. Recall that we are dealing with  $n_r$  reactions involving  $n_s$  species. The chemical reactions are described by the reaction functions

$$f_j = \sum_{i=1}^{n_r} \nu_{ij} K_i, \quad j = 1, \dots, n_s,$$

where  $K_i : \mathbb{R}^{n_s} \rightarrow \mathbb{R}$  is the rate at which the  $i$ th reaction proceeds. In the previous chapter the rates were assumed to be linear and to satisfy conditions (C1) and (C2) from Section 2.1.1.

As a warm-up for nonlinear problems we treat here the case of possibly nonlinear but Lipschitz continuous rates  $K_i$ , i.e., in addition to (C1) and (C2) they are supposed to satisfy the condition

$$|K_i(\mathbf{u}) - K_i(\mathbf{v})| \leq c_L \|\mathbf{u} - \mathbf{v}\|, \quad \mathbf{u}, \mathbf{v} \in \mathbb{R}^{n_s}, \quad i = 1, \dots, n_r, \quad (\text{L})$$

for a constant  $c_L > 0$ . As we shall see, this is only a slight generalisation of the linear case. The macroscopic PDE system with Dirichlet boundary conditions still reads

$$\begin{cases} \partial_t u_j - D_j \Delta u_j = f_j(\mathbf{u}) & \text{on } Q_T \\ u_j = 0 & \text{on } \partial G \times [0, T] \\ u_j(\cdot, 0) = u_{j,0} & \text{on } G, \end{cases} \quad (3.1)$$

$j = 1, \dots, n_s$ . Here  $G \subset \mathbb{R}^m$  is a bounded domain with Lipschitz boundary representing the chemical reactor, and  $D_1, \dots, D_{n_s} > 0$  are the diffusion coefficients.

The weak formulation of the PDE system (3.1) is obtained as in the previous chapter by multiplying with a test function  $\mathbf{v} \in C_0^\infty(G)$ , integrating over  $G$  and an integration by parts. We adopt the notation of Sections 2.2 and 2.4 concerning the various continuous and discrete function spaces and define again the (bounded and coercive) bilinear form  $a(\cdot, \cdot)$  on  $\mathbf{H}_0^1(G) \times \mathbf{H}_0^1(G)$  by

$$a(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^{n_s} \sum_{k=1}^m D_j (\partial_{x_k} u_j, \partial_{x_k} v_j)_{L^2(G)}, \quad \mathbf{u}, \mathbf{v} \in \mathbf{H}_0^1(G).$$

Observe that the vector of reaction functions  $\mathbf{f}$  induces a now nonlinear but Lipschitz continuous operator  $F : \mathbf{L}^2(G) \rightarrow \mathbf{L}^2(G)$  by  $F(\mathbf{u})(\cdot) = \mathbf{f}(\mathbf{u}(\cdot))$ . That is,

$$\|F(\mathbf{u}) - F(\mathbf{v})\|_{\mathbf{L}^2} \leq C_L \|\mathbf{u} - \mathbf{v}\|_{\mathbf{L}^2}, \quad \mathbf{u}, \mathbf{v} \in \mathbf{L}^2, \quad (3.2)$$

for a constant  $C_L > 0$ . In the weak formulation of the PDE a function  $\mathbf{u} \in H^1(0, T; \mathbf{H}_0^1, \mathbf{L}^2)$  is sought such that

$$\frac{d}{dt}(\mathbf{u}(t), \mathbf{v})_{\mathbf{L}^2} + a(\mathbf{u}(t), \mathbf{v}) = (\mathbf{f}(\mathbf{u}), \mathbf{v})_{\mathbf{L}^2} \quad (3.3a)$$

for all  $\mathbf{v} \in \mathbf{H}_0^1$  and a.e.  $t \in (0, T)$ , and

$$\mathbf{u}(0) = \mathbf{u}_0 \in \mathbf{L}^2. \quad (3.3b)$$

The equivalent operator equation in the space  $H^1(0, T; \mathbf{H}_0^1, \mathbf{L}^2)$  reads

$$\mathbf{u}' + A\mathbf{u} = F(\mathbf{u}), \quad \mathbf{u}(0) = \mathbf{u}_0 \in \mathbf{L}^2. \quad (3.4)$$

The weak problem can be solved with the Faedo-Galerkin method in combination with the Aubin-Lions compactness theorem, the continuous analogue of Theorem 2.3.4 (see, e.g., Lions (1969), Section 5 of Chapter 1). Alternatively, one can use the existence result for linear equations together with the Banach fixed-point theorem (Evans, 1998). Moreover, the convergence proof for the solutions of the approximating problem below can easily be extended to an existence proof. Note that the uniqueness of a solution of (3.3) can readily be established because the operator  $F$  is Lipschitz continuous.

The approximating problem

Let  $\mathcal{G}_h$  be the interior lattice points of the domain  $G$  representing the chemical reactor generated by a lattice  $\mathcal{Z}_h(z_0)$  (cf. Section 2.1.2), and recall that  $\mathcal{L}^2(\mathcal{G}_h)$  and  $\mathcal{H}_0^1(\mathcal{G}_h)$  are the discrete versions of  $\mathbf{L}^2(G)$  and  $\mathbf{H}_0^1(G)$ . The discrete analogue of the PDE system is again given by

$$\begin{cases} u'_{h,j} - D_{h,j} \nabla^- \cdot \nabla^+ u_{h,j} = f_j(\mathbf{u}_h) & \text{on } \mathcal{G}_h^1 \times (0, T) \\ u_{h,j} = 0 & \text{on } (\mathcal{G}_h \setminus \mathcal{G}_h^1) \times [0, T] \\ u_{h,j} = u_{h,j,0} & \text{on } \mathcal{G}_h^1, \end{cases} \quad (3.5)$$

$j = 1, \dots, n_s$ , which is now an initial-value problem for a (nonlinear) finite-dimensional system of ODEs with Lipschitz continuous right-hand side. Hence, it has a unique local solution according to the Picard-Lindelöf theorem. Existence of a unique solution on the entire interval  $[0, T]$  follows from the derivation of the a-priori estimate (3.11) below.

The discrete version  $a_h(\cdot, \cdot)$  of the bilinear form  $a(\cdot, \cdot)$  is again

$$a_h(\mathbf{u}_h, \mathbf{v}_h) = \sum_{j=1}^{n_s} \sum_{k=1}^m D_{h,j} (\partial_k^+ u_{h,j}, \partial_k^+ v_{h,j})_{\mathcal{L}^2}, \quad \mathbf{u}_h, \mathbf{v}_h \in \mathcal{H}_0^1, \quad (3.6)$$

and the solution of (3.5) can be regarded as a function in  $C^1([0, T], \mathcal{H}_0^1)$  that satisfies the discrete weak formulation

$$\frac{d}{dt}(\mathbf{u}_h(t), \mathbf{v}_h)_{\mathcal{L}^2} + a_h(\mathbf{u}_h(t), \mathbf{v}_h) = (\mathbf{f}(\mathbf{u}_h(t)), \mathbf{v}_h)_{\mathcal{L}^2} \quad (3.7)$$

for all  $\mathbf{v}_h \in \mathcal{H}_0^1$  and  $t \in (0, T)$ .

### The mesoscopic stochastic particle model

The state space  $S_l$  of the particle density process  $(\mathbf{u}_l(t))_{t \geq 0}$  is the same as for the general linear model from Section 2.1.2. That is,  $S_l$  is the set of vector lattice functions in  $\mathcal{L}^2(\mathcal{G}_l)$  taking values in  $\frac{1}{n}\mathbb{N}_0^{n_s}$ . We still assume Dirichlet boundary conditions, i.e.,  $\mathbf{u}_l(z, \cdot) \equiv 0$  if  $z \in \mathcal{G}_l \setminus \mathcal{G}_l^1$ , and that  $\mathbf{u}_l(0)$  is non-random. The transitions intensities  $q_l(\cdot, \cdot)$  are formally the same as in Section 2.1.2.

- A particle of species  $j$  may leave cell  $z \in \mathcal{G}_l^1$  and jump to  $z \pm le_k$ .

$$\begin{aligned} q_l(\mathbf{u}_l, \mathbf{u}_l - \frac{1}{n}\chi_{j,z} + \frac{1}{n}\chi_{j,(z-le_k)}) &= n \frac{d_j}{2m} u_{l,j}(z), \\ q_l(\mathbf{u}_l, \mathbf{u}_l - \frac{1}{n}\chi_{j,z} + \frac{1}{n}\chi_{j,(z+le_k)}) &= n \frac{d_j}{2m} u_{l,j}(z), \end{aligned} \quad (3.8)$$

where  $d_j > 0$ .

- The number of particles in cell  $z \in \mathcal{G}_l^1$  changes according to reaction  $i$ .

$$q_l(\mathbf{u}_l, \mathbf{u}_l + \frac{1}{n} \sum_{j=1}^{n_s} \nu_{ij} \chi_{j,z}) = n K_i(\mathbf{u}_l(z)) \quad \text{if } \mathbf{u}_l + \frac{1}{n} \sum_{j=1}^{n_s} \nu_{ij} \chi_{j,z} \in S_l. \quad (3.9)$$

The rates  $K_i$  are now nonlinear but satisfy the Lipschitz condition (L).

The existence of a particle density process  $(\mathbf{u}_l(t))_{t \geq 0}$  with state space  $S_l$  and generator  $L_l$  given by

$$L_l g(\mathbf{u}_l) = \sum_{\tilde{\mathbf{u}}_l \neq \mathbf{u}_l} q_l(\mathbf{u}_l, \tilde{\mathbf{u}}_l) (g(\tilde{\mathbf{u}}_l) - g(\mathbf{u}_l)), \quad g \in \hat{C}(S_l), \quad (3.10)$$

follows as for the general linear model from Theorem 3.1 in Chapter 8 of Ethier & Kurtz (1986) by making use of the Lipschitz conditions (L).

### Law of large numbers

We first show convergence of the semi-discrete approximation. (Recall the definition of the external approximation of the space  $\mathbf{H}_0^1(G)$  in terms of the spaces  $\mathcal{H}_0^1(\mathcal{G}_h)$  as discussed in Sections 2.2.3 and 2.4.)

**Theorem 3.1.1.** *Let  $\mathbf{u}$  be the solution of the weak PDE problem (3.3) to the initial value  $\mathbf{u}_0$ , and let  $(\mathbf{u}_h)$ ,  $h \searrow 0$ , be a sequence of solutions of the approximating problem (3.5) to the initial values  $\mathbf{u}_{h,0}$ . Assume  $\mathbf{u}_{h,0} \rightarrow \mathbf{u}_0$  strongly in  $\mathbf{L}^2$ , and  $D_{h,j} \rightarrow D_j$ ,  $j = 1, \dots, n_s$ . Then  $J_h \mathbf{u}_h$  converges weakly to  $J\mathbf{u}$  in  $L^2(0, T; X)$ , and  $\mathbf{u}_h$  converges strongly to  $\mathbf{u}$  in  $L^2(0, T; \mathbf{L}^2)$ .*

**Proof.** The proof is similar to the proof for linear reaction rates.

1. We first derive the a-priori estimates

$$\sup_h \max_{0 \leq t \leq T} \|\mathbf{u}_h\|_{\mathcal{L}^2} < \infty, \quad (3.11)$$

$$\sup_h \|\mathbf{u}_h\|_{L^2(0,T;\mathcal{H}_0^1)} < \infty, \quad (3.12)$$

$$\sup_h \|\mathbf{u}'_h\|_{L^2(0,T;\mathcal{H}^{-1})} < \infty. \quad (3.13)$$

By inserting  $\mathbf{u}_h(t)$  for  $\mathbf{v}_h$  in the discrete weak formulation, integrating over time and making use of the coerciveness of the bilinear form  $a_h(\cdot, \cdot)$ , we get the estimate

$$\|\mathbf{u}_h(t)\|_{\mathcal{L}^2}^2 + 2\alpha \int_0^t \|\mathbf{u}_h(s)\|_{\mathcal{H}_0^1}^2 ds \leq \|\mathbf{u}_h(0)\|_{\mathcal{L}^2}^2 + 2 \int_0^t (\mathbf{f}(\mathbf{u}_h(s)), \mathbf{u}_h(s))_{\mathcal{L}^2} ds.$$

Since the operator  $F : \mathbf{L}^2 \rightarrow \mathbf{L}^2$  induced by  $\mathbf{f}$  is Lipschitz continuous, we can deduce the estimate

$$\|\mathbf{u}_h(t)\|_{\mathcal{L}^2}^2 + 2\alpha \int_0^t \|\mathbf{u}_h(s)\|_{\mathcal{H}_0^1}^2 ds \leq \|\mathbf{u}_h(0)\|_{\mathcal{L}^2}^2 + C \int_0^t (1 + \|\mathbf{u}_h(s)\|_{\mathcal{L}^2}^2) ds,$$

where  $C > 0$  is a constant independent of  $h$ . Gronwall's inequality yields

$$\sup_h \max_{0 \leq t \leq T} \|\mathbf{u}_h\|_{\mathcal{L}^2} < \infty. \quad (3.14)$$

Hence,

$$\sup_h \|\mathbf{u}_h\|_{L^2(0,T;\mathcal{H}_0^1)} < \infty. \quad (3.15)$$

The third a-priori estimate follows from the boundedness of the bilinear form  $a_h(\cdot, \cdot)$  and the Lipschitz continuity of the operator  $F$ .

2. Owing to the a-priori estimates we can extract a subsequence (still denoted by  $(\mathbf{u}_h)$ ) such that  $J_h \mathbf{u}_h \rightharpoonup J\mathbf{u}$  in  $L^2(0,T;X)$  for a  $\mathbf{u} \in L^2(0,T;\mathbf{H}_0^1)$ . In addition, we may assume that  $\mathbf{u}_h$  converges strongly to  $\mathbf{u}$  in  $L^2(0,T;\mathbf{L}^2)$  due to Theorem 2.3.4. From the discrete weak formulation it follows by integration by parts that

$$\begin{aligned} -(\mathbf{u}_{h,0}, \mathbf{v}_h)_{\mathcal{L}^2} \varphi(0) - \int_0^T (\mathbf{u}_h(t), \mathbf{v}_h)_{\mathcal{L}^2} \varphi'(t) dt \\ = - \int_0^T a_h(\mathbf{u}_h(t), \mathbf{v}_h) \varphi(t) dt + \int_0^T (\mathbf{f}(\mathbf{u}_h(t)), \mathbf{v}_h)_{\mathcal{L}^2} \varphi(t) dt \end{aligned} \quad (3.16)$$

for all  $\mathbf{v} \in \mathbf{H}_0^1$  and  $\varphi \in C^1[0,T]$  with  $\varphi(T) = 0$ , where we have set  $\mathbf{v}_h = R_h \mathbf{v}$ . Since the operator  $F$  is Lipschitz continuous as mapping from the space  $L^2(0,T;\mathbf{L}^2)$  to itself, we can pass to the limit in all the terms (see also Lemma 2.4.1), which yields

$$\begin{aligned} -(\mathbf{u}(0), \mathbf{v})_{\mathcal{L}^2} \varphi(0) - \int_0^T (\mathbf{u}(t), \mathbf{v})_{\mathcal{L}^2} \varphi'(t) dt \\ = - \int_0^T a(\mathbf{u}(t), \mathbf{v}) \varphi(t) dt + \int_0^T (\mathbf{f}(\mathbf{u}(t)), \mathbf{v})_{\mathcal{L}^2} \varphi(t) dt. \end{aligned} \quad (3.17)$$

Now it follows by the same arguments as in the proof of Theorem 2.3.2 that  $\mathbf{u}$  is indeed the solution of (3.3) to the initial condition  $\mathbf{u}_0$ .

4. Since the solution of (3.3) is unique, the whole sequence converges.  $\square$

We denote from now on the stochastic particle density by  $\mathbf{u}_l$  and the solutions of the approximating problem (3.5) by  $\mathbf{v}_l$ . (Recall that  $l$  is the edge length of a cell.)

As in Section 2.4 we replace the reaction rates  $K_i$  by measurable functions  $K_{l,i} : \mathbb{R}^{n_s} \rightarrow \mathbb{R}$  that converge uniformly to the  $K_i$  so that the transition intensities  $q_l(\cdot, \cdot)$  are left unchanged and forbidden reactive jumps (jumps that would leave the state space) are automatically excluded. (Cf. the discussion in Section 2.4.2.) We assume the same scaling relations as for the general linear model.

$$l \rightarrow 0, \quad n \rightarrow \infty, \tag{S1}$$

$$\frac{d_j}{2m} l^2 \rightarrow D_j, \tag{S2}$$

$$\frac{d_j}{n} \rightarrow 0, \tag{S3}$$

$j = 1, \dots, n_s$ , where the  $D_j > 0$  are the macroscopic diffusion coefficients. We obtain the same type of law of large numbers as for the case of linear reaction rates.

**Theorem 3.1.2** (Law of large numbers). *Let  $\mathbf{u}$  be the solution of the weak PDE problem (3.3) to the initial value  $\mathbf{u}_0$ . Assume that (S1)–(S3) hold and that  $\mathbf{u}_l(0)$  converges strongly to  $\mathbf{u}_0$  in  $\mathbf{L}^2$ . Then*

$$E \left[ \|\mathbf{u}_l - \mathbf{u}\|_{L^2(0,T;\mathbf{L}^2)}^2 \right] = E \left[ \|\mathbf{u}_l - \mathbf{u}\|_{(L^2(Q_T))^{n_s}}^2 \right] \rightarrow 0.$$

The law of large numbers follows immediately from the next result.

**Theorem 3.1.3.** *Assume (S1)–(S3), and denote by  $\mathbf{v}_l$  be the solutions of the approximating problem (3.5) with  $D_{l,j} = \frac{d_j}{2m} l^2$ ,  $j = 1, \dots, n_s$ , to the initial values  $\mathbf{v}_{l,0}$ . Moreover, assume that*

$$\|\mathbf{u}_l(0) - \mathbf{v}_{l,0}\|_{\mathcal{L}^2} \rightarrow 0.$$

Then

$$\sup_{t \leq T} E \left[ \|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2 \right] \rightarrow 0.$$

As in the previous chapter, the proof of Theorem 3.1.3 is based on a lemma that identifies a local martingale related to the process  $\|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2$ ,  $t \leq T$ . We define again for  $p \in \mathbb{N}$  the stopping times  $\tau_p$  as in (2.111).

**Lemma 3.1.4.** *Let  $(M_l(t))_{t \leq T}$  be the process given by*

$$\begin{aligned} M_l(t) &= \|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2}^2 - \|\mathbf{u}_l(0) - \mathbf{v}_l(0)\|_{\mathcal{L}^2}^2 \\ &\quad + 2 \int_0^t a_l(\mathbf{u}_l(s) - \mathbf{v}_l(s), \mathbf{u}_l(s) - \mathbf{v}_l(s)) ds \\ &\quad - 2 \int_0^t (\mathbf{f}_l(\mathbf{u}_l(s)) - \mathbf{f}_l(\mathbf{v}_l(s)), \mathbf{u}_l(s) - \mathbf{v}_l(s))_{\mathcal{L}^2} ds - R_l(t), \end{aligned} \tag{3.18}$$

where

$$R_l(t) = \frac{1}{n} \sum_{j=1}^{n_s} 2d_j \int_0^t (u_{l,j}(s), 1)_{\mathcal{L}^2} ds + \frac{1}{n} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} \nu_{ij}^2 \int_0^t (K_{l,i}(\mathbf{u}_l(s)), 1)_{\mathcal{L}^2} ds. \quad (3.19)$$

Then the stopped process  $(M_l(t \wedge \tau_p))_{t \leq T}$  is a martingale for each  $p \in \mathbb{N}$ .

**Proof.** The proof is very similar to the proof of Lemma 2.4.5, since the linearity of  $\mathbf{f}$  was not used there.  $\square$

**Proof of Theorem 3.1.3.** The proof is almost identical to the proof of Theorem 2.4.4. Note that, in order to obtain the essential estimate (2.125), it is sufficient that  $\mathbf{f}$  be Lipschitz continuous. Moreover, we can still control the growth of the process  $\|\mathbf{u}_l(t)\|_{\mathcal{L}^2}$ ,  $t \geq 0$ , by an appropriate Yule process.  $\square$

### 3.1.2 Invariant regions

In this section we treat a class of nonlinear reaction rates that is more often met in practice than the Lipschitz continuous rates of the previous section. We suppose that the vector field  $\mathbf{f} = (f_1, \dots, f_{n_s})$  admits an *invariant region*. Let us explain this in more detail. In practice, the reaction rates  $K_i$  are rarely globally Lipschitz continuous functions, since they are often polynomials in the concentrations of quadratic or even higher order. However, by the structure of the chemical equations, an explosion of the concentrations is usually avoided, whatever nonlinear the reaction rates may be. The model for CO oxidation on Pt that has been developed in Sections 1.2 and 1.3, for instance, involves reaction rates that are not globally Lipschitz continuous functions, but the concentrations are obviously confined to the cube  $[0, 1]^3$ . This motivates assumption (I) below.

We still assume that all reaction rates  $K_i : \mathbb{R}^{n_s} \rightarrow \mathbb{R}$  are *locally* Lipschitz continuous functions that satisfy conditions (C1) and (C2) from Section 2.1.1. In addition, we now assume the following condition (for  $i = 1, \dots, n_r$ ,  $j = 1, \dots, n_s$ ).

$$\text{If } \nu_{ij} > 0 \text{ then } K_i(\mathbf{w}) = 0 \text{ for all } \mathbf{w} \in [0, 1]^{n_s} \text{ with } w_j = 1. \quad (\text{I})$$

Hence, the vector field  $\mathbf{f} : \mathbb{R}^{n_s} \rightarrow \mathbb{R}^{n_s}$  given by

$$f_j = \sum_{i=1}^{n_r} \nu_{ij} K_i, \quad j = 1, \dots, n_s,$$

points inwards (although, in general, not strictly) everywhere on the boundary of the cube  $[0, 1]^{n_s}$ . The physical interpretation of condition (I) is that there exists an upper limit for each species above of which immigration to a cell from the outside or creation of new particles of the species by chemical reactions is prohibited due to lack of space. The choice of the upper bound one is, of course, arbitrary and could be replaced by any other positive constant (or a different constant for each species).

In addition, we admit in this section that some of the diffusion coefficients  $D_j$ ,  $j = 1, \dots, n_s$ , be zero. We assume that the first  $n_d$  species do diffuse (have diffusion coefficient greater than zero) and the remaining  $n_s - n_d$  species are immobile.

The macroscopic PDE and its weak formulation

The system of PDEs (with Dirichlet boundary conditions for the diffusing species) that describes the dynamics of the concentrations on the macroscopic level is given by

$$\begin{cases} \partial_t u_j - D_j \Delta u_j = f_j(\mathbf{u}) & \text{on } Q_T \\ u_j = 0 & \text{on } \partial G \times [0, T] \\ u_j(\cdot, 0) = u_{j,0} & \text{on } G \end{cases} \quad (3.20a)$$

( $j = 1, \dots, n_d$ ), and

$$\begin{cases} \partial_t u_j = f_j(\mathbf{u}) & \text{on } Q_T \\ u_j(\cdot, 0) = u_{j,0} & \text{on } G \end{cases} \quad (3.20b)$$

( $j = n_d + 1, \dots, n_s$ ). The PDE system (3.20) is usually not treated with energy methods because it is not quite obvious how to use the hypothesis on the vector field  $\mathbf{f}$  to obtain a-priori estimates. Existence of a solution is proved, e.g., in Smoller (1983), Chapter 14, in a somewhat different functional setting by showing local existence with a fixed-point argument and extending the local result to a global one by making use of invariant regions. Although spatial semi-discretisation with finite differences is usually less convenient than the Faedo-Galerkin method, we shall see that for the particular case of the present section the necessary a-priori estimates can easily be obtained and the difficulties that one encounters with the Faedo-Galerkin method do not arise. We include an existence result in the proof of convergence of the solution of the approximating problem (3.25) below.

Note that because of condition (I) it only makes sense to look for solutions  $u_j$  of (3.20) that take values in the interval  $[0, 1]$ . We set

$$Y = \{u \in L^\infty(G) : 0 \leq u \leq 1 \text{ a.e. on } G\}, \quad (3.21)$$

$$Z = \{u \in L^\infty(Q_T) : 0 \leq u \leq 1 \text{ a.e. on } Q_T\}. \quad (3.22)$$

In the weak formulation of the PDE system (3.20) we seek functions  $u_j \in H^1(0, T; H_0^1, L^2) \cap Z$  ( $j = 1, \dots, n_d$ ), respectively  $u_j \in H^1(0, T; H^1, L^2) \cap Z$  ( $j = n_d + 1, \dots, n_s$ ), such that the following equations are satisfied:

$$\frac{d}{dt} (u_j(t), v_j)_{L^2} + D_j (\nabla u_j(t), \nabla v_j)_{(L^2)^m} = (f_j(\mathbf{u}(t)), v_j)_{L^2} \quad (3.23a)$$

for all  $v_j \in H_0^1$  and a.e.  $t \in (0, T)$ , and

$$u_j(0) = u_{j,0} \in L^2 \cap Y \quad (3.23b)$$

( $j = 1, \dots, n_d$ ), respectively

$$\frac{d}{dt}(u_j(t), v_j)_{L^2} = (f_j(\mathbf{u}(t)), v_j)_{L^2} \quad (3.23c)$$

for all  $v_j \in H^1$  and a.e.  $t \in (0, T)$ , and

$$u_j(0) = u_{j,0} \in H^1 \cap Y \quad (3.23d)$$

( $j = n_d + 1, \dots, n_s$ ). Here, in general, the reaction functions  $f_j$  do not induce a Lipschitz continuous operator  $F : \mathbf{L}^2 \rightarrow \mathbf{L}^2$  by  $F(\mathbf{u})(\cdot) = \mathbf{f}(\mathbf{u}(\cdot))$ , but the operator  $F$  thus defined is Lipschitz continuous as mapping from  $(L^2 \cap Y)^{n_s}$  to  $\mathbf{L}^2$  because the reaction rates are still locally Lipschitz continuous. That is,

$$\|F(\mathbf{u}) - F(\mathbf{v})\|_{\mathbf{L}^2} \leq C \|\mathbf{u} - \mathbf{v}\|_{\mathbf{L}^2} \quad \text{for all } \mathbf{u}, \mathbf{v} \in (L^2 \cap Y)^{n_s}. \quad (3.24)$$

Therefore a solution of (3.23) is unique.

The approximating problem

Let  $\mathcal{G}_h$  be the interior lattice points of the bounded domain  $G \subset \mathbb{R}^m$  with Lipschitz boundary generated by the lattice  $\mathcal{Z}_h(z_0)$  (cf. Section 2.1.2). Since now some of the functions  $u_j$  that solve the PDE system (3.23) are at a fixed time elements of the space  $H^1(G)$ , we have to use a larger grid to define an appropriate approximation.

**Definition 3.1.5.** We define the sets of lattice points  $\bar{\mathcal{G}}_h$  and  $\bar{\mathcal{G}}_h^1$  as

$$\begin{aligned} \bar{\mathcal{G}}_h^1 &= \left\{ z \in \mathcal{Z}_h(z_0) : c_h(z) \cap G \neq \emptyset \right\}, \\ \bar{\mathcal{G}}_h &= \left\{ z \pm h e_k : z \in \bar{\mathcal{G}}_h^1, k = 1, \dots, m \right\}. \end{aligned}$$

We are now going to define a second internal approximation of  $L^2(G)$  and an external approximation of the Sobolev space  $H^1(G)$  in terms of lattice functions that are defined on the grid  $\bar{\mathcal{G}}_h$ . (Recall the definition of internal and external approximation from Section 2.2.3.) We still do not distinguish in notation between lattice functions, their extended versions and the restrictions of the extended versions to the domain  $G$  (cf. Remark 2.1.4).

**Definition 3.1.6.** The *discrete Lebesgue space*  $\mathcal{L}^2(\bar{\mathcal{G}}_h)$  is the space of lattice functions that vanish outside  $\bar{\mathcal{G}}_h$  equipped with the scalar product

$$(u_h, v_h)_{\mathcal{L}^2(\bar{\mathcal{G}}_h)} = h^m \sum_{z \in \bar{\mathcal{G}}_h} u_h(z) v_h(z) = \int_{\mathbb{R}^m} u_h(x) v_h(x) dx.$$

Note that here, in general,

$$(u_h, v_h)_{\mathcal{L}^2(\bar{\mathcal{G}}_h)} \neq \int_G u_h(x) v_h(x) dx.$$

The discrete derivatives  $\partial_k^\pm u_h$  are defined in the same way as before. We are now in a position to define the discrete Sobolev space  $\mathcal{H}^1(\bar{\mathcal{G}}_h)$ , the discrete analogue of  $H^1(G)$ .

**Definition 3.1.7.** By the *discrete Sobolev space*  $\mathcal{H}^1(\bar{\mathcal{G}}_h)$  we understand the space of lattice functions that vanish outside  $\bar{\mathcal{G}}_h$  equipped with the scalar product

$$(u_h, v_h)_{\mathcal{H}^1(\bar{\mathcal{G}}_h)} = (u_h, v_h)_{\mathcal{L}^2(\bar{\mathcal{G}}_h)} + h^m \sum_{k=1}^m \sum_{z \in \bar{\mathcal{G}}_h^1} \partial_k^+ u_h(z) \partial_k^+ v_h(z).$$

The approximating problem is given by the following system of ODEs.

$$\begin{cases} u'_{h,j} - D_{h,j} \nabla^- \cdot \nabla^+ u_{h,j} = f_j(\mathbf{u}_h) & \text{on } \mathcal{G}_h^1 \times (0, T) \\ u_{h,j} = 0 & \text{on } (\bar{\mathcal{G}}_h \setminus \mathcal{G}_h^1) \times [0, T] \\ u_{h,j}(\cdot, 0) = u_{h,j,0} & \text{on } \mathcal{G}_h^1 \end{cases} \quad (3.25a)$$

( $j = 1, \dots, n_d$ ), and

$$\begin{cases} u'_{h,j} = f_j(\mathbf{u}_h) & \text{on } \bar{\mathcal{G}}_h \times (0, T) \\ u_{h,j} = u_{h,j,0} & \text{on } \bar{\mathcal{G}}_h \end{cases} \quad (3.25b)$$

( $j = n_d + 1, \dots, n_s$ ). The solutions of (3.25) can, for  $j = 1, \dots, n_d$ , be regarded as elements of  $C^1([0, T], \mathcal{H}_0^1(\mathcal{G}_h))$  and, for  $j = n_d + 1, \dots, n_s$ , as elements of  $C^1([0, T], \mathcal{H}^1(\bar{\mathcal{G}}_h))$ , respectively. They also solve the following discrete analogue of the weak PDE problem (3.23).

$$\frac{d}{dt} (u_{h,j}(t), v_{h,j})_{\mathcal{L}^2(\mathcal{G}_h)} + D_{h,j} (\nabla^+ u_{h,j}(t), \nabla^+ v_{h,j})_{(\mathcal{L}^2(\mathcal{G}_h))^m} = (f_j(\mathbf{u}_h(t)), v_{h,j})_{\mathcal{L}^2(\bar{\mathcal{G}}_h)} \quad (3.26a)$$

for all  $v_{h,j} \in \mathcal{H}_0^1(\mathcal{G}_h)$  and  $t \in (0, T)$  ( $j = 1, \dots, n_d$ ), respectively

$$\frac{d}{dt} (u_{h,j}(t), v_{h,j})_{\mathcal{L}^2(\bar{\mathcal{G}}_h)} = (f_j(\mathbf{u}_h(t)), v_{h,j})_{\mathcal{L}^2(\bar{\mathcal{G}}_h)} \quad (3.26b)$$

for all  $v_{h,j} \in \mathcal{H}^1(\bar{\mathcal{G}}_h)$  and  $t \in (0, T)$  ( $j = n_d + 1, \dots, n_s$ ).

### The mesoscopic stochastic particle model

As state space  $S_l$  of the particle density process  $(\mathbf{u}_l(t))_{t \geq 0}$  we use the discrete subset of vector lattice functions from  $\mathcal{L}^2(\bar{\mathcal{G}}_l) = (\mathcal{L}^2(\bar{\mathcal{G}}_l))^{n_s}$  that take values in  $\frac{1}{n} \mathbb{N}_0^{n_s}$ . We assume Dirichlet boundary conditions for the diffusing species, i.e.,  $u_{l,j}(z, \cdot) \equiv 0$  for  $z \in \bar{\mathcal{G}}_l \setminus \mathcal{G}_l^1$  ( $j = 1, \dots, n_d$ ). Furthermore, it is supposed that  $\mathbf{u}_l(0)$  is non-random and that  $0 \leq u_{l,j}(z, 0) \leq 1$  for all  $z \in \bar{\mathcal{G}}_l$ ,  $j = 1, \dots, n_s$ . Let, for  $j = 1, \dots, n_s$  and  $z \in \mathcal{G}_l^1$ ,  $\chi_{j,z}$  be the vector lattice function that has  $j$ th component equal to one in the point  $z$  and is equal to zero elsewhere. For  $z \in \bar{\mathcal{G}}_l \setminus \mathcal{G}_l^1$ , let  $\chi_{j,z} \equiv 0$ . In addition we define, for  $j = 1, \dots, n_s$  and  $z \in \bar{\mathcal{G}}_l$ ,  $\bar{\chi}_{j,z}$  as the vector lattice function that is zero everywhere except for the  $j$ th component in the point  $z$  which is equal to one.

Here we exclude immigration of particles of a certain species to a cell from the reservoirs and creation of new particles of the species by chemical reactions if there are already  $n$  or more individuals of the same species present (i.e., the concentration is already greater than

or equal to one). However, for the sake of simplicity, we do not suppress immigration to a cell from neighbouring cells by diffusion, even if the particle density in the cell is already greater than or equal to one. That is, we allow overshooting of the concentration by diffusion. A particular two-species model that excludes overshooting has been discussed in Guiaş (2002), and the exclusion condition did not alter the limit equation. From a physical point of view it should not matter if an exclusion condition is used or not, since the number  $n$  of sites per cell is in practice relatively large (about  $10^3$ ) and not strictly fixed.

As usual, we may assume that (by possibly modifying the original  $K_i$ ) there are measurable functions  $K_{l,i} : \mathbb{R}^{n_s} \rightarrow \mathbb{R}$  such that the transition intensities are left unchanged and intensity zero is automatically assigned to forbidden reactive jumps (i.e., jumps that would lead to negative concentrations or concentrations greater than one). More precisely, for  $\mathbf{w} \in \{0, 1/n, 2/n, \dots, 1\}^{n_s}$ ,

$$K_{l,i}(\mathbf{w}) = \begin{cases} K_i(\mathbf{w}) & \text{if } 0 \leq w_j + \frac{1}{n}\nu_{ij} \leq 1 \text{ for all } j = 1, \dots, n_s \\ 0 & \text{otherwise.} \end{cases} \quad (3.27)$$

Moreover, we may assume that  $K_{l,i} \rightarrow K_i$  uniformly on  $\mathbb{R}^{n_s}$  for  $l \rightarrow 0$ .

The transition intensities are the following.

- A particle of species  $j$  may leave cell  $z \in \mathcal{G}_l^1$  and jump to  $z \pm le_k$ .

$$\begin{aligned} q_l(\mathbf{u}_l, \mathbf{u}_l - \frac{1}{n}\chi_{j,z} + \frac{1}{n}\chi_{j,(z-le_k)}) &= n \frac{d_j}{2m} u_{l,j}(z), \\ q_l(\mathbf{u}_l, \mathbf{u}_l - \frac{1}{n}\chi_{j,z} + \frac{1}{n}\chi_{j,(z+le_k)}) &= n \frac{d_j}{2m} u_{l,j}(z), \end{aligned} \quad (3.28)$$

where  $d_j > 0$ .

- The number of particles in cell  $z \in \mathcal{G}_l^1$  changes according to the  $i$ th reaction.

$$q_l(\mathbf{u}_l, \mathbf{u}_l + \frac{1}{n}\sum_{j=1}^{n_s}\nu_{ij}\chi_{j,z}) = n K_{l,i}(\mathbf{u}_l(z) \wedge 1) \quad \text{if } \mathbf{u}_l + \frac{1}{n}\sum_{j=1}^{n_s}\nu_{ij}\chi_{j,z} \in S_l. \quad (3.29)$$

Here  $\mathbf{w} \wedge 1$  for  $\mathbf{w} \in \mathbb{R}^{n_s}$  is defined componentwise.

- The number of particles in a boundary cell  $z \in \bar{\mathcal{G}}_l \setminus \mathcal{G}_l^1$  changes according to the  $i$ th reaction.

$$q_l(\mathbf{u}_l, \mathbf{u}_l + \frac{1}{n}\sum_{j=n_d+1}^{n_s}\nu_{ij}\bar{\chi}_{j,z}) = n K_{l,i}(\mathbf{u}_l(z) \wedge 1) \quad \text{if } \mathbf{u}_l + \frac{1}{n}\sum_{j=n_d+1}^{n_s}\nu_{ij}\bar{\chi}_{j,z} \in S_l. \quad (3.30)$$

The transition intensities again characterise a Markov jump process with generator

$$L_l g(\mathbf{u}_l) = \sum_{\tilde{\mathbf{u}}_l \neq \mathbf{u}_l} q_l(\mathbf{u}_l, \tilde{\mathbf{u}}_l) (g(\tilde{\mathbf{u}}_l) - g(\mathbf{u}_l)), \quad g \in \hat{C}(S_l), \quad (3.31)$$

according to Theorem 3.1 in Chapter 8 of (Ethier & Kurtz, 1986).

## Law of large numbers

We first show convergence of the semi-discrete approximation, which ensures at the same time the existence of a solution for the weak PDE problem (3.23). To this end, we first discuss the approximation properties of the spaces  $\mathcal{L}^2(\bar{\mathcal{G}}_h)$  and  $\mathcal{H}^1(\bar{\mathcal{G}}_h)$ . Let again  $I = (0, h_0] \subset \mathbb{R}^+$ . For the proofs of the following theorems we refer to similar results for symmetric difference operators that can be found in Raviart (1967) and Temam (1973).

We extend a function  $u \in L^2(G)$  to a function  $\tilde{u} \in L^2(\mathbb{R}^m)$  by setting  $\tilde{u} = u$  on  $G$  and  $\tilde{u} \equiv 0$  on  $\mathbb{R}^m \setminus G$ . Let the lattice function  $\bar{u}_h$  be defined as in (2.65). The restriction operator  $\bar{Q}_h : L^2(G) \rightarrow \mathcal{L}^2(\bar{\mathcal{G}}_h)$  is then defined as

$$\bar{Q}_h u(z) = \begin{cases} \bar{u}_h(z) & \text{if } z \in \bar{\mathcal{G}}_h \\ 0 & \text{otherwise.} \end{cases} \quad (3.32)$$

If  $u$  is in  $H^1(G)$  we can extend it to a function  $\tilde{u}$  in  $H^1(\mathbb{R}^m)$  such that  $u$  is left unchanged on  $G$  by applying a bounded linear prolongation operator (see, e.g., Temam (1973)). We define again  $\bar{u}$  as in (2.65). The restriction operator  $\bar{R}_h : H^1(G) \rightarrow \mathcal{H}^1(\bar{\mathcal{G}}_h)$  is then given as above by setting

$$\bar{R}_h u(z) = \begin{cases} \bar{u}_h(z) & \text{if } z \in \bar{\mathcal{G}}_h \\ 0 & \text{otherwise.} \end{cases} \quad (3.33)$$

The restriction operators have the following properties.

**Lemma 3.1.8.** *The restriction operators  $\bar{Q}_h$  and  $\bar{R}_h$  are bounded linear operators, i.e., they are elements of  $\mathcal{L}(L^2(G), \mathcal{L}^2(\bar{\mathcal{G}}_h))$  and  $\mathcal{L}(H^1(G), \mathcal{H}^1(\bar{\mathcal{G}}_h))$ , respectively, and the families  $(\bar{Q}_h)_{h \in I}$  and  $(\bar{R}_h)_{h \in I}$  are stable.*

**Proof.** Cf. Raviart (1967), Chapter 0, or Temam (1973), Chapter 9. □

We define the extension operators  $\bar{P}_h : \mathcal{L}^2(\bar{\mathcal{G}}_h) \rightarrow L^2(G)$  by

$$\bar{P}_h u_h = u_h|_G, \quad u_h \in \mathcal{L}^2(\bar{\mathcal{G}}_h). \quad (3.34)$$

**Theorem 3.1.9.** *The internal approximation of  $L^2(G)$  by  $(\mathcal{L}^2(\bar{\mathcal{G}}_h), \bar{Q}_h, \bar{P}_h)_{h \in I}$  is stable and convergent.*

**Proof.** Cf. Raviart (1967), Chapter 0. □

We turn now to an external approximation of the Sobolev space  $H^1(G)$  in terms of the spaces  $\mathcal{H}^1(\bar{\mathcal{G}}_h)$ . To this end, we set again  $X = (L^2(G))^{m+1}$ . The isometric embedding  $\bar{J} : H^1(G) \rightarrow X$  is given by

$$u \mapsto \bar{J}u = (u, \partial_{x_1} u, \dots, \partial_{x_m} u). \quad (3.35)$$

The extension operators  $\bar{J}_h : \mathcal{H}^1(\bar{\mathcal{G}}_h) \rightarrow X$  are defined as

$$u_h \mapsto \bar{J}_h u_h = (u_h|_G, \partial_1^+ u_h|_G, \dots, \partial_m^+ u_h|_G). \quad (3.36)$$

**Theorem 3.1.10.** *The external approximation of  $H^1(G)$  by  $X$ ,  $\bar{J}$  and  $(\mathcal{H}^1(\bar{\mathcal{G}}_h), \bar{R}_h, \bar{J}_h)_{h \in I}$  is stable and convergent.*

**Proof.** Cf. Raviart (1967), Chapter 0, or Temam (1973), Chapter 9.  $\square$

Furthermore we have an analogue of Theorem 2.3.3 (cf. Raviart (1967), Chapter 0) and the following analogue of the compactness theorem 2.3.4.

**Theorem 3.1.11.** *Let  $(u_h)$ ,  $h \searrow 0$ , be a sequence of functions in  $L^2(0, T; \mathcal{H}^1(\bar{\mathcal{G}}_h))$  that have time derivatives  $u'_h$  in  $L^2(0, T; (\mathcal{H}^1(\bar{\mathcal{G}}_h))^*)$  and satisfy*

$$\sup_h \int_0^T \left( \|u_h(t)\|_{\mathcal{H}^1}^2 + \|u'_h(t)\|_{(\mathcal{H}^1)^*}^2 \right) dt < \infty.$$

*Then there is a subsequence  $(u_{h'})$ ,  $h' \searrow 0$ , that converges strongly in  $L^2(0, T; L^2(G))$ .*

**Proof.** The proof can again be carried out by imitating the proof of Theorem 12.1 in Chapter 1 of Lions (1969).  $\square$

We are now going to prove the convergence of the solutions of the approximating problem (3.25) to the solution of the weak PDE problem (3.23).

**Theorem 3.1.12.** *Suppose that  $\mathbf{u}_0 = (u_{1,0}, \dots, u_{n_s,0})$  is as in the weak formulation (3.23). Let  $(u_{h,j})$ ,  $h \searrow 0$ , be a sequence of solutions of the  $j$ th equation of the approximating problem (3.25) to the initial values  $\bar{Q}_h u_{j,0}$  ( $j = 1, \dots, n_d$ ), respectively  $\bar{R}_h u_{j,0}$  ( $j = n_d + 1, \dots, n_s$ ), and assume that  $D_{h,j} \rightarrow D_j$ ,  $j = 1, \dots, n_d$ . Then  $J_h u_{h,j} \rightharpoonup J u_j$  in  $L^2(0, T; X)$  ( $j = 1, \dots, n_d$ ), respectively  $\bar{J}_h u_{h,j} \rightharpoonup J u_j$  in  $L^2(0, T; X)$  ( $j = n_d + 1, \dots, n_s$ ), where  $\mathbf{u} = (u_1, \dots, u_{n_s})$  is the solution vector of the weak PDE problem (3.23) to the initial value  $\mathbf{u}_0$ . Moreover,  $\mathbf{u}_h = (u_{h,1}, \dots, u_{h,n_s})$  converges strongly to  $\mathbf{u}$  in  $L^2(0, T; L^2(G))$ .*

**Proof.** 1. From the definition of the restriction operators  $\bar{Q}_h$  and  $\bar{R}_h$  it follows that

$$0 \leq u_{h,j}(z, 0) \leq 1 \quad (3.37)$$

for all  $z \in \bar{\mathcal{G}}_h$  and  $j = 1, \dots, n_s$ . The local solutions  $\mathbf{u}_h(z, \cdot)$  of the approximating problem (3.25) obtained from the Picard-Lindelöf theorem are confined to the cube  $[0, 1]^{n_s}$  because the vector field  $\mathbf{f}$  points inwards on its boundary and the discrete Laplacian  $\nabla^- \cdot \nabla^+$  dampens maxima and minima. Suppose that one of the functions  $u_{h,j}(z, \cdot)$  would leave the interval  $[0, 1]$  for the first time at  $t_0$ , say,  $u_{h,j_0}(z_0, t_0) = 1$ ,  $u_{h,j_0}(z_0, t) > 1$  for  $t \in (t_0, t_0 + \varepsilon)$ , where  $\varepsilon > 0$ , and  $0 \leq u_{h,j}(z, t_0) \leq 1$  for  $z \neq z_0$ ,  $j \neq j_0$ . Then  $u'_{h,j_0}(z_0, t_0) \leq 0$  because of hypothesis

(I) and the definition of the discrete Laplacian, a contradiction. This ensures the existence of a solution of the approximating problem on the whole interval  $[0, T]$  and yields the estimate

$$0 \leq u_{h,j} \leq 1 \quad \text{on } \bar{\mathcal{G}}_h \times [0, T] \quad (3.38)$$

for  $j = 1, \dots, n_s$ . In particular,

$$\sup_h \max_{0 \leq t \leq T} \|u_{h,j}\|_{\mathcal{L}^2(\bar{\mathcal{G}}_h)} < \infty. \quad (3.39)$$

Note that the restriction of  $\mathbf{f}$  to the cube  $[0, 1]^{n_s}$  is Lipschitz continuous. By inserting  $u_{h,j}(t)$  for  $v_{h,j}$  in the discrete weak formulation and summing up the equations, we get

$$\begin{aligned} \|\mathbf{u}_h(t)\|_{\mathcal{L}^2(\bar{\mathcal{G}}_h)}^2 + 2 \sum_{j=1}^{n_d} \sum_{k=1}^m \int_0^t D_{h,j} \|\partial_k^+ u_{h,j}(s)\|_{\mathcal{L}^2(\mathcal{G}_h)}^2 ds \\ \leq \|\mathbf{u}_h(0)\|_{\mathcal{L}^2(\bar{\mathcal{G}}_h)}^2 + C \int_0^t \left(1 + \|\mathbf{u}_h(s)\|_{\mathcal{L}^2(\bar{\mathcal{G}}_h)}^2\right) ds. \end{aligned} \quad (3.40)$$

By making use of the estimate (3.38) we find (due to the discrete Poincaré inequality) that

$$\sup_h \|u_{h,j}\|_{L^2(0,T;\mathcal{H}_0^1(\mathcal{G}_h))} < \infty, \quad j = 1, \dots, n_d. \quad (3.41)$$

The delicate point in the proof is that here the usual procedure doesn't provide us with an estimate for  $\|u_{h,j}\|_{L^2(0,T;\mathcal{H}^1(\bar{\mathcal{G}}_h))}$  for  $j = n_d + 1, \dots, n_s$ . We now exploit that, for  $j = n_d + 1, \dots, n_s$ , we have  $u_{j,0} \in H^1(G)$  by hypothesis. Note that, for  $j = n_d + 1, \dots, n_s$ ,  $z \in \bar{\mathcal{G}}_h^1$ ,  $k = 1, \dots, m$ ,

$$|u_{h,j}(z + h e_k, t) - u_{h,j}(z, t)| \leq \exp(C_L t) |u_{h,j}(z + h e_k, 0) - u_{h,j}(z, 0)|, \quad (3.42)$$

where  $C_L$  denotes the Lipschitz constant of  $\mathbf{f}$  restricted to  $[0, 1]^{n_s}$ . Hence,

$$\begin{aligned} h^m \sum_{z \in \bar{\mathcal{G}}_h^1} \frac{1}{h^2} |u_{h,j}(z + h e_k, t) - u_{h,j}(z, t)|^2 \\ \leq h^m \sum_{z \in \bar{\mathcal{G}}_h^1} \frac{1}{h^2} \exp(2 C_L t) |u_{h,j}(z + h e_k, 0) - u_{h,j}(z, 0)|^2 \\ \leq \exp(2 C_L t) \|\bar{R}_h\|_{\mathcal{L}(H^1, \mathcal{H}^1)}^2 \|u_{j,0}\|_{H^1}^2. \end{aligned} \quad (3.43)$$

Owing to the stability of the family  $(\bar{R}_h)$  and the estimate (3.39) we obtain

$$\sup_h \max_{0 \leq t \leq T} \|u_{h,j}\|_{\mathcal{H}^1(\bar{\mathcal{G}}_h)} < \infty, \quad j = n_d + 1, \dots, n_s. \quad (3.44)$$

The additional estimates

$$\sup_h \|u'_{h,j}\|_{L^2(0,T;\mathcal{H}^{-1})} < \infty, \quad j = 1, \dots, n_d, \quad (3.45)$$

$$\sup_h \|u'_{h,j}\|_{L^2(0,T;(\mathcal{H}^1)^*)} < \infty, \quad j = n_d + 1, \dots, n_s, \quad (3.46)$$

are easily obtained from the discrete weak formulation.

2. In view of the a-priori estimates, we can extract a subsequence (still denoted by  $(\mathbf{u}_h)$ ) so that  $J_h u_{h,j} \rightharpoonup J u_j$  in  $L^2(0, T; X)$  ( $j = 1, \dots, n_d$ ) for a  $u_j \in L^2(0, T; H_0^1(G))$ . Furthermore, for  $j = n_d + 1, \dots, n_s$ ,  $\bar{J} u_{h,j} \rightharpoonup u_j$  in  $L^2(0, T; X)$ , where  $u_j$  is an element of  $L^2(0, T; H^1(G))$ . In addition, we may assume that  $\mathbf{u}_h$  converges strongly to  $\mathbf{u}$  in  $L^2(0, T; \mathbf{L}^2(G))$  due to the compactness theorems 2.3.4 and 3.1.11.

3. The identification of  $\mathbf{u}$  as solution of the weak PDE problem (3.23) follows by similar arguments as in the proof of Theorem 2.3.2.

4. The uniqueness of the limit yields convergence of the whole sequence.  $\square$

We denote from now on the stochastic particle density by  $\mathbf{u}_l$  and the solutions of the approximating problem (3.25) by  $\mathbf{v}_l$ . Let (S1)–(S3) be the scaling relations from the previous section.

**Theorem 3.1.13** (Law of large numbers). *Let  $\mathbf{u}$  be the solution of the weak PDE problem (3.23) to the initial value  $\mathbf{u}_0$ . Assume that (S1)–(S3) hold and that  $\mathbf{u}_l(0)$  converges strongly to  $\mathbf{u}_0$  in  $\mathbf{L}^2$ . Then*

$$E \left[ \|\mathbf{u}_l - \mathbf{u}\|_{L^2(0, T; \mathbf{L}^2)}^2 \right] = E \left[ \|\mathbf{u}_l - \mathbf{u}\|_{(L^2(Q_T))^{n_s}}^2 \right] \rightarrow 0.$$

The law of large numbers follows, as usual, immediately from the next auxiliary theorem.

**Theorem 3.1.14.** *Assume (S1)–(S3), and denote by  $\mathbf{v}_l$  the solution vector of the approximating problem (3.25) with  $D_{l,j} = \frac{d_j}{2m} l^2$  to the initial values  $v_{l,j,0} = \bar{Q}_l u_{j,0}$  ( $j = 1, \dots, n_d$ ) and  $\bar{R}_l u_{j,0}$  ( $j = n_d + 1, \dots, n_s$ ), respectively. Then*

$$\sup_{t \leq T} E \left[ \|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2(\bar{\mathcal{G}}_l)}^2 \right] \rightarrow 0.$$

The proof of the above result rests again on a lemma that identifies a local martingale related to the process  $\|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2}$ ,  $t \leq T$ . The stopping times  $\tau_p$ ,  $p \in \mathbb{N}$ , are defined as in (2.111).

**Lemma 3.1.15.** *Let  $(M_l(t))_{t \leq T}$  be the process given by*

$$\begin{aligned} M_l(t) &= \|\mathbf{u}_l(t) - \mathbf{v}_l(t)\|_{\mathcal{L}^2(\bar{\mathcal{G}}_l)}^2 - \|\mathbf{u}_l(0) - \mathbf{v}_l(0)\|_{\mathcal{L}^2(\bar{\mathcal{G}}_l)}^2 \\ &\quad + 2 \sum_{j=1}^{n_d} \sum_{k=1}^m \int_0^t D_{l,j} (\partial_k^+ u_{h,j} - \partial_k^+ v_{h,j}, \partial_k^+ u_{h,j} - \partial_k^+ v_{h,j})_{\mathcal{L}^2(\mathcal{G}_h)} ds \\ &\quad - 2 \int_0^t (\mathbf{f}_l(\mathbf{u}_l(s) \wedge 1) - \mathbf{f}(\mathbf{v}_l(s)), \mathbf{u}_l(s) - \mathbf{v}_l(s))_{\mathcal{L}^2(\bar{\mathcal{G}}_l)} ds - R_l(t), \end{aligned} \tag{3.47}$$

where

$$\begin{aligned} R_l(t) &= \frac{1}{n} \sum_{j=1}^{n_d} 2d_j \int_0^t (u_{l,j}(s), 1)_{\mathcal{L}^2(\bar{\mathcal{G}}_l)} ds \\ &\quad + \frac{1}{n} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} \nu_{ij}^2 \int_0^t (K_{l,i}(\mathbf{u}_l(s) \wedge 1), 1)_{\mathcal{L}^2(\bar{\mathcal{G}}_l)} ds. \end{aligned} \tag{3.48}$$

Then the stopped process  $(M_l(t \wedge \tau_p))_{t \leq T}$  is a martingale for each  $p \in \mathbb{N}$ .

**Proof.** The proof of the lemma above works in the same way as for the general linear case and the case of Lipschitz conditions. The growth of the process  $\|\mathbf{u}_l(t)\|_{\mathcal{L}^2(\bar{\mathcal{G}}_h)}$ ,  $t \geq 0$ , can again be estimated by an appropriate Yule process.  $\square$

**Proof of Theorem 3.1.14.** By stopping the local martingale (3.47), taking expectations, and observing that the term involving the discrete derivatives is still monotone, we get

$$\begin{aligned}
E \left[ \|\mathbf{u}_l(t \wedge \tau_p) - \mathbf{v}_l(t \wedge \tau_p)\|_{\mathcal{L}^2(\bar{\mathcal{G}}_l)}^2 \right] &\leq \|\mathbf{u}_l(0) - \mathbf{v}_l(0)\|_{\mathcal{L}^2(\bar{\mathcal{G}}_l)}^2 \\
&+ E \int_0^{t \wedge \tau_p} \left| (\mathbf{f}_l(\mathbf{u}_l(s) \wedge 1) - \mathbf{f}(\mathbf{v}_l(s)), \mathbf{u}_l(s) - \mathbf{v}_l(s))_{\mathcal{L}^2(\bar{\mathcal{G}}_l)} \right| ds \\
&+ \frac{1}{n} \sum_{j=1}^{n_s} 2d_j E \int_0^{t \wedge \tau_p} (u_{l,j}(s), 1)_{\mathcal{L}^2(\bar{\mathcal{G}}_l)} ds \\
&+ \frac{1}{n} \sum_{j=1}^{n_s} \sum_{i=1}^{n_r} \nu_{ij}^2 E \int_0^{t \wedge \tau_p} (K_{l,i}(\mathbf{u}_l(s) \wedge 1), 1)_{\mathcal{L}^2(\bar{\mathcal{G}}_l)} ds.
\end{aligned} \tag{3.49}$$

We can derive again the estimate (2.125) (with  $\mathcal{L}^2(\mathcal{G}_l)$  replaced by  $\mathcal{L}^2(\bar{\mathcal{G}}_l)$ ) by making use of the fact that the reaction rates  $K_i$  are locally Lipschitz continuous in the second and fourth term on the right. The second term, for instance, can be estimated as follows. Let  $C_L$  be the Lipschitz constant of the vector field  $\mathbf{f}$  restricted to  $[0, 1]^{n_s}$ . The triangle inequality yields

$$\begin{aligned}
&\left| (\mathbf{f}_l(\mathbf{u}_l(s) \wedge 1) - \mathbf{f}(\mathbf{v}_l(s)), \mathbf{u}_l(s) - \mathbf{v}_l(s))_{\mathcal{L}^2(\bar{\mathcal{G}}_l)} \right| \\
&\leq \left| (\mathbf{f}_l(\mathbf{u}_l(s) \wedge 1) - \mathbf{f}(\mathbf{u}_l(s) \wedge 1), \mathbf{u}_l(s) - \mathbf{v}_l(s))_{\mathcal{L}^2(\bar{\mathcal{G}}_l)} \right| \\
&\quad + \left| (\mathbf{f}(\mathbf{u}_l(s) \wedge 1) - \mathbf{f}(\mathbf{v}_l(s)), \mathbf{u}_l(s) - \mathbf{v}_l(s))_{\mathcal{L}^2(\bar{\mathcal{G}}_l)} \right|.
\end{aligned} \tag{3.50}$$

Here the first term on the right is not problematic because  $\mathbf{f}_l$  converges uniformly to  $\mathbf{f}$ . (Recall that  $\mathbf{f}_l$  is the vector of reaction functions obtained by using the rates  $K_{l,i}$  instead of the  $K_i$ .) As for the second term, observe that

$$\begin{aligned}
&\left| (\mathbf{f}(\mathbf{u}_l(s) \wedge 1) - \mathbf{f}(\mathbf{v}_l(s)), \mathbf{u}_l(s) - \mathbf{v}_l(s))_{\mathcal{L}^2(\bar{\mathcal{G}}_l)} \right| \\
&= \left| l^m \sum_{j=1}^{n_s} \sum_{z \in \bar{\mathcal{G}}_l} (f_j(\mathbf{u}_l(z, s) \wedge 1) - f_j(\mathbf{v}_l(z, s))) (u_{l,j}(z, s) - v_{l,j}(z, s)) \right| \\
&\leq C_L l^m \sum_{j=1}^{n_s} \sum_{z \in \bar{\mathcal{G}}_l} |\mathbf{u}_l(z, s) \wedge 1 - \mathbf{v}_l(z, s)| |u_{l,j}(z, s) - v_{l,j}(z, s)| \\
&\leq C_L l^m \sum_{j=1}^{n_s} \sum_{z \in \bar{\mathcal{G}}_l} |\mathbf{u}_l(z, s) - \mathbf{v}_l(z, s)| |u_{l,j}(z, s) - v_{l,j}(z, s)| \\
&\leq C l^m \sum_{z \in \bar{\mathcal{G}}_l} |\mathbf{u}_l(z, s) - \mathbf{v}_l(z, s)|^2 = C \|\mathbf{u}_l(s) - \mathbf{v}_l(s)\|_{\mathcal{L}^2(\bar{\mathcal{G}}_l)}^2.
\end{aligned} \tag{3.51}$$

With the aid of the estimate (2.125) the proof can be finished as before.  $\square$

## 3.2 Nonlinear diffusion

In this section we study in more detail which kind of limit equation can arise when the intensities for diffusive jumps depend on the concentration or the concentration gradient. For the sake of simplicity, we restrict the discussion to single-species models without chemical reactions. Moreover, in order to save notation, the proof of the law of large numbers is only carried out for one space dimension, i.e., for a chemical reactor that is represented by an interval  $(0, L)$ .

### 3.2.1 Crowding effects

We are going to study what happens when the intensity for a diffusive jump increases with the concentration in the cells, i.e., the intensity for a jump to a neighboring cell is  $d = d(u_l)$ , where  $d$  is a monotonously increasing function. This models repulsive interactions between the particles.

The macroscopic PDE and its weak formulation

For the following discussion  $G$  may still be a bounded domain in  $\mathbb{R}^m$  with Lipschitz boundary. The PDE that will be approached in the limit of large particle numbers in this section (which is now perhaps less obvious) is

$$\begin{cases} \partial_t u - \Delta(D(u) u) = 0 & \text{on } Q_T \\ u = 0 & \text{on } \partial G \times [0, T] \\ u(\cdot, 0) = u_0(\cdot) & \text{on } G, \end{cases} \quad (3.52)$$

where the function  $D : \mathbb{R} \rightarrow \mathbb{R}_0^+$  is assumed to satisfy certain conditions that will be specified below. As before, we look for an appropriate weak formulation. Let us assume for the following calculations that  $D$  is sufficiently smooth and that we have found a smooth solution  $u$  of (3.52). We multiply the equation with a test function  $v \in C_0^\infty(G)$  and integrate over  $G$ , which yields

$$\frac{d}{dt} \int_G u v \, dx - \int_G \Delta(D(u) u) v = 0. \quad (3.53)$$

By partial integration we get

$$\frac{d}{dt} \int_G u v \, dx + \int_G \nabla(D(u) u) \cdot \nabla v \, dx - \int_{\partial G} v \nabla(D(u) u) \cdot \vec{n} \, dS = 0, \quad (3.54)$$

where  $\vec{n}$  denotes the outer normal. The integral over the boundary vanishes because  $v \in C_0^\infty(G)$ . We notice, however, that the term

$$\int_G \nabla(D(u) u) \cdot \nabla v \, dx$$

is in general not monotone (not even for monotone  $D$ ), and therefore Eq. (3.54) does not allow for an immediate interpretation as monotone evolution equation. Therefore we integrate by parts one more time, which yields

$$\frac{d}{dt} \int_G u v - \int_G D(u) u \Delta v \, dx + \int_{\partial G} D(u) u \nabla v \cdot \vec{n} \, dS = 0. \quad (3.55)$$

The boundary integral vanishes again. By setting  $-\Delta v = w$  we get

$$\frac{d}{dt} \int_G u (-\Delta)^{-1} w \, dx + \int_G D(u) u w \, dx = 0. \quad (3.56)$$

The above equation can be interpreted as monotone evolution equation (cf. Lions (1969), Section 3 of Chapter 2). To this end, we endow the Hilbert space  $H_0^1(G)$  with the equivalent scalar product

$$(u, v)_{H_0^1} = (\nabla u, \nabla v)_{L^2}, \quad u, v \in H_0^1. \quad (3.57)$$

We interpret  $-\Delta$  as operator from  $H_0^1$  to  $H^{-1}$  in the usual way:

$$\langle -\Delta u, v \rangle_{H_0^1} = (\nabla u, \nabla v)_{L^2}, \quad u, v \in H_0^1. \quad (3.58)$$

Apparently, the operator  $-\Delta$  defined in this way is identical to the Riesz isomorphism between the Hilbert space  $H_0^1$  and its dual  $H^{-1}$ , since the scalar product in  $H_0^1$  is given by (3.57). We define on  $H^{-1}$  the scalar product

$$(u, v)_{H^{-1}} = \langle u, -\Delta^{-1} v \rangle_{H_0^1}, \quad u, v \in H^{-1}, \quad (3.59)$$

and we denote the corresponding norm by  $||| \cdot |||_{H^{-1}}$ . The norm  $||| \cdot |||_{H^{-1}}$  is in fact equal to the standard norm in  $H^{-1}$  which we denote by  $\| \cdot \|_{H^{-1}}$ . To see this, set  $\tilde{w} = (-\Delta)^{-1} w$  for  $w \in H^{-1}$ , and observe that

$$||| w |||_{H^{-1}}^2 = (w, w)_{H^{-1}} = \langle w, \tilde{w} \rangle_{H_0^1} = \langle -\Delta \tilde{w}, \tilde{w} \rangle_{H_0^1} = (\tilde{w}, \tilde{w})_{H_0^1} = \| w \|_{H^{-1}}^2. \quad (3.60)$$

The spaces  $L^2(G)$ ,  $H^{-1}(G)$  (endowed with the scalar product (3.59)), and  $(L^2(G))^*$  form a Gelfand triple:

$$L^2(G) \hookrightarrow H^{-1}(G) \cong (H^{-1}(G))^* \hookrightarrow (L^2(G))^*. \quad (3.61)$$

We have to take care that here the second embedding  $j_2 : H^{-1} \hookrightarrow (L^2)^*$  is given by

$$\langle j_2(u), v \rangle_{L^2} = (u, v)_{H^{-1}} = \langle u, (-\Delta)^{-1} v \rangle_{H_0^1}, \quad u \in H^{-1}, \quad v \in L^2, \quad (3.62)$$

while the first embedding is still the usual one.

We make the following hypotheses for the function  $D : \mathbb{R} \rightarrow \mathbb{R}_0^+$ .

$$D \text{ is continuous and monotonously increasing on } \mathbb{R}_0^+. \quad (D1)$$

$$D(p) = D(-p) \text{ for all } p \in \mathbb{R}. \quad (D2)$$

$$\text{There are constants } C, \alpha > 0 \text{ such that } D(p) \leq C \text{ and } D(p) p^2 \geq \alpha p^2 \text{ for all } p \in \mathbb{R}. \quad (D3)$$

**Lemma 3.2.1.** *Let the mapping  $a : L^2 \times L^2 \rightarrow \mathbb{R}$  be given by*

$$a(u, v) = \int_G D(u) u v \, dx, \quad u, v \in L^2, \quad (3.63)$$

*and assume that  $D$  satisfies (D1)–(D3). Then the mapping  $a(\cdot, \cdot)$  induces a (generally nonlinear) operator  $A : L^2 \rightarrow (L^2)^*$  by*

$$\langle A(u), v \rangle_{L^2} = a(u, v), \quad u, v \in L^2, \quad (3.64)$$

*which is bounded, coercive, hemicontinuous and monotone.*

(For the definition of hemicontinuity see, e.g., Zeidler (1990c).)

**Proof.** For  $u, v \in L^2$  we have the estimate

$$|\langle A(u), v \rangle_{L^2}| = \left| \int_G D(u) u v \, dx \right| \leq C \|u\|_{L^2} \|v\|_{L^2}, \quad (3.65)$$

and thus  $\|A(u)\|_{(L^2)^*} \leq C \|u\|_{L^2}$ . Furthermore,

$$\langle A(u), u \rangle_{L^2} = \int_G D(u) u^2 \, dx \geq \alpha \|u\|_{L^2}^2. \quad (3.66)$$

Hence,  $A$  is bounded and coercive. As for the hemicontinuity, note that for  $u, v, w \in L^2$ ,

$$|\langle A(u + \lambda v) - A(u), w \rangle_{L^2}| = \left| \int_G (D(u + \lambda v)(u + \lambda v) - D(u)u) w \, dx \right| \quad (3.67)$$

converges to zero for  $\lambda \rightarrow 0$  by the dominated convergence theorem. Finally, the monotonicity condition

$$\langle A(u) - A(v), u - v \rangle_{L^2} \geq 0 \quad (3.68)$$

follows from the properties (D1) and (D2) of the function  $D$ .  $\square$

In the weak interpretation of the PDE (3.52) a function  $u \in H^1(0, T; L^2, H^{-1})$  is sought such that

$$\frac{d}{dt}(u, v)_{H^{-1}} + a(u, v) = 0 \quad (3.69a)$$

for all  $v \in L^2$  and a.e.  $t \in [0, T]$ , and

$$u(0) = u_0 \in H^{-1}. \quad (3.69b)$$

It has a unique solution according to a general theorem on monotone first-order evolution equations (see, e.g., Theorem 30.A in Zeidler (1990c) or Theorem 1.2 in Chapter 2 of Lions (1969)).

### The approximating problem

We now specialise to one space dimension, i.e.,  $G = (0, L)$ , and  $\Delta_h = \partial^- \partial^+$  is the discrete Laplacian. The discrete analogue of the PDE (3.52) on the interior lattice points  $\mathcal{G}_h$  is given by

$$\begin{cases} u'_h - \Delta_h(D_h(u_h) u_h) = 0 & \text{on } \mathcal{G}_h^1 \times (0, T) \\ u_h = 0 & \text{on } (\mathcal{G}_h \setminus \mathcal{G}_h^1) \times [0, T] \\ u_h(\cdot, 0) = u_{h,0} & \text{on } \mathcal{G}_h^1. \end{cases} \quad (3.70)$$

We assume that the functions  $D_h$  satisfy the same hypotheses (D1)–(D3) as  $D$  and that  $\sup_{\mathbb{R}} |D_h - D| \rightarrow 0$  for  $h \rightarrow 0$ . The discretised PDE (3.70) is in fact a system of ODEs with continuous right-hand side which has a local solution according to the Peano theorem. Existence of a solution on the whole interval  $[0, T]$  follows from the derivation of the a-priori estimate (3.82) below.

We endow the discrete Sobolev space  $\mathcal{H}_0^1(\mathcal{G}_h)$  with the scalar product

$$(u_h, v_h)_{\mathcal{H}_0^1} = (\partial^+ u_h, \partial^+ v_h)_{\mathcal{L}^2}, \quad u_h, v_h \in \mathcal{H}_0^1, \quad (3.71)$$

which induces a norm equivalent to the original one due to the discrete Poincaré inequality. In analogy to the treatment of the PDE (3.52) we regard  $-\Delta_h$  as operator from  $\mathcal{H}_0^1$  to  $\mathcal{H}^{-1}$  given by

$$\langle -\Delta_h u_h, v_h \rangle_{\mathcal{H}_0^1} = (\partial^+ u_h, \partial^+ v_h)_{\mathcal{L}^2}, \quad u_h, v_h \in \mathcal{H}_0^1, \quad (3.72)$$

and  $\mathcal{H}^{-1}$  is equipped with the scalar product

$$(u_h, v_h)_{\mathcal{H}^{-1}} = \langle u_h, -\Delta_h^{-1} v_h \rangle_{\mathcal{H}_0^1}, \quad u_h, v_h \in \mathcal{H}^{-1}. \quad (3.73)$$

The corresponding norm is denoted by  $\|\cdot\|_{\mathcal{H}^{-1}}$ . It is equal to the standard norm which we denote by  $\|\cdot\|_{\mathcal{H}^{-1}}$ . The discrete spaces  $\mathcal{L}^2(\mathcal{G}_h)$ ,  $\mathcal{H}^{-1}(\mathcal{G}_h)$  and  $(\mathcal{L}^2(\mathcal{G}_h))^*$  also form a Gelfand triple:

$$\mathcal{L}^2(\mathcal{G}_h) \hookrightarrow \mathcal{H}^{-1}(\mathcal{G}_h) \cong (\mathcal{H}^{-1}(\mathcal{G}_h))^* \hookrightarrow (\mathcal{L}^2(\mathcal{G}_h))^*. \quad (3.74)$$

The solution of (3.70) can be regarded as a function in  $C^1([0, T], \mathcal{L}^2)$  that satisfies

$$\frac{d}{dt} (u_h, v_h)_{\mathcal{H}^{-1}} + a_h(u_h, v_h) = 0 \quad (3.75)$$

for all  $v_h \in \mathcal{L}^2$  and  $t \in (0, T)$ , where the mapping  $a_h : \mathcal{L}^2 \times \mathcal{L}^2 \rightarrow \mathbb{R}$  is defined as

$$a_h(u_h, v_h) = (D_h(u_h) u_h, v_h)_{\mathcal{L}^2}, \quad u_h, v_h \in \mathcal{L}^2. \quad (3.76)$$

In analogy to  $a(\cdot, \cdot)$ , the mapping  $a_h(\cdot, \cdot)$  induces a bounded and monotone operator  $A_h : \mathcal{L}^2 \rightarrow (\mathcal{L}^2)^*$  by

$$\langle A_h(u_h), v_h \rangle_{\mathcal{L}^2} = a_h(u_h, v_h), \quad u_h, v_h \in \mathcal{L}^2. \quad (3.77)$$

### The mesoscopic stochastic particle model

We work in the same setting as for the linear example model from Section 2.1.2, i.e., the state space  $S_l$  of the particle density process  $(u_l(t))_{t \geq 0}$  is the subset of functions in  $\mathcal{L}^2(\mathcal{G}_l)$  that take values in  $\frac{1}{n}\mathbb{N}_0$ . Moreover, we assume that  $u_l(0)$  is non-random and that  $u_l$  satisfies Dirichlet boundary conditions, i.e.,  $u_l(z, \cdot) \equiv 0$  for  $z \in \mathcal{G}_l \setminus \mathcal{G}_l^1$ . The possible transitions are the following.

- A particle may leave cell  $z \in \mathcal{G}_l^1$  and jump to  $z \pm l$ .

$$\begin{aligned} q_l(u_l, u_l - \frac{1}{n}\chi_z + \frac{1}{n}\chi_{(z-l)}) &= n \frac{1}{2} d(u_l(z)) u_l(z), \\ q_l(u_l, u_l - \frac{1}{n}\chi_z + \frac{1}{n}\chi_{(z+l)}) &= n \frac{1}{2} d(u_l(z)) u_l(z), \end{aligned} \quad (3.78)$$

where  $l^2 d : \mathbb{R} \rightarrow \mathbb{R}_0^+$  is supposed to satisfy conditions (D1)–(D3).

The existence of a particle density process  $(u_l(t))_{t \geq 0}$  with generator

$$L_l g(u_l) = \sum_{\tilde{u}_l \neq u_l} q_l(u_l, \tilde{u}_l) (g(\tilde{u}_l) - g(u_l)), \quad g \in \hat{C}(S_l), \quad (3.79)$$

follows again from Theorem 3.1 in Chapter 8 of Ethier & Kurtz (1986). Note that by construction

$$\sup_{z \in \mathcal{G}_l, t \geq 0} |u_l(z, t)| < \infty, \quad (3.80)$$

and

$$\|u_l(t)\|_{L^1(G)} \leq \|u_l(0)\|_{L^1(G)}, \quad t \geq 0, \quad (3.81)$$

almost surely.

### Law of large numbers

Comparing to the rest of our work on laws of large numbers we have introduced in this section a different functional setting to establish the existence of a unique solution of the macroscopic PDE. As a consequence we are (without further regularity considerations) only able to show weak convergence of the approximation in the space  $L^2(0, T; L^2)$ , which, in turn, results in a weaker law of large numbers.

**Theorem 3.2.2.** *Let  $u$  be the solution of the weak PDE problem (3.69) to the initial value  $u_0$ . Let  $(u_h)$ ,  $h \searrow 0$ , be a sequence of solutions of the approximating problem (3.75) to the initial values  $u_{h,0}$ , and assume that  $u_{h,0}$  converges strongly to  $u_0$  in  $L^2$ . Then  $u_h$  converges weakly to  $u$  in  $L^2(0, T; L^2)$ .*

**Proof.** 1. The a-priori estimates

$$\sup_h \max_{0 \leq t \leq T} \|u_h\|_{\mathcal{H}^{-1}} < \infty, \quad (3.82)$$

$$\sup_h \|u_h\|_{L^2(0,T;L^2)} < \infty \quad (3.83)$$

follow immediately from the discrete weak formulation by inserting  $u_h(t)$  for  $v_h$  and integrating over time. Therefore we can extract a subsequence such that  $u_h \rightharpoonup u$ , and  $D_h(u_h)u_h \rightharpoonup \xi$  in  $L^2(0,T;L^2)$ . Recall that the restriction operators  $R_h : H_0^1 \rightarrow \mathcal{H}_0^1$  constitute a stable family with  $\|R_h\|_{\mathcal{L}(H_0^1, \mathcal{H}_0^1)} \leq 1$  (even with changed scalar products, cf. the proof of Lemma 2.2.12), and consider the embeddings  $J_h^* : \mathcal{H}^{-1} \rightarrow H^{-1}$  given by

$$\langle J_h^*(u_h), v \rangle_{H_0^1} = \langle u_h, R_h v \rangle_{\mathcal{H}_0^1}, \quad u_h \in \mathcal{H}^{-1}, v \in H_0^1. \quad (3.84)$$

Observe that

$$\langle J_h^*(u_h), v \rangle_{H_0^1} \leq \|u_h\|_{\mathcal{H}^{-1}} \|R_h\| \|v\|_{H_0^1}, \quad (3.85)$$

and therefore  $\|J_h^*(u_h)\|_{H^{-1}} \leq \|u_h\|_{\mathcal{H}^{-1}}$ . Hence, by possibly passing to a further subsequence, we may assume that  $J_h^*(u_h(T)) \rightharpoonup \eta$  in  $H^{-1}$ .

2. We claim that the weak limit  $u$  satisfies the equation

$$u' + \xi = 0 \quad (3.86)$$

in  $H^1(0,T;L^2, H^{-1})$  with  $u(0) = u_0$  and  $u(T) = \eta$ . To see this, consider an arbitrary function  $\varphi \in C^1[0,T]$ , and set  $v_h = Q_h v$  for an arbitrary  $v \in L^2$ . From the discrete weak formulation we get by integration by parts that

$$\begin{aligned} & (u_h(T), \varphi(T)v_h)_{\mathcal{H}^{-1}} - (u_h(0), \varphi(0)v_h)_{\mathcal{H}^{-1}} \\ &= \int_0^T \left( \langle u_h'(t), \varphi(t)v_h \rangle_{\mathcal{L}^2} + \langle \varphi'(t)v_h, u_h(t) \rangle_{\mathcal{L}^2} \right) dt \\ &= \int_0^T \left( -a_h(u_h(t), v_h) \varphi(t) + (u_h(t), (-\Delta_h)^{-1}v_h)_{\mathcal{L}^2} \varphi'(t) \right) dt. \end{aligned} \quad (3.87)$$

Recall that by definition

$$(u_h(0), v_h)_{\mathcal{H}^{-1}} = \langle u_h(0), (-\Delta_h)^{-1}v_h \rangle_{\mathcal{H}_0^1} = (u_h(0), (-\Delta_h)^{-1}v_h)_{\mathcal{L}^2}, \quad (3.88)$$

and

$$(u_h(T), v_h)_{\mathcal{H}^{-1}} = \langle u_h(T), (-\Delta_h)^{-1}v_h \rangle_{\mathcal{H}_0^1}. \quad (3.89)$$

We set  $\tilde{v} = (-\Delta)^{-1}v \in H_0^1$ . Note that  $(-\Delta_h)^{-1}v_h \rightarrow \tilde{v}$  strongly in  $L^2$ , and

$$\|(-\Delta_h)^{-1}v_h - R_h \tilde{v}\|_{\mathcal{H}_0^1} \rightarrow 0 \quad (h \rightarrow 0). \quad (3.90)$$

This follows from Theorem 3.1 in Chapter 1 of Temam (2001) and the fact that  $v_h = Q_h v \rightarrow v$  strongly in  $L^2$ . Passing to the limit in all terms in Eq. (3.87) yields

$$\begin{aligned} & \langle \eta, \tilde{v} \rangle_{H_0^1} \varphi(T) - \langle u_0, \tilde{v} \rangle_{H_0^1} \varphi(0) \\ &= (\eta, v)_{H^{-1}} \varphi(T) - (u_0, v)_{H^{-1}} \varphi(0) \\ &= \int_0^T (u(t), v)_{H^{-1}} \varphi'(t) dt - \int_0^T (\xi, v)_{L^2} \varphi(t) dt. \end{aligned} \quad (3.91)$$

Hence, by taking  $\varphi \in C_0^\infty(0, T)$ , we conclude that  $u$  is indeed in  $H^1(0, T; L^2, H^{-1})$  with  $u' = -\xi$ . It follows from the integration by parts formula for functions in  $H^1(0, T; L^2, H^{-1})$  by choosing a  $\varphi$  with  $\varphi(T) = 1$  and  $\varphi(0) = 0$ , respectively  $\varphi(T) = 0$  and  $\varphi(0) = 1$ , that  $u(T) = \eta$  and  $u(0) = u_0$ .

3. It remains to show that  $\xi = A(u)$ . This equality would follow from the hemicontinuity of  $A$  by a well-known argument, often called the Minty trick or monotonicity trick (see, e.g., Lions (1969), Chapter 2, or Zeidler (1990c)), if we were able to prove that

$$\int_0^T \langle \xi(t) - A(v(t)), u(t) - v(t) \rangle_{L^2} dt \geq 0 \quad \text{for all } v \in L^2(0, T; L^2). \quad (3.92)$$

It is sufficient to show (3.92) for elements of  $L^2(0, T; L^2)$  of the form  $\varphi v$ , where  $v \in L^2$  and  $\varphi \in C^1[0, T]$ , because linear combinations of such elements are dense. We set  $v_h(t) = Q_h v \varphi(t)$  and consider the nonnegative expression

$$X_h = \int_0^T (D_h(u_h(t)) u_h(t) - D_h(v_h(t)) v_h(t), u_h(t) - v_h(t))_{\mathcal{L}^2} dt. \quad (3.93)$$

From the approximating problem it follows that

$$\int_0^T (D_h(u_h(t)) u_h(t), u_h(t))_{\mathcal{L}^2} dt = \frac{1}{2} \| \|u_h(0)\| \|_{\mathcal{H}^{-1}}^2 - \frac{1}{2} \| \|u_h(T)\| \|_{\mathcal{H}^{-1}}^2. \quad (3.94)$$

Hence,

$$\begin{aligned} X_h &= \frac{1}{2} \| \|u_h(0)\| \|_{\mathcal{H}^{-1}}^2 - \frac{1}{2} \| \|u_h(T)\| \|_{\mathcal{H}^{-1}}^2 - \int_0^T (D_h(u_h(t)) u_h(t), v_h(t))_{\mathcal{L}^2} dt \\ &\quad - \int_0^T (D_h(v_h(t)) v_h(t), u_h(t) - v_h(t))_{\mathcal{L}^2} dt. \end{aligned} \quad (3.95)$$

Note that  $(-\Delta_h)^{-1} u_h(0)$  converges strongly to  $(-\Delta)^{-1} u_0$  in  $L^2$ , since by hypothesis  $u_h(0)$  converges strongly to  $u_0$  in  $L^2$ . By taking limits superior we thus get

$$\begin{aligned} \overline{\lim} X_h &= \frac{1}{2} \| \|u_0\| \|_{H^{-1}}^2 - \frac{1}{2} \underline{\lim} \| \|u_h(T)\| \|_{\mathcal{H}^{-1}}^2 - \int_0^T (\xi(t), v(t))_{L^2} dt \\ &\quad - \int_0^T (D(v(t)) v(t), u(t) - v(t))_{L^2}. \end{aligned} \quad (3.96)$$

Observe that

$$\underline{\lim} \| \|u_h(T)\| \|_{\mathcal{H}^{-1}} \geq \underline{\lim} \| J_h^*(u_h(T)) \|_{H^{-1}} \geq \| \eta \|_{H^{-1}} = \| u(T) \|_{H^{-1}} = \| \|u(T)\| \|_{H^{-1}}, \quad (3.97)$$

Hence,

$$\begin{aligned} \overline{\lim} X_h \leq & \frac{1}{2} \| \|u_0\| \|_{H^{-1}}^2 - \frac{1}{2} \| \|u(T)\| \|_{H^{-1}}^2 - \int_0^T (\xi(t), v(t))_{L^2} dt \\ & - \int_0^T (D(v(t))v(t), u(t) - v(t))_{L^2}. \end{aligned} \quad (3.98)$$

Since, by integration by parts,

$$\frac{1}{2} \| \|u(T)\| \|_{H^{-1}}^2 - \frac{1}{2} \| \|u_0\| \|_{H^{-1}}^2 = \int_0^T (\xi(t), u(t))_{L^2} dt, \quad (3.99)$$

the estimate (3.92) follows.

4. Since the limit  $u$  is unique, the whole sequence  $(u_h)$  converges weakly to  $u$  in  $L^2(0, T; L^2)$ .  $\square$

Let from now on  $v_l$  be the solution of the approximating problem and  $u_l$  the stochastic particle density. We shall see below that by proceeding as usual we can show that

$$\sup_{t \leq T} E \left[ \| \|u_l(t) - v_l(t)\| \|_{\mathcal{H}^{-1}}^2 \right] \rightarrow 0 \quad (l \rightarrow 0). \quad (3.100)$$

Unfortunately, we do not have a nice compatibility of the norms in  $\mathcal{H}^{-1}$  and  $H^{-1}$ . If a sequence of lattice functions converges discretely to zero with respect to the norm in  $\mathcal{H}^{-1}(\mathcal{G}_h)$  we are (to the best of our knowledge) not able to conclude that the same is true for the extended versions with respect to the norm in  $H^{-1}(G)$ . This difficulty is circumvented below by proving only a weaker law of large numbers.

Note that if we define for a function  $\psi \in C_0^\infty(Q_T)$  approximating lattice functions  $\psi_l(\cdot, t)$  simply by setting

$$\psi_l(z, t) = \begin{cases} \psi(z, t) & \text{for } z \in \mathcal{G}_l^1 \\ 0 & \text{otherwise,} \end{cases} \quad (3.101)$$

then  $\psi_l$  converges uniformly to  $\psi$  on  $Q_T$  for  $l \rightarrow 0$ . Moreover the discrete derivative  $\partial^+ \psi_l$  converges uniformly to  $\partial_x \psi$  on  $Q_T$ . Furthermore, observe that for  $\psi \in C_0^\infty(Q_T)$

$$\int_0^T \langle u_l(t), \psi_l(t) \rangle_{\mathcal{H}_0^1} dt = \int_0^T (u_l(t), \psi_l(t))_{\mathcal{L}^2} dt = \int_0^T (u_l(t), \psi_l(t))_{L^2} dt \quad (3.102)$$

(cf. (2.53) and (2.27)). We make the following hypotheses for the scale parameters.

$$l \rightarrow 0, \quad n \rightarrow \infty, \quad (S1)$$

$$\sup_{\mathbb{R}} \left| \frac{1}{2} d l^2 - D \right| \rightarrow 0, \quad (S2)$$

$$\frac{1}{n} \sup_{\mathbb{R}} d \rightarrow 0. \quad (S3)$$

**Theorem 3.2.3** (Law of large numbers). *Let  $u$  be the solution of the weak PDE problem (3.69) to the initial value  $u_0$ . Assume that (S1)–(S3) hold and that  $u_l(0)$  converges strongly to  $u_0$  in  $L^2(G)$ . Then the particle density  $u_l$  converges to  $u$  in the following sense: For all  $\psi \in C_0^\infty(Q_T)$  and  $\varepsilon > 0$ ,*

$$P \left[ \left| \int_0^T \int_G u_l \psi \, dx \, dt - \int_0^T \int_G u \psi \, dx \, dt \right| > \varepsilon \right] \rightarrow 0.$$

The proof is based on the following auxiliary result that will be shown below.

**Theorem 3.2.4.** *Assume (S1)–(S3), and denote by  $v_l$  the solutions of the approximating problem (3.70) with  $D_l = \frac{1}{2}dl^2$  to the initial values  $v_{l,0} = u_l(0)$ . Then*

$$\sup_{t \leq T} E \left[ \|u_l(t) - v_l(t)\|_{\mathcal{H}^{-1}}^2 \right] \rightarrow 0.$$

**Proof of the law of large numbers.** Let  $\psi_l$  be the approximating lattice function of an arbitrary function  $\psi \in C_0^\infty(Q_T)$  defined above. (Assume that  $\psi$  is not identically zero to avoid trivialities.) Then

$$\begin{aligned} & P \left[ \left| \int_0^T \int_G u_l \psi \, dx \, dt - \int_0^T \int_G u \psi \, dx \, dt \right| > \varepsilon \right] \\ &= P \left[ \left| \int_0^T \int_G (u_l - v_l) \psi \, dx \, dt - \int_0^T \int_G (u - v_l) \psi \, dx \, dt \right| > \varepsilon \right] \\ &\leq P \left[ \left| \int_0^T \int_G (u_l - v_l) \psi \, dx \, dt \right| > \varepsilon/2 \right] + P \left[ \left| \int_0^T \int_G (u - v_l) \psi \, dx \, dt \right| > \varepsilon/2 \right]. \end{aligned} \tag{3.103}$$

The second term in the sum vanishes for  $l \rightarrow 0$  because  $v_l \rightarrow u$  in  $L^2(0, T; L^2)$ . As for the first term, observe that

$$\begin{aligned} & P \left[ \left| \int_0^T \int_G (u_l - v_l) \psi \, dx \, dt \right| > \varepsilon/2 \right] \leq P \left[ \left| \int_0^T \int_G (u_l - v_l) \psi_l \, dx \, dt \right| > \varepsilon/4 \right] \\ & \quad + P \left[ \left| \int_0^T \int_G (u_l - v_l) (\psi - \psi_l) \, dx \, dt \right| > \varepsilon/4 \right]. \end{aligned} \tag{3.104}$$

Again the second term in the sum tends to zero for  $l \rightarrow 0$ , since  $\sup_l \|u_l - v_l\|_{L^1(Q_T)} \leq C$  because of (3.81), and  $\|\psi - \psi_l\|_{L^\infty(Q_T)} \rightarrow 0$ . Let now  $C > 0$  be a constant such that

$\|\psi_l\|_{L^2(0,T;\mathcal{H}_0^1)} \geq C$  for sufficiently small  $l$ . Then

$$\begin{aligned}
& P \left[ \left| \int_0^T \int_G (u_l - v_l) \psi_l \, dx \, dt \right| > \varepsilon/4 \right] \\
&= P \left[ \left| \int_0^T \langle u_l(t) - v_l(t), \psi_l(t) \rangle_{\mathcal{H}_0^1} \, dt \right| > \varepsilon/4 \right] \\
&\leq P \left[ \left( \int_0^T \|u_l(t) - v_l(t)\|_{\mathcal{H}^{-1}}^2 \, dt \right)^{1/2} \left( \int_0^T \|\psi_l(t)\|_{\mathcal{H}_0^1}^2 \, dt \right)^{1/2} > \varepsilon/4 \right] \quad (3.105) \\
&\leq P \left[ \left( \int_0^T \|u_l(t) - v_l(t)\|_{\mathcal{H}^{-1}}^2 \, dt \right)^{1/2} > \varepsilon/(4C) \right] \\
&\leq \frac{(4C)^2}{\varepsilon^2} E \left[ \int_0^T \|u_l(t) - v_l(t)\|_{\mathcal{H}^{-1}}^2 \, dt \right] \rightarrow 0.
\end{aligned}$$

□

It remains to prove the auxiliary theorem 3.2.4. The proof is based on the next lemma that identifies a martingale related to the process  $\|u_l(t) - v_l(t)\|_{\mathcal{H}^{-1}}^2$ ,  $t \leq T$ .

**Lemma 3.2.5.** *The process  $(M_l(t))_{t \leq T}$  given by*

$$\begin{aligned}
M_l(t) &= \|u_l(t) - v_l(t)\|_{\mathcal{H}^{-1}}^2 - \|u_l(0) - v_l(0)\|_{\mathcal{H}^{-1}}^2 \\
&\quad + 2 \int_0^t \langle A_l(u_l(s)) - A_l(v_l(s)), u_l(s) - v_l(s) \rangle_{\mathcal{L}^2} \, ds - R_l(t), \quad (3.106)
\end{aligned}$$

where

$$R_l(t) = \frac{1}{n} \int_0^t (d(u_l(s)) u_l(s), \beta_l)_{\mathcal{L}^2}, \quad (3.107)$$

is a martingale. Here the  $\beta_l$  are lattice functions that satisfy

$$\max_{z \in \mathcal{G}_l} |\beta_l(z)| \leq C$$

independent of  $l$ .

**Proof.** Consider for fixed  $w_l \in \mathcal{L}^2$  the function  $g(\cdot, w_l) : S_l \rightarrow \mathbb{R}$  given by

$$u_l \mapsto g(u_l, w_l) = \|u_l - w_l\|_{\mathcal{H}^{-1}}^2,$$

and recall that

$$\|u_l - w_l\|_{\mathcal{H}^{-1}}^2 = \| \|u_l - w_l\|_{\mathcal{H}^{-1}}^2 = (u_l - w_l, (-\Delta_l)^{-1}(u_l - w_l))_{\mathcal{L}^2}. \quad (3.108)$$

We are going to compute  $L_l g(u_l, w_l)$ .

$$\begin{aligned}
L_l g(u_l, w_l) &= \sum_{z \in \mathcal{G}_l^1} n \frac{1}{2} d(u_l(z)) u(z) \left( \|u - \frac{1}{n} \chi_z + \frac{1}{n} \chi_{(z+l)} - w\|_{\mathcal{H}^{-1}}^2 \right. \\
&\quad \left. - 2 \|u - w\|_{\mathcal{H}^{-1}}^2 + \|u - \frac{1}{n} \chi_z + \frac{1}{n} \chi_{(z-l)} - w\|_{\mathcal{H}^{-1}}^2 \right) \\
&= \sum_{z \in \mathcal{G}_l^1} n \frac{1}{2} d(u_l(z)) u(z) \left( \frac{2}{n} (u - w, \chi_{(z-l)} - 2\chi_z + \chi_{(z+l)})_{\mathcal{H}^{-1}} \right. \\
&\quad \left. + \frac{1}{n^2} \|\chi_{(z-l)} - \chi_z\|_{\mathcal{H}^{-1}}^2 + \frac{1}{n^2} \|\chi_{(z+l)} - \chi_z\|_{\mathcal{H}^{-1}}^2 \right). \tag{3.109}
\end{aligned}$$

We set  $\tilde{u}_l = (-\Delta_l)^{-1} u_l$ , and  $\tilde{w}_l = (-\Delta_l)^{-1} w_l$ . Hence, we get

$$L_l g(u_l, w_l) = \sum_{z \in \mathcal{G}_l^1} n \frac{1}{2} d(u_l(z)) u(z) \left( \frac{2}{n} l^2 (\Delta_l \tilde{u}_l(z) - \Delta_l \tilde{w}_l(z)) + \frac{1}{n^2} \tilde{\beta}_l(z) \right), \tag{3.110}$$

where

$$\begin{aligned}
\tilde{\beta}_l(z) &= (\chi_{(z-l)} - \chi_z, (-\Delta_l)^{-1} (\chi_{(z-l)} - \chi_z))_{\mathcal{L}^2} \\
&\quad + (\chi_{(z+l)} - \chi_z, (-\Delta_l)^{-1} (\chi_{(z+l)} - \chi_z))_{\mathcal{L}^2}. \tag{3.111}
\end{aligned}$$

Note that

$$\left| \tilde{\beta}_l(z) \right| \leq C \left( \|\chi_{(z-l)} - \chi_z\|_{\mathcal{L}^2}^2 + \|\chi_{(z+l)} - \chi_z\|_{\mathcal{L}^2}^2 \right) \leq C l, \tag{3.112}$$

since the discrete Laplacian, as its continuous analogue, has a bounded inverse. We set  $\beta_l = \tilde{\beta}_l / (2l)$ . Finally, we get

$$\begin{aligned}
L_l g(u_l, w_l) &= -2 \frac{1}{2} l^2 (d(u_l) u_l, u_l - w_l)_{\mathcal{L}^2} + \frac{1}{n} (d(u_l) u_l, \beta_l)_{\mathcal{L}^2} \\
&= -2 a_l(u_l, u_l - w_l) + \frac{1}{n} (d(u_l) u_l, \beta_l)_{\mathcal{L}^2}. \tag{3.113}
\end{aligned}$$

Consider now for fixed  $w_l \in \mathcal{L}^2$  the function  $h(\cdot, w_l) : [0, T] \rightarrow \mathbb{R}$ ,

$$t \mapsto h(t, w_l) = \|w_l - v_l(t)\|_{\mathcal{H}^{-1}}^2,$$

and observe that

$$h'(t, w_l) = -2 (v_l'(t), w_l - v_l(t))_{\mathcal{H}^{-1}} = 2 a_l(v_l(t), w_l - v_l(t)). \tag{3.114}$$

Let  $\Phi : S_l \times [0, T]$ ,  $(u_l, t) \mapsto \Phi(u_l, t) = \|u_l - v_l(t)\|_{\mathcal{H}^{-1}}^2$ . It follows from Dynkin's formula that the process  $(M_l(t))_{t \leq T}$  given by

$$\begin{aligned}
M_l(t) &= \Phi(u_l(t), t) - \Phi(u_l(0), 0) - \int_0^t (L_l \Phi(u_l(s), s) + \partial_s \Phi(u_l(s), s)) ds \\
&= \|u_l(t) - v_l(t)\|_{\mathcal{H}^{-1}}^2 - \|u_l(0) - v_l(0)\|_{\mathcal{H}^{-1}}^2 \\
&\quad - \int_0^t (L_l g(u_l(s), v_l(s)) + h'(s, u_l(s))) ds \tag{3.115}
\end{aligned}$$

is a martingale. (Here we do not have to worry about  $\Phi$  being unbounded, since the particle density process is, for fixed  $l$ , bounded by construction.) Substituting the explicit computations in the equation above yields formula (3.106).  $\square$

**Proof of Theorem 3.2.4.** We set  $\hat{d} = \sup_{\mathbb{R}} d$ . By taking expectations in the martingale formula (3.106) and making use of the monotonicity of  $A_l$  we get the estimate

$$\begin{aligned} E \left[ \|u_l(t) - v_l(t)\|_{\mathcal{H}^{-1}}^2 \right] &\leq \|u_l(0) - v_l(0)\|_{\mathcal{H}^{-1}}^2 + C E \int_0^t \frac{\hat{d}}{n} (u_l(s), 1)_{\mathcal{L}^2} ds \\ &\leq \|u_l(0) - v_l(0)\|_{\mathcal{H}^{-1}}^2 + C \frac{\hat{d}}{n} \|u_l(0)\|_{\mathcal{L}^2} \rightarrow 0. \end{aligned} \quad (3.116)$$

Here the second inequality is due to the fact that by construction

$$0 \leq (u_l(t), 1)_{\mathcal{L}^2} \leq (u_l(0), 1)_{\mathcal{L}^2}, \quad t \geq 0. \quad \square$$

### 3.2.2 Gradient-activated diffusion

In the present section we are going to study an example for nonlinear diffusion where the intensity for a diffusive jump to a neighboring cell increases with the concentration gradient. That is, the intensity for a jump to the left or right is  $d(\partial^+ u)$  or  $d(-\partial^- u)$ , respectively. To save notation we again restrict the discussion to a single-species model in a one-dimensional reactor represented by an interval  $G = (0, L)$ . For the function  $d$  we again assume that  $d(p) = d(-p)$  for  $p \in \mathbb{R}$ , i.e., the jump intensity changes according to the absolute value of the concentration gradient. We call this behaviour gradient-activated diffusion.

The macroscopic PDE and its weak formulation

The macroscopic PDE that will be approached in the limit of large particle numbers reads, in one space dimension,

$$\begin{cases} \partial_t u - \partial_x (D(\partial_x u) \partial_x u) = 0 & \text{on } Q_T \\ u = 0 & \text{on } \partial G \times [0, T] \\ u(0) = u_0 & \text{on } G. \end{cases} \quad (3.117)$$

In order to discuss existence of a solution, we return here to our standard functional setting, i.e., we look for a function in  $H^1(0, T; H_0^1(G), L^2(G))$  that solves an appropriate weak formulation of (3.117). Again a monotonicity property plays a crucial role.

**Lemma 3.2.6.** *Let the mapping  $a : H_0^1 \times H_0^1 \rightarrow \mathbb{R}$  be given by*

$$a(u, v) = \int_G D(\partial_x u) \partial_x u \partial_x v \, dx, \quad u, v \in H_0^1, \quad (3.118)$$

and assume that  $D$  satisfies conditions (D1)–(D3) of the previous section. Then the mapping  $a(\cdot, \cdot)$  induces a (generally nonlinear) operator  $A : H_0^1 \rightarrow H^{-1}$  by

$$\langle A(u), v \rangle_{H_0^1} = a(u, v), \quad u, v \in H_0^1, \quad (3.119)$$

which is bounded, coercive, hemicontinuous and monotone.

**Proof.** The proof is similar to the proof of Lemma 3.2.1.  $\square$

The weak formulation of the PDE is obtained in the usual way by multiplying with a test function and integrating by parts:

$$\frac{d}{dt}(u(t), v)_{L^2} + a(u(t), v) = 0 \quad (3.120a)$$

for all  $v \in H_0^1$  and a.e.  $t \in (0, T)$ , and

$$u(0) = u_0 \in L^2. \quad (3.120b)$$

By a general theorem on first-order monotone evolution equations (see, e.g., Theorem 30.A in Zeidler (1990c) or Theorem 1.2 in Chapter 2 of Lions (1969)), the weak problem (3.120) has a unique solution.

The approximating problem

The approximating problem on the interior lattice points  $\mathcal{G}_h$  is given by

$$\begin{cases} u'_h - \partial^- D_h(\partial^+ u_h) \partial^+ u_h = 0 & \text{on } \mathcal{G}_h^1 \times (0, T) \\ u_h = 0 & \text{on } (\mathcal{G}_h \setminus \mathcal{G}_h^1) \times [0, T] \\ u_h(\cdot, 0) = u_{h,0} & \text{on } \mathcal{G}_h^1. \end{cases} \quad (3.121)$$

The functions  $D_h$  are supposed to satisfy conditions (D1)–(D3). Moreover, we assume that  $\sup_{\mathbb{R}} |D_h - D| \rightarrow 0$  for  $h \rightarrow 0$ . This is again a finite-dimensional ODE system with continuous right-hand side, which has a local solution according to the Peano theorem. The existence of a solution on the entire interval  $[0, T]$  follows from the derivation of the a-priori estimate (3.127) below. We define the mapping  $a_h : \mathcal{H}_0^1 \times \mathcal{H}_0^1 \rightarrow \mathbb{R}$  by

$$a_h(u_h, v_h) = (D_h(\partial^+ u_h) \partial^+ u_h, \partial^+ v_h)_{\mathcal{L}^2}, \quad u_h, v_h \in \mathcal{H}_0^1. \quad (3.122)$$

Note that  $a_h(\cdot, \cdot)$  induces again a bounded monotone operator  $A_h : \mathcal{H}_0^1 \rightarrow \mathcal{H}^{-1}$  by

$$\langle A_h(u_h), v_h \rangle_{\mathcal{H}_0^1} = a_h(u_h, v_h), \quad u_h, v_h \in \mathcal{H}_0^1. \quad (3.123)$$

A solution of (3.121) can be regarded as a function in  $C^1([0, T], \mathcal{H}_0^1)$  that solves the discrete weak problem

$$\frac{d}{dt}(u_h(t), v_h)_{\mathcal{L}^2} + a_h(u_h(t), v_h) = 0 \quad (3.124)$$

for all  $v_h \in \mathcal{H}_0^1$  and  $t \in (0, T)$ .

### The mesoscopic stochastic particle model

We work in the same setting as for the linear example model, i.e., the state space  $S_l$  of the particle density process  $(u_l(t))_{t \geq 0}$  is the subset of functions in  $\mathcal{L}^2(\mathcal{G}_l)$  taking values in  $\frac{1}{n}\mathbb{N}_0$ . Moreover, we assume that  $u_l(0)$  is non-random and that  $u_l$  satisfies Dirichlet boundary conditions, i.e.,  $u_l(z, \cdot) \equiv 0$  for  $z \in \mathcal{G}_l \setminus \mathcal{G}_l^1$ . The possible transitions are the following.

- A particle may leave cell  $z \in \mathcal{G}_l^1$  and jump to  $z \pm l$ .

$$\begin{aligned} q_l(u_l, u_l - \frac{1}{n}\chi_z + \frac{1}{n}\chi_{(z+l)}) &= n d(\partial^+ u_l(z)) u_l(z), \\ q_l(u_l, u_l - \frac{1}{n}\chi_z + \frac{1}{n}\chi_{(z-l)}) &= n d(-\partial^- u_l(z)) u_l(z), \end{aligned} \quad (3.125)$$

where it is assumed that  $l^2 d : \mathbb{R} \rightarrow \mathbb{R}_0^+$  satisfies conditions (D1)–(D3) from the previous section.

The existence of a particle density process  $(u_l(t))_{t \geq 0}$  with state space  $S_l$  and generator

$$L_l g(u_l) = \sum_{\tilde{u}_l \neq u_l} q_l(u_l, \tilde{u}_l) (g(\tilde{u}_l) - g(u_l)), \quad g \in \hat{C}(S_l), \quad (3.126)$$

follows from Theorem 3.1 in Chapter 8 of Ethier & Kurtz (1986).

### Law of large numbers

Here, as usual, we first show strong convergence of the solutions of the approximating problem (3.121) in  $L^2(0, T; L^2)$ .

**Theorem 3.2.7.** *Let  $u$  be the solution of the weak PDE problem (3.120) to the initial value  $u_0$ . Let  $(u_h)$ ,  $h \searrow 0$ , be a sequence of solutions of the approximating problem (3.124) to the initial values  $u_{h,0}$ , and assume that  $u_{h,0}$  converges strongly to  $u_0$  in  $L^2$ . Then  $J_h u_h$  converges weakly to  $Ju$  in  $L^2(0, T; X)$ , and  $u_h$  converges strongly to  $u$  in  $L^2(0, T; L^2)$ .*

**Proof.** 1. As usual, the a-priori estimates

$$\sup_h \max_{0 \leq t \leq T} \|u_h\|_{L^2} < \infty, \quad (3.127)$$

$$\sup_h \|u_h\|_{L^2(0, T; \mathcal{H}_0^1)} < \infty, \quad (3.128)$$

$$\sup_h \|u_h'\|_{L^2(0, T; \mathcal{H}^{-1})} < \infty, \quad (3.129)$$

follow immediately from the weak formulation of the approximating problem by inserting  $u_h(t)$  for  $v_h$  and integrating over time. Thus, we can conclude that there is a subsequence (still denoted by  $(u_h)$ ) such that  $J_h u_h \rightharpoonup Ju$  for a  $u \in L^2(0, T; H_0^1)$  and  $D_h(\partial^+ u_h) \partial^+ u_h \rightharpoonup \xi$  in  $L^2(0, T; L^2)$ . Moreover, by passing to a further subsequence if necessary, we have  $u_h(T) \rightharpoonup \eta$  in  $L^2$ .

2. We are going to prove that the limit  $u$  satisfies

$$u' + \xi = 0, \quad u(0) = u_0, \quad u(T) = \eta, \quad (3.130)$$

in the space  $H^1(0, T; H_0^1, L^2)$ . To this end, consider arbitrary functions  $v \in H_0^1, \varphi \in C^1[0, T]$ , and set  $v_h = R_h v$ . From the discrete weak formulation it follows that

$$\begin{aligned} (u_h(T), \varphi(T)v_h)_{\mathcal{L}^2} - (u_h(0), \varphi(0)v_h)_{\mathcal{L}^2} \\ = \int_0^T \left( \langle u_h'(t), \varphi(t)v_h \rangle_{\mathcal{H}_0^1} + \langle \varphi'(t)v_h, u_h(t) \rangle_{\mathcal{H}_0^1} \right) dt \end{aligned} \quad (3.131)$$

by integration by parts. Passing to the limit in all terms yields

$$\begin{aligned} (\eta, v)_{L^2} \varphi(T) - (u_0, v)_{L^2} \varphi(0) \\ = \int_0^T \langle u(t), v \rangle_{H_0^1} \varphi'(t) dt - \int_0^T \langle \xi(t), v \rangle_{H_0^1} \varphi(t) dt. \end{aligned} \quad (3.132)$$

Hence, by taking  $\varphi \in C_0^\infty(0, T)$ , we deduce that  $u$  is indeed in the space  $H^1(0, T; H_0^1, L^2)$  and has the time derivative  $u' = -\xi$ . It follows from the integration by parts formula for functions in  $H^1(0, T; H_0^1, L^2)$  by choosing a  $\varphi$  with  $\varphi(T) = 1$  and  $\varphi(0) = 0$ , respectively  $\varphi(T) = 0$  and  $\varphi(0) = 1$ , that  $u(T) = \eta$  and  $u(0) = u_0$ .

3. It remains to show that  $A(u) = \xi$ . This would follow from the hemicontinuity of  $A$  with the Minty trick if we were able to prove that

$$\int_0^T \langle \xi(t) - A(v(t)), u(t) - v(t) \rangle_{H_0^1} dt \geq 0 \quad \text{for all } v \in L^2(0, T; H_0^1). \quad (3.133)$$

It is again sufficient to show the estimate (3.133) for functions of the form  $v \varphi$ , where  $v \in H_0^1$  and  $\varphi \in C^1[0, T]$ , because linear combinations of such functions are dense. Let therefore  $v_h(t) = R_h v \varphi(t)$  and consider the non-negative expression

$$X_h = \int_0^T (D_h(\partial^+ u_h(t)) \partial^+ u_h(t) - D_h(\partial^+ v_h(t)) \partial^+ v_h(t), \partial^+ u_h(t) - \partial^+ v_h(t))_{\mathcal{L}^2} dt. \quad (3.134)$$

From the weak form of the approximating problem we get

$$\int_0^T (D_h(\partial^+ u_h(t)) \partial^+ u_h(t), \partial^+ u_h(t))_{\mathcal{L}^2} dt = \frac{1}{2} \|u_h(0)\|_{\mathcal{L}^2}^2 - \frac{1}{2} \|u_h(T)\|_{\mathcal{L}^2}^2. \quad (3.135)$$

Hence,

$$\begin{aligned} X_h = \frac{1}{2} \|u_h(0)\|_{\mathcal{L}^2}^2 - \frac{1}{2} \|u_h(T)\|_{\mathcal{L}^2}^2 - \int_0^T (D_h(\partial^+ u_h(t)) \partial^+ u_h(t), \partial^+ v_h(t))_{\mathcal{L}^2} dt \\ - \int_0^T (D_h(\partial^+ v_h(t)) \partial^+ v_h(t), \partial^+ u_h(t) - \partial^+ v_h(t))_{\mathcal{L}^2} dt. \end{aligned} \quad (3.136)$$

By taking limits superior we get

$$\begin{aligned} \overline{\lim} X_h = \frac{1}{2} \|u_0\|_{L^2}^2 - \frac{1}{2} \underline{\lim} \|u_h(T)\|_{L^2}^2 - \int_0^T (\xi(t), \partial_x v(t))_{L^2} dt \\ - \int_0^T (D(\partial_x v(t)) \partial_x v(t), \partial_x u(t) - \partial_x v(t))_{L^2} dt. \end{aligned} \quad (3.137)$$

Note that

$$\underline{\lim} \|u_h(T)\|_{\mathcal{L}^2} = \underline{\lim} \|u_h(T)\|_{L^2} \geq \|\eta\|_{L^2} = \|u(T)\|_{L^2}. \quad (3.138)$$

Therefore

$$\begin{aligned} \overline{\lim} X_h &\leq \frac{1}{2}\|u_0\|_{L^2}^2 - \frac{1}{2}\|u(T)\|_{L^2}^2 - \int_0^T (\xi(t), \partial_x v(t))_{L^2} dt \\ &\quad - \int_0^T (D(\partial_x v(t)) \partial_x v(t), \partial_x u(t) - \partial_x v(t))_{L^2}. \end{aligned} \quad (3.139)$$

Since

$$\frac{1}{2}\|u(T)\|_{L^2}^2 - \frac{1}{2}\|u_0\|_{L^2}^2 = \int_0^T (\xi(t), u(t))_{L^2} dt \quad (3.140)$$

by integration by parts, inequality (3.133) follows.

4. By possibly passing to a further subsequence we may assume that  $u_h$  converges even strongly to  $u$  in  $L^2(0, T; L^2)$  due to the compactness theorem 2.3.4. Convergence of the whole sequence follows from the uniqueness of the limit  $u$ .  $\square$

Let from now on  $v_l$  be the solution of the approximating problem (3.117) and  $u_l$  the particle density. We assume the following scaling relations:

$$l \rightarrow 0, \quad n \rightarrow \infty, \quad (S1)$$

$$\sup_{\mathbb{R}} |l^2 d - D| \rightarrow 0, \quad (S2)$$

$$\frac{1}{n} \sup_{\mathbb{R}} d \rightarrow 0. \quad (S3)$$

**Theorem 3.2.8** (Law of large numbers). *Let  $u$  be the solution of the weak PDE problem (3.120) to the initial value  $u_0$ . Assume that (S1)–(S3) hold and that  $u_l(0)$  converges strongly to  $u_0$  in  $L^2$ . Then*

$$E \left[ \|u_l - u\|_{L^2(0, T; L^2)}^2 \right] = E \left[ \|u_l - u\|_{L^2(Q_T)}^2 \right] \rightarrow 0.$$

In view of Theorem 3.2.7, the law of large numbers follows immediately from the next auxiliary result.

**Theorem 3.2.9.** *Assume (S1)–(S3), and denote by  $v_l$  the solutions of the approximating problem (3.121) with  $D_l = l^2 d$  to the initial values  $v_{l,0} = u_l(0)$ . Then*

$$\sup_{t \leq T} E \left[ \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 \right] \rightarrow 0.$$

The next lemma identifies a martingale related to the process  $\|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2$ ,  $t \leq T$ .

**Lemma 3.2.10.** *The process  $(M_l(t))_{t \leq T}$  given by*

$$\begin{aligned} M_l(t) &= \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 - \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 \\ &\quad + \int_0^t \langle A_l(u_l(s)) - A_l(v_l(s)), u_l(s) - v_l(s) \rangle_{\mathcal{H}_0^1} ds - R_l(t), \end{aligned} \quad (3.141)$$

where

$$R_l(t) = \frac{1}{n} \int_0^t ((d(\partial^+ u_l(s)) + d(\partial^- u_l(s))) u_l(s), 1)_{\mathcal{L}^2} ds, \quad (3.142)$$

is a martingale.

**Proof.** Consider for fixed  $w_l \in \mathcal{H}_0^1$  the function  $g(\cdot, w_l) : S_l \rightarrow \mathbb{R}$  given by

$$u_l \mapsto g(u_l, w_l) = \|u_l - w_l\|_{\mathcal{L}^2}^2.$$

We compute  $L_l g(u_l, w_l)$ .

$$\begin{aligned} L_l g(u_l, w_l) &= \sum_{z \in \mathcal{G}_l^1} n d(\partial^+ u_l(z)) u_l(z) \left( \|u_l - \frac{1}{n} \chi_z + \frac{1}{n} \chi_{(z+l)} - w_l\|_{\mathcal{L}^2}^2 - \|u_l - w_l\|_{\mathcal{L}^2}^2 \right) \\ &\quad + \sum_{z \in \mathcal{G}_l^1} n d(-\partial^- u_l(z)) u_l(z) \left( \|u_l - \frac{1}{n} \chi_z + \frac{1}{n} \chi_{(z-l)} - w_l\|_{\mathcal{L}^2}^2 - \|u_l - w_l\|_{\mathcal{L}^2}^2 \right) \\ &= \sum_{z \in \mathcal{G}_l^1} n d(\partial^+ u_l(z)) u_l(z) \left( \frac{2}{n} (u_l - w_l, \chi_{(z+l)} - \chi_z)_{\mathcal{L}^2} + \frac{1}{n^2} \|\chi_{(z+l)} - \chi_z\|_{\mathcal{L}^2}^2 \right) \\ &\quad + \sum_{z \in \mathcal{G}_l^1} n d(-\partial^- u_l(z-l)) u_l(z) \times \\ &\quad \times \left( \frac{2}{n} (u_l - w_l, \chi_{(z-l)} - \chi_z)_{\mathcal{L}^2} + \frac{1}{n^2} \|\chi_{(z-l)} - \chi_z\|_{\mathcal{L}^2}^2 \right) \\ &= \sum_{z \in \mathcal{G}_l^1} n d(\partial^+ u_l(z)) 2 \frac{l}{n} \left( u_l(z+l) - w_l(z+l) - (u_l(z) - w_l(z)) \right) \\ &\quad + \sum_{z \in \mathcal{G}_l^1} n d(-\partial^- u_l(z-l)) u_l(z) \times \\ &\quad \times 2 \frac{l}{n} \left( u_l(z-l) - w_l(z-l) - (u_l(z) - w_l(z)) \right) \\ &\quad + \sum_{z \in \mathcal{G}_l^1} n u_l(z) \left( d(\partial^+ u_l(z)) + d(-\partial^- u_l(z-l)) \right) 2 \frac{l}{n^2}. \end{aligned} \quad (3.143)$$

By introducing discrete derivatives and making use of assumption (D2) it follows that

$$\begin{aligned}
L_l g(u_l, w_l) &= 2 \sum_{z \in \mathcal{G}_l^1} l^2 d(\partial^+ u_l(z)) u_l(z) (\partial^+ u_l(z) - \partial^+ w_l(z)) \\
&\quad - 2 \sum_{z \in \mathcal{G}_l^1} l^2 d(\partial^+ u_l(z-l)) u_l(z) (\partial^+ u_l(z-l) - \partial^+ w_l(z-l)) \\
&\quad + 2 \sum_{z \in \mathcal{G}_l^1} \frac{l}{n} u_l(z) (d(\partial^+ u_l(z)) + d(\partial^- u_l(z))) \\
&= 2l \sum_{z \in \mathcal{G}_l^1} u_l(z) \partial^- (l^2 d(\partial^+ u_l)(\partial^+ u_l - \partial^+ w_l))(z) \\
&\quad + 2 \sum_{z \in \mathcal{G}_l^1} \frac{l}{n} u_l(z) (d(\partial^+ u_l(z)) + d(\partial^- u_l(z)))
\end{aligned} \tag{3.144}$$

Hence, by a discrete integration by parts,

$$\begin{aligned}
L_l g(u_l, w_l) &= -2 (l^2 d(\partial^+ u_l) \partial^+ u_l, \partial^+ u_l - \partial^+ w_l)_{\mathcal{L}^2} \\
&\quad + \frac{2}{n} ((d(\partial^+ u_l) + d(\partial^- u_l)) u_l, 1)_{\mathcal{L}^2} \\
&= -2 a_l(u_l, u_l - w_l) + \frac{2}{n} ((d(\partial^+ u_l) + d(\partial^- u_l)) u_l, 1)_{\mathcal{L}^2}.
\end{aligned} \tag{3.145}$$

We now consider for arbitrary but fixed  $w_l \in \mathcal{H}_0^1$  the function  $h(\cdot, w_l) : [0, T] \rightarrow \mathbb{R}$ ,

$$t \mapsto h(t, w_l) = \|w_l - v_l(t)\|_{\mathcal{L}^2}^2. \tag{3.146}$$

Note that

$$h'(t, w_l) = -2 (v_l'(t), w_l - v_l(t))_{\mathcal{L}^2} = 2 a_l(v_l(t), w_l - v_l(t)). \tag{3.147}$$

Let  $\Phi : S_l \times [0, T] \rightarrow \mathbb{R}$  be given by  $(u_l, t) \mapsto \Phi(u_l, t) = \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2$ , and recall that the process  $(u_l(t))$  is (for fixed  $l$ ) bounded by construction. Thus it follows from Dynkin's formula that the process

$$\begin{aligned}
M_l(t) &= \Phi(u_l(t), t) - \Phi(u_l(0), 0) - \int_0^t \left( L_l \Phi(u_l(s), s) + \partial_s \Phi(u_l(s), s) \right) ds \\
&= \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 - \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 \\
&\quad - \int_0^t \left( L_l g(u_l(s), v_l(s)) + h'(s, u_l(s)) \right) ds,
\end{aligned} \tag{3.148}$$

$t \leq T$ , is a martingale. Substituting the explicit computations above yields Eq. (3.141).  $\square$

**Proof of Theorem 3.2.9.** By taking expectations in Eq. (3.141) and making use of the monotonicity of  $A_l$  we get the estimate

$$\begin{aligned}
E \left[ \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 \right] &\leq \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 + 2E \int_0^t \frac{\hat{d}}{n} (u_l(s), 1)_{\mathcal{L}^2} ds \\
&\leq \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 + C \frac{\hat{d}}{n} \|u_l(0)\|_{\mathcal{L}^2},
\end{aligned} \tag{3.149}$$

where  $\hat{d} = \sup_{\mathbb{R}} d$ . In view of the hypotheses we can conclude that the right hand side tends to zero, which finishes the proof.  $\square$

### 3.3 A refined law of large numbers

In this section we discuss how some of the laws of large numbers obtained so far may (in a certain sense) be refined. However, we restrict the discussion to the linear example model (cf. Section 2.3). We first show stronger convergence of the solutions  $u_h$  of the approximating problem (2.54).

**Theorem 3.3.1.** *We make the same hypotheses as in Theorem 2.3.2. If, in addition,  $u_0 \in H_0^1$ , and  $\partial^+ u_{h,0} \rightarrow \partial_x u_0$  strongly in  $L^2$ , then the solutions  $u_h$  of the approximating problem (2.58) converge strongly to the solution  $u$  of the weak problem (2.40) in  $C([0, T], L^2)$ . That is,*

$$\sup_{0 \leq t \leq T} \|u_h(t) - u(t)\|_{L^2} \rightarrow 0 \quad (h \rightarrow 0).$$

**Proof.** We multiply the discretised PDE (2.54) by  $u'_h(z, t)$  and sum over  $z$ , which yields, after a discrete integration by parts,

$$\begin{aligned} \|u'_h(t)\|_{\mathcal{L}^2}^2 + \frac{1}{2} \frac{d}{dt} \|\partial^+ u_h(z, t)\|_{\mathcal{L}^2}^2 &= (f(u_h(t)), u'_h(t))_{\mathcal{L}^2} \\ &\leq \frac{1}{2\varepsilon} \|f(u_h(t))\|_{\mathcal{L}^2}^2 + 2\varepsilon \|u'_h(t)\|_{\mathcal{L}^2}^2. \end{aligned} \quad (3.150)$$

We choose  $\varepsilon = 1/4$  and make use of the linearity of  $f$  to get the estimate

$$\frac{1}{2} \|u'_h(t)\|_{\mathcal{L}^2}^2 + \frac{1}{2} \frac{d}{dt} \|\partial^+ u_h(z, t)\|_{\mathcal{L}^2}^2 \leq C \|u_h(t)\|_{\mathcal{L}^2}^2. \quad (3.151)$$

By recalling the a-priori estimate (2.72), we have

$$\frac{d}{dt} \|\partial^+ u(z, t)\|_{\mathcal{L}^2}^2 \leq C \max_{0 \leq t \leq T} \|u_h(t)\|_{\mathcal{L}^2}^2 \leq C, \quad (3.152)$$

where the constant  $C$  is independent of  $h$ . Thus it follows from the discrete Poincaré inequality that the solutions  $u_h$  constitute an equicontinuous family of functions in  $C([0, T], L^2)$ . Moreover,

$$\sup_h \max_{0 \leq t \leq T} \|u_h(t)\|_{\mathcal{H}_0^1} \leq C. \quad (3.153)$$

Because of the compactness theorem 2.3.3 there is, for each  $t \in [0, T]$ , a subsequence of  $(u_h(t))$  that converges strongly in  $L^2$ . Hence, we can apply an appropriate version of the Arzéla-Ascoli theorem (see, e.g., Zeidler (1990a)) from which the claim follows.  $\square$

Denote from now on the solutions of the approximating problem by  $v_l$  and the stochastic particle density by  $u_l$ .

**Theorem 3.3.2** (Refined law of large numbers). *We make the same assumptions as in Theorem 2.3.6. If, in addition,  $u_0 \in H_0^1$ , and  $\partial^+ u_l(0) \rightarrow \partial_x u_0$  strongly in  $L^2$ , then  $u_l$*

converges to the solution  $u$  of the weak PDE problem (2.40) to the initial value  $u_0$  in the following sense:

$$P \left[ \sup_{t \leq T} \|u_l(t) - u(t)\|_{L^2} > \varepsilon \right] \rightarrow 0 \quad \text{for all } \varepsilon > 0.$$

**Proof.** Recall that the process

$$\begin{aligned} M_l(t) &= \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 - \|u_l(0) - v_l(0)\|_{\mathcal{L}^2}^2 \\ &\quad + 2 \int_0^t a_l(u_l(s) - v_l(s), u_l(s) - v_l(s)) ds \\ &\quad - 2 \int_0^t (f(u_l(s)) - f(v_l(s)), u_l(s) - v_l(s))_{\mathcal{L}^2} ds - R_l(t), \end{aligned}$$

$t \leq T$ , where

$$R_l(t) = \frac{1}{n} \int_0^t \left( 2d(u_l(s), 1)_{\mathcal{L}^2} + (k_1 u_l(s), 1)_{\mathcal{L}^2} + (k_2 u_l(s), 1)_{\mathcal{L}^2} \right) ds,$$

is a local martingale (cf. Lemma 2.3.8). From the monotonicity of the bilinear form  $a_l(\cdot, \cdot)$  we deduce that

$$X_l(t) = \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 - 2 \int_0^t (f(u_l(s)) - f(v_l(s)), u_l(s) - v_l(s))_{\mathcal{L}^2} ds - R_l(t), \quad (3.154)$$

$t \leq T$ , is a (local) supermartingale. We denote the two latter terms on the right by  $Y_l(t)$  and  $Z_l(t)$ . By stopping at  $\tau_p$  and letting  $p \rightarrow \infty$ , it follows from a well-known maximal inequality (see, e.g., Kallenberg (2002), Chapter 7) that, for arbitrary  $\varepsilon > 0$ ,

$$\varepsilon P \left[ \sup_{t \leq T} |X_l(t)| > \varepsilon \right] \leq 3 \sup_{t \leq T} E \left[ |X_l(t)| \right]. \quad (3.155)$$

With the aid of Theorem 2.3.7 we can conclude that  $\sup_{t \leq T} E \left[ |X_l(T)| \right] \rightarrow 0$ . Moreover,

$$\begin{aligned} P \left[ \sup_{t \leq T} |X_l(t)| > \varepsilon \right] &\geq \\ &P \left[ \sup_{t \leq T} |Y_l(t)| < \varepsilon/2, \sup_{t \leq T} |Z_l(t)| < \varepsilon/2, \sup_{t \leq T} \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 > 2\varepsilon \right]. \end{aligned} \quad (3.156)$$

Therefore

$$\begin{aligned} P \left[ \sup_{t \leq T} |X_l(t)| > \varepsilon \right] &\geq P \left[ \sup_{t \leq T} \|u_l(t) - v_l(t)\|_{\mathcal{L}^2}^2 > 2\varepsilon \right] \\ &\quad - P \left[ \sup_{t \leq T} |Y_l(t)| \geq \varepsilon/2 \right] - P \left[ \sup_{t \leq T} |Z_l(t)| \geq \varepsilon/2 \right]. \end{aligned} \quad (3.157)$$

It follows again from Theorem 2.3.7 that each of the two latter terms converges to zero, which finishes the proof.  $\square$

## Discussion

In this chapter we have seen how the ideas introduced in Chapter 2 can be extended to treat certain nonlinear particle models. Moreover, we have indicated for the example model how the laws of large numbers may be refined. (Note, however, that Theorem 3.3.1 does not contain Theorem 2.3.6.) Again we have to postulate scaling relations that are the same or similar to those used in the previous chapter. If the jump intensities depend on the local concentration ( $d = d(u)$ ) we get the same limit equation as in the context of scaling limits for the zero-range process (see, e.g., Kipnis & Landim (1999)). There the limit is obtained by rescaling space and time variables. Nonconstant diffusion coefficients play a role in the modelling of self-organisation of microorganisms (Ben-Jacob et al., 2000) and surface reactions (Naumovets, 2005).

In the present chapter and the previous one we have always imposed homogeneous Dirichlet boundary conditions, but other boundary conditions can probably be handled with similar techniques. However, we have not carried out the proofs. Although we were able to treat quite general classes of reaction-diffusion systems, the cases considered in the present work are by no means exhaustive. The same or similar techniques may perhaps be applied to models that include convection, cross-diffusion of different species, or ‘freezing’ of particles (Stefan problems).



## Chapter 4

# Stochastic simulations

### Overview

In this chapter we turn to the more practical question of simulating paths of the particle density processes (the rescaled versions of the mesoscopic stochastic particle model, cf. Section 2.1.2) for the models discussed in the previous chapters. The discussion applies, in particular, to the particle density process associated to the model for CO oxidation on Pt(110) presented in Section 1.3. In the absence of drift terms it is, in principle, possible to perform ‘exact’ simulations in the sense that the probability distribution of the simulated paths is *exactly* the same as for the paths of the original process. The simulation method is, however, very time-consuming and thus often not applicable in practice. In the next section we therefore introduce an algorithm for the *approximate* simulation of the paths of the particle density processes. In Section 4.2 we use this algorithm for the simulation of raindrop patterns (cf. Figs. 1.8 and 1.9) with the mesoscopic stochastic particle model for CO oxidation on Pt(110) from Section 1.3.

### 4.1 The simulation algorithm

We present here a simple algorithm for the approximate simulation of the density processes associated to mesoscopic stochastic particle models of the kind introduced in Chapters 2 and 3. The simulation method we propose is not exact, but the simulated approximating process converges in distribution to the original one if the time step  $h$  tends to zero. Exact simulation of spatial models does not seem feasible at the present stage, since the well-known ‘direct method’ (Gillespie, 1977) is far too slow, and the faster ‘next reaction method’ (Gibson & Bruck, 2000) needs too much memory. Therefore we have to resort to an approximate method. A spatially homogeneous variant of the algorithm given below has been described in Gillespie (2001) as the ‘ $\tau$ -leap method’.

### A basic algorithm without temperature variable

Generally, if we neglect temperature effects, the density process  $(\mathbf{u}_l(t))$  jumps from a state  $\mathbf{u}_l = (u_{l,1}, \dots, u_{l,n_s})$  to another state  $\tilde{\mathbf{u}}_l$  in its state space  $S_l = \prod_{k=1}^{|\mathcal{G}_l|} \frac{1}{n} \mathbb{N}_0^{n_s}$  with a certain intensity  $q_l(\mathbf{u}_l, \tilde{\mathbf{u}}_l)$ . Here  $n$  denotes the number of sites per cell,  $\mathcal{G}_l$  is the set of midpoints of the cells with edge length  $l$ , and  $|\mathcal{G}_l|$  is the number of cells. In the isothermal case such jumps are the only possible transitions; the process  $(\mathbf{u}_l(t))$  is a pure jump process. In this chapter we henceforth skip the subscript  $l$ , since we are only interested in simulating paths of  $(\mathbf{u}_l(t))$  for fixed  $l$ . The transition intensities  $q(\mathbf{u}, \tilde{\mathbf{u}})$  are non-zero only for a certain finite number of  $\tilde{\mathbf{u}}$  of the form  $\tilde{\mathbf{u}} = \mathbf{u} + \boldsymbol{\delta}$ , where  $\boldsymbol{\delta}$  is taken from a finite set  $\mathcal{T}$ .

Although it is, in principle, possible to simulate the paths of the process  $(\mathbf{u}(t))$  directly, the complexity of the spatial models forces us to use an approximate simulation method. We shall simulate paths of an approximating process  $(\mathbf{u}^h(t))$ , where the parameter  $h > 0$  is the size of a time step. The process  $(\mathbf{u}^h(t))$  is an approximation of  $(\mathbf{u}(t))$  in so far as it converges in distribution to  $(\mathbf{u}(t))$  for  $h \searrow 0$ .

Let  $Y_{\boldsymbol{\delta}}$ ,  $\boldsymbol{\delta} \in \mathcal{T}$ , be mutually independent standard Poisson processes. We define iteratively the process in discrete time  $(\mathbf{u}^h(k))_{k \in \mathbb{N}_0}$  as follows.

(0) Let  $\mathbf{u}^h(0) = \mathbf{u}(0)$ . (We assume that  $\mathbf{u}(0)$  is deterministic.)

(1) Next, set for each  $\boldsymbol{\delta} \in \mathcal{T}$

$$\tau_{\boldsymbol{\delta}}(0) = \begin{cases} q(\mathbf{u}^h(0), \mathbf{u}^h(0) + \boldsymbol{\delta}) h & \text{if } \mathbf{u}^h(0) + \boldsymbol{\delta} \in S \\ 0 & \text{otherwise,} \end{cases} \quad (4.1)$$

and let

$$N_{\boldsymbol{\delta}}(0) = Y_{\boldsymbol{\delta}}(\tau_{\boldsymbol{\delta}}(0)). \quad (4.2)$$

Then we compute  $\mathbf{u}^h(1)$  as

$$\mathbf{u}^h(1) = \left( \mathbf{u}^h(0) + \sum_{\boldsymbol{\delta} \in \mathcal{T}} \boldsymbol{\delta} N_{\boldsymbol{\delta}}(0) \right) \vee 0. \quad (4.3)$$

Here, for  $\mathbf{w} \in S$ ,  $\mathbf{w} \vee 0$  is the state  $\tilde{\mathbf{w}} \in S$  with  $\tilde{w}_j(z) = \max(w_j(z), 0)$ ,  $z \in \mathcal{G}$ ,  $j = 1, \dots, n_s$ . (Recall that  $n_s$  is the number of species.) This is necessary to avoid unphysical concentrations.

(2) For the second time step we define for each  $\boldsymbol{\delta} \in \mathcal{T}$

$$\tau_{\boldsymbol{\delta}}(1) = \begin{cases} q(\mathbf{u}^h(1), \mathbf{u}^h(1) + \boldsymbol{\delta}) h & \text{if } \mathbf{u}^h(1) + \boldsymbol{\delta} \in S \\ 0 & \text{otherwise.} \end{cases} \quad (4.4)$$

Furthermore, let  $\vartheta(s)$  be the time shift by  $s$ , i.e.,  $\vartheta(s)Y_{\delta}(\cdot) = Y_{\delta}(s + \cdot)$ . We set

$$N_{\delta}(1) = \vartheta(\tau_{\delta}(0))Y_{\delta}(\tau_{\delta}(1)) - \vartheta(\tau_{\delta}(0))Y_{\delta}(0) \quad (4.5)$$

and compute  $\mathbf{u}^h(2)$  as

$$\mathbf{u}^h(2) = \left( \mathbf{u}^h(1) + \sum_{\delta \in \mathcal{T}} \delta N_{\delta}(1) \right) \vee 0. \quad (4.6)$$

(k) For the general time step from  $k$  to  $k + 1$  we define for each  $\delta \in \mathcal{T}$

$$\tau_{\delta}(k) = \begin{cases} q(\mathbf{u}^h(k), \mathbf{u}^h(k) + \delta)h & \text{if } \mathbf{u}^h(k) + \delta \in S \\ 0 & \text{otherwise} \end{cases} \quad (4.7)$$

and set

$$N_{\delta}(k) = \vartheta(\tau_{\delta}(k-1))\vartheta(\tau_{\delta}(k-2)) \cdots \vartheta(\tau_{\delta}(0))Y_{\delta}(\tau_{\delta}(k)) \\ - \vartheta(\tau_{\delta}(k-1))\vartheta(\tau_{\delta}(k-2)) \cdots \vartheta(\tau_{\delta}(0))Y_{\delta}(0).$$

We again compute  $\mathbf{u}^h(k+1)$  as

$$\mathbf{u}^h(k+1) = \left( \mathbf{u}^h(k) + \sum_{\delta \in \mathcal{T}} \delta N_{\delta}(k) \right) \vee 0. \quad (4.8)$$

The algorithm introduced above obviously defines a Markov chain  $(\mathbf{u}^h(k))_{k \in \mathbb{N}_0}$  in discrete time in the same state space as the density process  $(\mathbf{u}(t))$ . In order to simulate its paths, we have to produce at each time step samples of the random variables  $N_{\delta}(k)$ ,  $\delta \in \mathcal{T}$ . Clearly, conditional on  $\mathbf{u}^h(k)$ , the  $N_{\delta}(k)$  are independent and distributed according to a Poisson law with parameter  $\tau_{\delta}(k)$ .

Consider now the process  $(\mathbf{u}^h([t/h]))$  as approximation of the particle density process  $(\mathbf{u}(t))$ . (Here  $[x]$  denotes the largest integer smaller than or equal to  $x$ .) As the size of the time step  $h$  converges to 0, the process  $(\mathbf{u}^h([t/h]))$  converges in distribution to the original particle density process  $(\mathbf{u}(t))$ . This follows from a convergence theorem in Kallenberg (2002) (Theorem 19.28), for instance, and the properties of the Poisson process. Loosely speaking, one has to show that the generator of the approximating process converges for  $h \searrow 0$  to the generator of the original one which is given by

$$Lg(\mathbf{u}) = \sum_{\substack{\delta \in \mathcal{T}: \\ \mathbf{u} + \delta \in S}} q(\mathbf{u}, \mathbf{u} + \delta)(g(\mathbf{u} + \delta) - g(\mathbf{u})). \quad (4.9)$$

#### An algorithm including temperature effects

The basic algorithm of the previous paragraph can easily be adapted for the simulation of the stochastic model from Section 1.3 including temperature variables  $\theta(z, t)$ . Let  $(\mathbf{u}(t), \theta(t))$  be

a Markov jump process with inter-jump drift in the state space  $S_l = \prod_{k=1}^{|\mathcal{G}_l|} (\frac{1}{n} \mathbb{N}_0^{n_s} \times \mathbb{R})$  with generator of the same form as in Section 1.3. Here  $\mathcal{G}_l$  denotes the set of midpoints of the cells of width  $l$ . We shall again skip the subscript  $l$ , since we are interested only in simulating the process corresponding to a certain fixed  $l$ . As before, for each state  $(\mathbf{u}, \theta) \in S$  there is only a finite number of other states  $(\tilde{\mathbf{u}}, \tilde{\theta}) = (\mathbf{u} + \boldsymbol{\delta}, \theta + \varepsilon)$  that are reachable through a jump, i.e.,  $(\boldsymbol{\delta}, \varepsilon)$  ranges over a finite set  $\mathcal{T}$ .

We define again auxiliary Poisson processes  $Y_{(\boldsymbol{\delta}, \varepsilon)}$  on some probability space. Let  $b_z(\mathbf{u}, \theta)$  be the rate at which the temperature changes ‘deterministically’ in cell  $z$  if the system is in the state  $(\mathbf{u}, \theta)$ . The approximating process is then constructed as follows.

(0) First, let  $(\mathbf{u}^h(0), \theta^h(0)) = (\mathbf{u}(0), \theta(0))$ . (We assume that  $(\mathbf{u}(0), \theta(0))$  is deterministic.)

(1) Next, set for each  $(\boldsymbol{\delta}, \varepsilon) \in \mathcal{T}$

$$\tau_{(\boldsymbol{\delta}, \varepsilon)}(0) = \begin{cases} q((\mathbf{u}^h(0), \theta^h(0)), (\mathbf{u}^h(0) + \boldsymbol{\delta}, \theta^h(0) + \varepsilon)) h & \text{if } (\mathbf{u}^h(0) + \boldsymbol{\delta}, \theta^h(0) + \varepsilon) \in S \\ 0 & \text{otherwise,} \end{cases} \quad (4.10)$$

and let

$$N_{(\boldsymbol{\delta}, \varepsilon)}(0) = Y_{(\boldsymbol{\delta}, \varepsilon)}(\tau_{(\boldsymbol{\delta}, \varepsilon)}(0)). \quad (4.11)$$

We set

$$\begin{aligned} \mathbf{u}^h(1) &= \left( \mathbf{u}^h(0) + \sum_{(\boldsymbol{\delta}, \varepsilon) \in \mathcal{T}} \boldsymbol{\delta} N_{(\boldsymbol{\delta}, \varepsilon)}(0) \right) \vee 0, \\ \theta^h(z, 1) &= \theta^h(z, 0) + \sum_{(\boldsymbol{\delta}, \varepsilon) \in \mathcal{T}} \varepsilon(z) N_{(\boldsymbol{\delta}, \varepsilon)}(0) + b_z(\mathbf{u}^h(0), \theta^h(0)) h, \quad z \in \mathcal{G}. \end{aligned} \quad (4.12)$$

(k) For the general time step from  $k$  to  $k + 1$  we define

$$\tau_{(\boldsymbol{\delta}, \varepsilon)}(k) = \begin{cases} q((\mathbf{u}^h(k), \theta^h(k)), (\mathbf{u}^h(k) + \boldsymbol{\delta}, \theta^h(k) + \varepsilon)) h & \text{if } (\mathbf{u}^h(k) + \boldsymbol{\delta}, \theta^h(k) + \varepsilon) \in S \\ 0 & \text{otherwise,} \end{cases} \quad (4.13)$$

and we set

$$\begin{aligned} N_{(\boldsymbol{\delta}, \varepsilon)}(k) &= \vartheta(\tau_{(\boldsymbol{\delta}, \varepsilon)}(k-1)) \vartheta(\tau_{(\boldsymbol{\delta}, \varepsilon)}(k-2)) \cdots \vartheta(\tau_{(\boldsymbol{\delta}, \varepsilon)}(0)) Y_{(\boldsymbol{\delta}, \varepsilon)}(\tau_{(\boldsymbol{\delta}, \varepsilon)}(k)) \\ &\quad - \vartheta(\tau_{(\boldsymbol{\delta}, \varepsilon)}(k-1)) \vartheta(\tau_{(\boldsymbol{\delta}, \varepsilon)}(k-2)) \cdots \vartheta(\tau_{(\boldsymbol{\delta}, \varepsilon)}(0)) Y_{(\boldsymbol{\delta}, \varepsilon)}(0). \end{aligned} \quad (4.14)$$

Finally,

$$\begin{aligned} \mathbf{u}^h(k+1) &= \left( \mathbf{u}^h(k) + \sum_{(\boldsymbol{\delta}, \varepsilon) \in \mathcal{T}} \boldsymbol{\delta} N_{(\boldsymbol{\delta}, \varepsilon)}(k) \right) \vee 0, \\ \theta^h(z, k+1) &= \theta^h(z, k) + \sum_{(\boldsymbol{\delta}, \varepsilon) \in \mathcal{T}} \varepsilon(z) N_{(\boldsymbol{\delta}, \varepsilon)}(k) + b_z(\mathbf{u}^h(k), \theta^h(k)) h, \quad z \in \mathcal{G}. \end{aligned} \quad (4.15)$$

The process  $(\mathbf{u}^h(k), \theta^h(k))_{k \in \mathbb{N}_0}$  constructed in this way is still a discrete time Markov chain. In order to simulate its paths, we have to produce at each time step samples of the random variables  $N_{(\delta, \varepsilon)}(k)$  and compute the drift velocities  $b_z(\mathbf{u}^h(k), \theta^h(k))$ . Conditional on  $(\mathbf{u}^h(k), \theta^h(k))$ , the  $N_{(\delta, \varepsilon)}(k)$  are again independent and distributed according to a Poisson law with parameter  $\tau_{(\delta, \varepsilon)}(k)$ .

The approximating process in continuous time corresponding to  $(\mathbf{u}^h(k), \theta^h(k))$  is  $(\mathbf{u}^h([t/h]), \theta_h([t/h]))_{t \geq 0}$ . Under suitable hypothesis on  $q(\cdot, \cdot)$  and  $b_z$  it should not be too difficult to show convergence in distribution of the approximating process to the original one for  $h \searrow 0$ . However, we shall not attempt to give a precise proof.

### Implementation

The two algorithms introduced above can conveniently be implemented in the C++ programming language if a physical cell is represented by an appropriate cell class object. (To get a copy of the code `spatCat` write an e-mail to the author.) For the simulation of Poisson random variables we have used the algorithm proposed in Press et al. (1992).

## 4.2 Simulation of raindrop patterns

Here the stochastic model for CO oxidation on Pt(110) from Section 1.3 is employed to simulate the raindrop patterns described in Section 1.1.3 (cf. Figs. 1.8 and 1.9). Since the simulated patch of Pt(110) surface is relatively small, heat diffusion is assumed infinitely fast. At low pressures the model behaves practically isothermally and follows the deterministic path very closely. In particular, no spontaneous nucleation in bistable or excitable parameter regions is observed. At intermediate pressures, however, significant fluctuations become visible (with  $n = 10^3$  adsorption sites per cell, transition rate  $d$  of CO molecules from cell to cell corresponding to  $D_u = 1.4 \times 10^{-14}$  m<sup>2</sup>/s) and critical nuclei do form spontaneously. A computer simulation for oxygen partial pressure  $p_v = 10^{-2}$  mbar of nucleation, pulse formation and subsequent propagation failure is reproduced in Fig. 4.2. A corresponding simulation with the PDE system (1.18a)–(1.18d) is depicted in Fig. 4.1. In order to match the time scale of the experiments, we had to increase the pre-exponential factors for desorption, reaction and structural phase transition by two orders of magnitude. That is, the simulations were performed with parameters  $\nu_2 = 5 \times 10^{17}$  s<sup>-1</sup>,  $\nu_3 = 5 \times 10^9$  s<sup>-1</sup>,  $\nu_5 = 10^5$  s<sup>-1</sup> and  $\nu_6 = 2 \times 10^4$  s<sup>-1</sup> (cf. Tables 1.2, 1.3 and 1.4). This can be rationalised by taking into account that at higher pressures interactions between adparticles become more important due to increased overall coverages. Nevertheless, the situation is not satisfactory and careful parameter estimation experiments would be highly desirable. Another slight flaw of the simulations is that, in order to limit computation time, CO diffusion was chosen about one to two orders of magnitude too slow. Consequently, the simulated raindrops are about

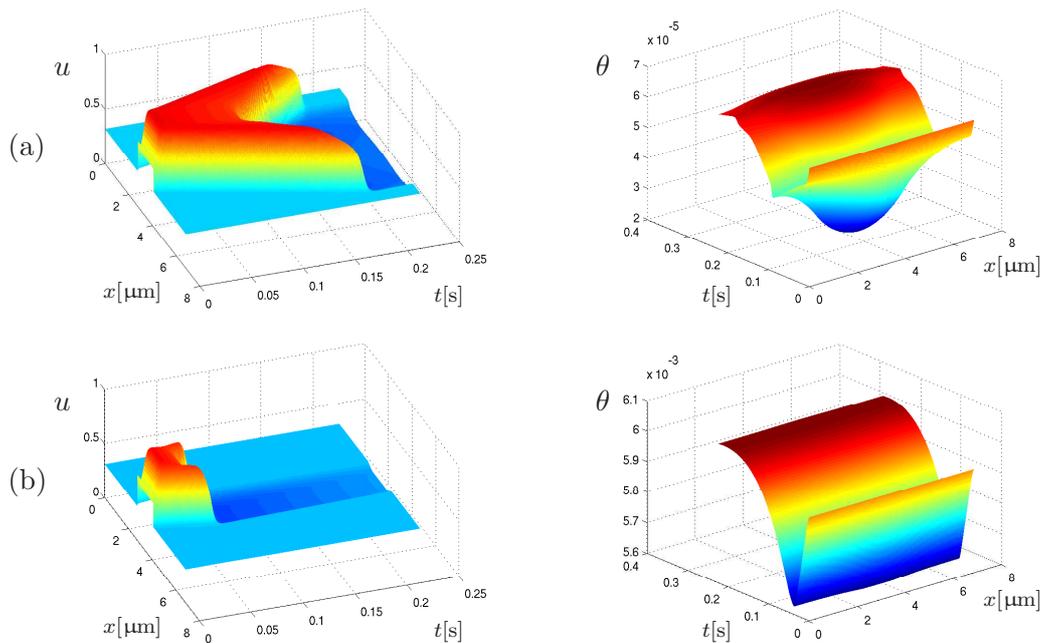


Figure 4.1: Simulation of pulse propagation performed with the thermo-kinetic model equations (1.18a)–(1.18d) in one space dimension with no-flux boundary conditions. A CO nucleus was put in the initial conditions on a reactive surface. (a) For  $\gamma = 10^3 \text{ s}^{-1}$  pulses form and propagate, but finally die due to rising temperature. (b) For  $\gamma = 10 \text{ s}^{-1}$  pulses can still be formed but die quickly. The parameters are  $p_u = 5.0 \times 10^{-3} \text{ mbar}$ ,  $p_v = 1.55 \times 10^{-2} \text{ mbar}$ ,  $D_u = 10^{-12} \text{ m}^2/\text{s}$ , and  $\bar{T} = 520 \text{ K}$ . In the depicted simulations heat conduction was chosen unrealistically slow, in order to visualise where heat production takes place. The effect of nucleation and propagation failure, however, persists even with realistic heat diffusion.

on order of magnitude smaller than those observed experimentally.

The role of thermal effects can be analysed using  $\gamma$  as bifurcation parameter (cf. Eq. (1.18d)), since for large  $\gamma$  and not too high reaction rates the system would remain isothermal. Thermokinetic effects are a consequence of the asymmetric inhibition of adsorption and the strong temperature dependence of CO desorption. A reactive surface with relatively high oxygen coverage exhibits a high reaction rate and therefore becomes hot, whereas a high CO coverage keeps the catalyst cool. Since, in turn, a lower temperature favours a high CO coverage through reduced desorption, the effect is autocatalytic.

A partial bifurcation analysis of the thermo-kinetic model without CO diffusion is reproduced in Fig. 4.3. For large  $\gamma$  and  $p_u$  close to the Hopf bifurcation CO pulses propagate on the O-covered surface for relatively long times. With decreasing  $\gamma$  this bifurcation shifts to slightly higher  $p_u$ , which moves the O-covered branch away from the region of excitability such that the pulses shrink faster (see Fig. 4.1). From a physical point of view this can be readily explained by temperature effects due to changes in the reaction rate. The rate drops sharply

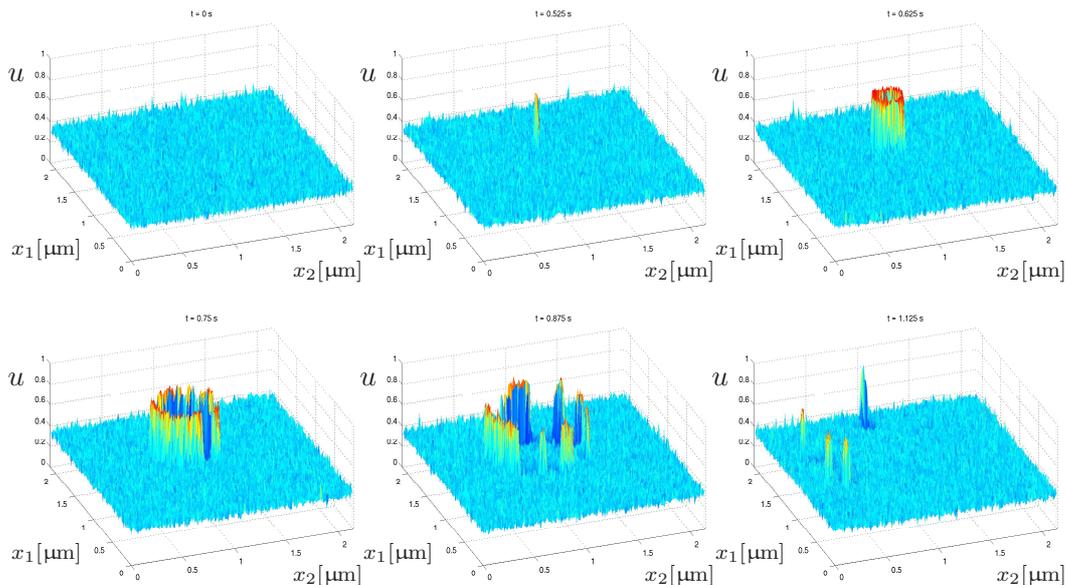


Figure 4.2: Two-dimensional stochastic simulation of raindrop patterns on a reactive Pt(110) surface using the stochastic model with periodic boundary conditions and thermo-kinetic effects included (cf. Fig. 4.1a). Spontaneous nucleation of a CO pulse occurs due to coverage fluctuations. The simulation was performed with  $200 \times 200$  cells,  $n = 10^3$  adsorption sites per cell, and  $p_u = 5.22 \times 10^{-3}$  mbar,  $p_v = 1.55 \times 10^{-2}$  mbar, CO diffusion corresponding to  $D_u = 1.4 \times 10^{-14}$  m<sup>2</sup>/s,  $\bar{T} = 520$  K,  $\gamma = 10^3$  s<sup>-1</sup>.

on predominantly CO-covered areas because oxygen adsorption is blocked there. Behind the CO pulses, however, the reconstruction has been lifted and the reaction rate increases to values even higher than on the original O-covered  $1 \times 2$  surface because of the higher sticking coefficient of oxygen on the  $1 \times 1$  surface. Consequently, the temperature locally rises to values even higher than at the beginning. Since heat conduction is fast, the CO pulses are overrun from the inside (because a hotter surface cannot maintain a high CO coverage due to increased desorption). In contrast, for slightly higher CO pressure ( $p_u \gtrsim 5.5 \times 10^{-3}$  mbar) the whole surface ends up in the CO-covered (cooler) state.

The rate at which critical nuclei are formed on an ideal surface can heuristically be estimated from the theory of large-deviations (see, e.g., (Durrett, 1996)). We assume that in an equilibrium situation the sites in a patch of critical size  $n_{cr}$  are CO-covered independently with probability  $\bar{u}$ , where  $\bar{u}$  is the deterministic equilibrium CO coverage. We speak of a critical nucleus if the CO coverage in the patch is greater than a critical coverage  $u^*$ . For given  $\bar{u}$ ,  $u^*$  and  $n_{cr}$  we can calculate an approximate value for the probability that the coverage is greater than or equal to  $u^*$ , assuming that  $n_{cr}$  is not too small. This probability can be interpreted as the fraction of time the patch has CO coverage  $u \geq u^*$ . It is given by

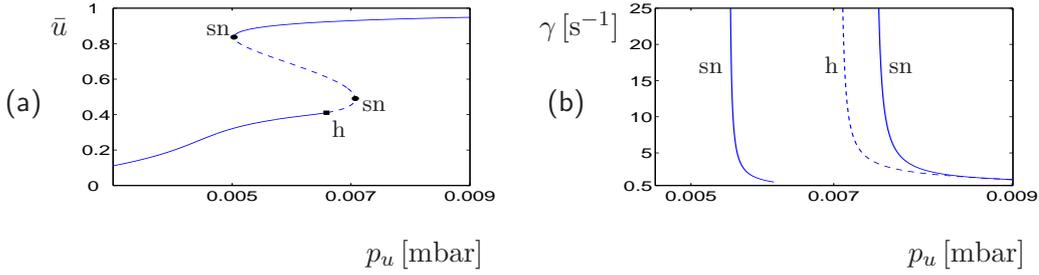


Figure 4.3: (a) Continuation of equilibrium CO coverage  $\bar{u}$  in  $p_u$  for the thermo-kinetic model,  $\gamma = 10^3 \text{ s}^{-1}$ . There is a Hopf bifurcation point (h) between two saddle-node bifurcation points (sn). (b) Continuation of Hopf and sn bifurcations in  $p_u$  and  $\gamma$ .

the expression

$$\left( \left( \frac{\bar{u}}{u^*} \right)^{u^*} \left( \frac{1 - \bar{u}}{1 - u^*} \right)^{1 - u^*} \right)^{n_{cr}}. \quad (4.16)$$

The values of  $\bar{u}$  and  $u^*$  can be taken from the null-clines of the deterministic model. The rate of nucleation events can then be estimated by multiplying expression (4.16) with the density of adsorption sites and dividing by the characteristic time scale  $\tau$  (the time required for the impingement of one monolayer, cf. Section 1.3). The resulting function is obviously very sensitive to  $n_{cr}$ , but it also depends crucially on the excitation threshold  $u^* - \bar{u}$ . Reasonable values of  $1 - 100 \text{ mm}^2 \text{ s}^{-1}$  (Rotermund, 1997b) are obtained with  $n_{cr} = 3000 - 4000$ ,  $u^* = 0.41$  and  $0.33 < \bar{u} < 0.35$ . Increasing  $n_{cr}$  by one order of magnitude (which would correspond to a pressure decrease by one order) results in values indistinguishable from zero.

## Discussion

Spatio-temporal pattern formation in CO oxidation on Pt has been studied experimentally over a wide range of parameters. The observed phenomena mostly appear deterministic, except for very small catalyst areas or at sufficiently high pressure where also temperature variations become observable. Stochastic effects at intermediate pressures such as random nucleation can be reproduced in simulations with a mesoscopic stochastic particle model. Obviously, surface inhomogeneities always play a role on real catalysts. Nevertheless, the experimental observations and their close correspondence to a realistic model clearly suggest that at least a significant fraction of the observed nuclei form uniformly distributed over the surface. The presented effect therefore constitutes the first example of mesoscopic pattern formation ( $1 - 100 \text{ }\mu\text{m}$ ) in a surface reaction that is initiated by internal fluctuations and cannot be captured in a deterministic description.

# Bibliography

- Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. (2004), *Essential Cell Biology*, Garland Science, 2nd edn.
- Andrade, R. F. S., Dewel, G. & Borckmans, P. (1989), ‘Modelling of the kinetic oscillations in the CO oxidation on Pt(100),’ *Journal of Chemical Physics*, vol. 91, pp. 2675–2682.
- Aris, R. (1978), *Mathematical modelling techniques*, Pitman, London.
- Arkin, A., Ross, J. & McAdams, H. H. (1998), ‘Stochastic kinetic analysis of developmental pathway bifurcation in phage  $\lambda$ -infected *Escherichia coli* cells,’ *Genetics*, vol. 149, pp. 1633–1648.
- Arnold, L. & Theodosopulu, M. (1980), ‘Deterministic Limit of the Stochastic Model of Chemical Reactions with Diffusion,’ *Advances in Applied Probability*, vol. 12(2), pp. 367–379.
- Auyang, S. Y. (1998), *Foundations of complex-system theories in economics, evolutionary biology, and statistical physics*, Cambridge University Press, Cambridge.
- Bär, M. (1993), ‘Räumliche Strukturbildung bei einer Oberflächenreaktion,’ Dissertation, Freie Universität Berlin.
- Ben-Jacob, E., Cohen, I. & Levine, H. (2000), ‘Cooperative self-organization of microorganisms,’ *Advances in Physics*, vol. 49(4), pp. 395–554.
- Blount, D. (1991), ‘Comparison of deterministic and stochastic models of a linear chemical reaction with diffusion,’ *The Annals of Probability*, vol. 19, pp. 1440–1462.
- Blount, D. (1993), ‘Limit theorems for a sequence of reaction-diffusion systems,’ *Stochastic Processes and their Applications*, vol. 42, pp. 1–30.
- Blount, D. (1994), ‘Density-dependent limits for a nonlinear reaction-diffusion model,’ *The Annals of Probability*, vol. 22(4), pp. 2040–2070.
- Brémaud, P. (1999), *Markov Chains: Gibbs fields, Monte Carlo simulation, and queues*, Springer, New York.

- Cross, M. C. & Hohenberg, P. C. (1993), 'Pattern formation outside of equilibrium,' *Reviews of Modern Physics*, vol. 65(3), pp. 851–1112.
- Da Prato, G. & Zabczyk, J. (1992), *Stochastic equations in infinite dimensions*, Cambridge University Press, Cambridge.
- Dayan, P. & Abbott, L. F. (2001), *Theoretical Neuroscience*, The MIT Press, Cambridge, Massachusetts.
- de Masi, A. & Presutti, E. (1991), *Mathematical methods for hydrodynamic limits*, Springer, Berlin.
- Doedel, E., Champneys, A. R., Fairgrieve, T. F., Kusnetsov, Y. A., Sandstede, B. & Wang, X. (1998), *AUTO 97: Continuation and Bifurcation Software for Ordinary Differential Equations*, <http://ftp.cs.concordia.ca/pub/doedel/auto>.
- Durrett, R. (1996), *Probability: Theory and Examples*, Wadsworth, Belmont, 2nd edn.
- Durrett, R. (1999), 'Stochastic spatial models,' *SIAM Review*, vol. 41(4), pp. 677–718.
- Eiswirth, M. (1987), 'Phänomene der Selbstorganisation bei der Oxidation von CO an Pt(110),' Dissertation, Ludwig-Maximilians-Universität München.
- Eiswirth, M. & Ertl, G. (1995), 'Pattern formation on Catalytic Surfaces,' in R. Kapral & K. Showalter (Eds.), 'Chemical Waves and Patterns,' pp. 447–483, Kluwer, Dordrecht.
- Ermentrout, G. B. (1998), 'Neural networks as spatio-temporal pattern-forming systems,' *Reports on Progress in Physics*, vol. 61, pp. 353–430.
- Ertl, G. (1991), 'Oscillatory Kinetics and Spatio-Temporal Self-Organization in Reactions at Solid Surfaces,' *Science*, vol. 254, pp. 1750–1755.
- Ethier, S. N. & Kurtz, T. G. (1986), *Markov Processes: Characterization and Convergence*, Wiley, New York.
- Evans, L. C. (1998), *Partial differential equations*, AMS, Providence.
- Friedman, A. (1982), *Variational Principles and Free Boundary Problems*, Wiley, New York.
- Gardiner, C. W. (2004), *Handbook of Stochastic Methods*, Springer, Berlin, 3rd edn.
- Gibson, M. A. & Bruck, J. (2000), 'Efficient exact simulation of chemical systems with many species and many channels,' *Journal of Physical Chemistry A*, vol. 104, pp. 1876 – 1889.
- Gillespie, D. T. (1977), 'Exact Stochastic Simulation of Coupled Chemical Reactions,' *Journal of Physical Chemistry*, vol. 81, pp. 2340–2361.

- Gillespie, D. T. (2001), 'Approximate accelerated stochastic simulation of chemically reacting systems,' *Journal of Chemical Physics*, vol. 115(4), pp. 1716–1733.
- Guckenheimer, J. (1986), 'Multiple Bifurcation Problems for Chemical Reactors,' *Physica D*, vol. 20(1), pp. 1–20.
- Guias, F. (2002), 'Mesoscopic models of reaction-diffusion processes with exclusion mechanism,' in N. Antonic (Ed.), 'Multiscale problems in science and technology,' pp. 161–173, Springer, Berlin.
- Haken, H. (1983), *Synergetics: An Introduction*, Springer, Berlin.
- Hale, J. & Koçak, H. (1991), *Dynamics and Bifurcations*, Springer, New York.
- Hayes, R. E. & Kolaczkowski, S. T. (1997), *Introduction to catalytic combustion*, Gordon & Breach, Amsterdam.
- Hopkinson, A., Bradley, J., Guo, X.-C. & King, D. A. (1993), 'Nonlinear Island Growth Dynamics in Adsorbate-Induced Restructuring of Quasihexagonal Reconstructed Pt(100) by CO,' *Physical Review Letters*, vol. 71(10), pp. 1597–1600.
- Hutter, K. & Jöhnk, K. (2004), *Continuum methods of physical modelling*, Springer, Berlin.
- Imbihl, R. (2005), 'Nonlinear dynamics on catalytic surfaces,' *Catalysis Today*, vol. 105, pp. 206–222.
- Imbihl, R., Cox, M. P., Ertl, G., Müller, H. & Brenig, W. (1985), 'Kinetic Oscillations in the Catalytic CO Oxidation on Pt(100): Theory,' *Journal of Chemical Physics*, vol. 83(4), pp. 1578–1587.
- Imbihl, R. & Ertl, G. (1995), 'Oscillatory Kinetics in Heterogeneous Catalysis,' *Chemical Reviews*, vol. 95(3), pp. 697–733.
- Jakubith, S., Rotermund, H. H., Engel, W., von Oertzen, A. & Ertl, G. (1990), 'Spatiotemporal concentration patterns in a surface reaction: propagating and standing waves, rotating spirals, and turbulence,' *Physical Review Letters*, vol. 65(24), pp. 3013–3016.
- Kallenberg, O. (2002), *Foundations of Modern Probability*, Springer, 2nd edn.
- Karlin, S. & Taylor, H. M. (1975), *A first course in stochastic processes*, Academic Press, London, 2nd edn.
- Kipnis, C. & Landim, C. (1999), *Scaling limits of interacting particle systems*, Springer, Berlin.
- Kotelenez, P. (1986), 'Law of large numbers and central limit theorem for linear chemical reactions with diffusion,' *The Annals of Probability*, vol. 14(1), pp. 173–193.

- Kotelenez, P. (1988), 'High density limit theorems for nonlinear chemical reactions with diffusion,' *Probability Theory and Related Fields*, vol. 78, pp. 11–37.
- Krischer, K., Eiswirth, M. & Ertl, G. (1992), 'Oscillatory CO oxidation on Pt(110): Modelling of temporal self-organization,' *Journal of Chemical Physics*, vol. 96(12), pp. 9162–9172.
- Krömker, S. (1997), 'Model and Analysis of Heterogeneous Catalysis with Phase Transition,' Dissertation, Universität Heidelberg.
- Kurtz, T. G. (1981), *Approximation of Population Processes*, SIAM, Philadelphia.
- Liggett, T. M. (2005), *Interacting Particle Systems*, Springer, Berlin.
- Lions, J. L. (1969), *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Paris.
- Mikhailov, A. S. (1994), *Foundations of Synergetics*, Springer, Berlin, 2nd edn.
- Murray, J. D. (2004a), *Mathematical biology 1*, Springer, Berlin, 3rd edn.
- Murray, J. D. (2004b), *Mathematical biology 2*, Springer, Berlin, 3rd edn.
- Naumovets, A. G. (2005), 'Collective surface diffusion: An experimentalist's view,' *Physica A*, vol. 357, pp. 189–215.
- Nicolis, G. & Prigogine, I. (1977), *Self-organization in Nonequilibrium Systems*, Wiley, New York.
- Oelschläger, K. (1989), 'On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes,' *Probability Theory and Related Fields*, vol. 82, pp. 565–586.
- Oura, K., Lifshits, V. G., Saranin, A. A., Zotov, A. V. & Katayama, M. (2003), *Surface Science: An Introduction*, Springer, Berlin.
- Press, T., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992), *Numerical Recipes in C*, Cambridge University Press, Cambridge, 2nd edn.
- Raviart, P. A. (1967), 'Sur l'approximation de certaines équations d'évolution linéaires et non linéaires,' *Journal de Mathématiques Pures et Appliquées*, vol. 46, pp. 11–107.
- Reichert, C. (2000), 'Ein mesoskopisches stochastisches Modell für die CO-Oxidation an Platinoberflächen,' Diplomarbeit, Universität Heidelberg.
- Reichert, C., Starke, J. & Eiswirth, M. (2001), 'Stochastic model of CO oxidation on platinum surfaces and deterministic limit,' *Journal of Chemical Physics*, vol. 115(10), pp. 4829 – 4838.

- Revuz, D. & Yor, M. (2005), *Continuous Martingales and Brownian Motion*, Springer, Berlin, 3rd edn.
- Roland, P. E. (1993), *Brain Activation*, Wiley, New York.
- Rosé, H., Hempel, H. & Schimansky-Geier, L. (1994), 'Stochastic dynamics of catalytic CO oxidation on Pt(100),' *Physica A*, vol. 206, pp. 421–440.
- Rotermund, H. H. (1997a), 'Imaging of dynamic processes on surfaces by light,' *Surface Science Reports*, vol. 29, pp. 265–364.
- Rotermund, H. H. (1997b), 'Imaging pattern formation in surface reactions from ultra-high vacuum to atmospheric pressures,' *Surface Science*, vol. 386, pp. 10–23.
- Rotermund, H. H., Haas, G., Franz, R. U., Tromp, R. M. & Ertl, G. (1995), 'Imaging pattern formation in surface reactions from Ultra High Vacuum up to atmospheric pressures,' *Science*, vol. 270, pp. 608–610.
- Sachs, C., Hildebrand, M., S., V., Winterlin, J. & Ertl, G. (2001), 'Spatiotemporal self-organization in a surface reaction: from the atomic to the mesoscopic scale,' *Science*, vol. 293, pp. 1635–1638.
- Smoller, J. (1983), *Shock waves and reaction-diffusion equations*, Springer, Berlin.
- Snow, C. D., Sorin, E. J., Rhee, Y. M. & Pande, V. S. (2005), 'How well can simulation predict protein folding kinetics and thermodynamics,' *Annual Review of Biophysics and Biomolecular Structure*, vol. 34, pp. 43–69.
- Starke, J., Reichert, C., Eiswirth, M. & Oelschläger, K. (2006), 'Stochastic modelling and deterministic limit of catalytic surface processes,' in W. Jäger, R. Rannacher & J. Warnatz (Eds.), 'Reactive Flow, Diffusion and Transport,' Springer, Berlin.
- Stroock, D. W. & Varadhan, S. R. S. (1979), *Multidimensional diffusion processes*, Springer, Berlin.
- Suchorski, Y., Beben, J., Imbihl, R., James, E. W., Liu, D.-J. & Evans, J. W. (2001), 'Fluctuations and Critical Phenomena in Catalytic CO-Oxidation on Pt Facets,' *Physical Review B*, vol. 63(16), p. 165417.
- Suchorski, Y., Beben, J., James, E. W., Evans, J. W. & Imbihl, R. (1999), 'Fluctuation-induced transitions in a bistable surface reaction: Catalytic CO oxidation on a Pt field emitter tip,' *Physical Review Letters*, vol. 82(9), pp. 1907–1910.
- Temam, R. (1973), *Numerical Analysis*, D. Reidel, Dordrecht.

- Temam, R. (2001), *Navier-Stokes Equations: Theory and Numerical Analysis*, AMS, Providence.
- Thomas, J. M. & Thomas, W. J. (1996), *Principles and Practice of Heterogeneous Catalysis*, VCH, Weinheim.
- van Kampen, N. (1992), *Stochastic Processes in Physics and Chemistry*, Elsevier, Amsterdam, 2nd edn.
- von Bertalanffy, L. (1968), *General System Theory*, G. Braziller, New York.
- von Oertzen, A., Rotermund, H. H., Mikhailov, A. S. & Ertl, G. (2000), 'Standing Wave Patterns in the CO Oxidation Reaction on a Pt(110) Surface: Experiments and Modeling,' *Journal of Physical Chemistry B*, vol. 104, pp. 3155–3178.
- Wiggins, S. (1990), *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer, New York.
- Zangwill, A. (1988), *Physics at Surfaces*, Cambridge University Press, Cambridge.
- Zeidler, E. (1990a), *Nonlinear Functional Analysis and its Applications*, vol. I, Springer, Berlin.
- Zeidler, E. (1990b), *Nonlinear Functional Analysis and its Applications*, vol. II/A, Springer, Berlin.
- Zeidler, E. (1990c), *Nonlinear Functional Analysis and its Applications*, vol. II/B, Springer, Berlin.
- Zhabotinsky, A. M. (1991), 'A history of chemical oscillations and waves,' *Chaos*, vol. 1(4), pp. 379–386.
- Ziff, R. M., Gulari, E. & Barshad, Y. (1986), 'Kinetic Phase Transitions in an Irreversible Surface-Reaction Model,' *Physical Review Letters*, vol. 56(24), p. 2553.

# List of Figures

1.1	Reconstructed and bulk-like surfaces of platinum. . . . .	16
1.2	Slow sinusoidal oscillations with large amplitude. . . . .	19
1.3	Relaxation oscillations. . . . .	19
1.4	Fast sinusoidal oscillations with small amplitude and period doubling . . . . .	19
1.5	Experimental bifurcation diagram for CO oxidation on Pt(110). . . . .	20
1.6	Nucleation of a CO island and spiral waves . . . . .	20
1.7	Target patterns and standing waves . . . . .	20
1.8	EMSI pictures of Pt surface showing raindrop patterns. . . . .	22
1.9	Formation of a single raindrop. . . . .	22
1.10	Computed bifurcation diagram for Pt(110) at 500–560 K . . . . .	29
1.11	Range of validity of different types of models. . . . .	34
4.1	Simulation of pulse propagation with thermo-kinetic model equations. . . . .	124
4.2	Two-dimensional stochastic simulation of raindrop patterns. . . . .	125
4.3	Basic bifurcation diagram at parameters where raindrops are formed. . . . .	126



# List of Tables

1.1	Reaction steps of CO oxidation on Pt at low to intermediate pressures . . . .	23
1.2	Rate terms for adsorption and desorption of CO and oxygen. . . . .	24
1.3	Rate term for reaction. . . . .	25
1.4	Parameters for surface reconstruction. . . . .	26
1.5	Additional important constants. . . . .	26
1.6	Commonly used denotations of bifurcations and their abbreviations. . . . .	27
1.7	Parameters for heat production through chemical processes. . . . .	30
1.8	Parameters for heat transfer by radiation. . . . .	30
1.9	Parameters for heat conduction. . . . .	31
1.10	Additional parameters for the energy balance. . . . .	31
1.11	Jumps caused by reaction events . . . . .	35

# Index

- adsorption
  - of O<sub>2</sub>, 16
  - of CO, 16
- adsorption rate
  - of O<sub>2</sub>, 24
  - of CO, 23
- bifurcation, 5
- bifurcation diagram for Pt(110), 27
- complex system
  - definition, 1
  - examples, 2
- desorption of CO, 17
- desorption rate of CO, 24
- discrete
  - error, 60
  - compactness theorem, 66, 93
  - derivative, 56
  - integration by parts, 57
  - Lebesgue space, 51, 89
  - Poincaré inequality, 57
  - Sobolev space, 57, 90
  - version of Rellich's theorem, 66
- dynamical system, 3
- Dynkin's formula, 49
- EMSI, 15
- error, 60
- extension operator, 60
  - $J_h$ , 63
  - $P_h$ , 62
  - $\bar{J}_h$ , 93
  - $\bar{P}_h$ , 92
- stability, 60
- external approximation, 60
  - convergence, 60
  - of  $H^1(G)$  by  $\mathcal{H}^1(\bar{\mathcal{G}}_h)$ , 93
  - of  $H_0^1(G)$  by  $\mathcal{H}_0^1(\mathcal{G}_h)$ , 63
  - stability, 60
- Feller semigroup, 49
- filtration, 49
- heat conduction, 30
- heat production by chemical reactions, 28
- heat transfer by radiation, 30
- internal approximation, 60
  - of  $L^2(G)$  by  $\mathcal{L}^2(\bar{\mathcal{G}}_h)$ , 92
  - of  $L^2(G)$  by  $\mathcal{L}^2(\mathcal{G}_h)$ , 63
- Langmuir-Hinshelwood mechanism, 13
- lattice function, 50
  - extended version, 50
- lattice points
  - $\bar{\mathcal{G}}_h$ , 89
  - $\bar{\mathcal{G}}_h^1$ , 89
  - $\mathcal{G}_h$ , 50
  - $\mathcal{G}_h^1$ , 52
- law of large numbers
  - crowding effects, 104
  - example, 67
  - general linear, 75
  - gradient-activated diffusion, 112
  - invariant regions, 95
  - Lipschitz conditions, 86

LEED, 14

Markov property, 49

martingale, 49

mathematical model, 3

mesoscopic stochastic particle model

- CO oxidation on Pt, 33
- crowding effects, 101
- example, 47
- general linear, 51
- gradient-activated diffusion, 110
- invariant regions, 90
- Lipschitz conditions, 84

oscillations in CO oxidation on Pt, 18

pattern formation, 2

- CO oxidation on Pt, 18

PEEM, 14

raindrop patterns

- experimental, 21
- nucleation rate, 125
- role of thermal effects, 124
- simulation, 123

rate of phase change, 25

reaction of CO and O, 18

reaction rate, 25

refined law of large numbers, 115

restriction operator, 60

- $Q_h$ , 61
- $R_h$ , 62
- $\bar{R}_h$ , 92
- $\bar{Q}_h$ , 92
- stability, 60

simple exclusion process, 17

stochastic process, 3

stopping time, 49

surface diffusion, 17

surface reconstruction, 15

truncation error, 60

UHV chamber, 14



‘[...] before her was another long passage, and the White Rabbit was still in sight, hurrying down it. There was not a moment to be lost: away went Alice like the wind, and was just in time to hear it say, as it turned a corner, “Oh my ears and whiskers, how late it’s getting!” She was close behind it when she turned the corner, but the Rabbit was no longer to be seen [...].’

Lewis Carroll

## Acknowledgements

First of all I thank Prof. Willi Jäger for his willingness to supervise such a truly interdisciplinary research project, for his confidence, and for his constant support.

I thank my co-operation partners Dr. Jens Starke, who has now left to DTU in Copenhagen, and Dr. Markus Eiswirth from the Fritz-Haber-Institut in Berlin for many fruitful discussions about modelling CO oxidation on Pt. A special thanks goes to Markus for hosting Jens and me in his office at home (full board) during a one-week stay at Fritz-Haber in June 2002.

Furthermore, I am indebted to Dr. Karl Oelschläger and Dr. Mariya Ptashnyk, who carefully read draft versions of the mathematical parts of the thesis, for useful comments and suggestions. (Of course they are not responsible for the remaining mistakes.)

I acknowledge financial support from the Deutsche Forschungsgemeinschaft in the form of a scholarship from the Graduiertenkolleg ‘Complex Processes’ at IWR in Heidelberg.

I thank all my colleagues from AMJ for the good times we had together, especially during several compact seminars at Fehrenbacher Hof (Black Forest) and Wallenfels, and for the friendly atmosphere at the institute. I thank Jan for sharing the office with me (‘Morgen, Ritchie!’ – ‘Hm.’).

Ich danke meinen Eltern für ihre Unterstützung während meiner Studienzeit und dafür, daß sie während der Doktorarbeit (fast) keine Fragen gestellt haben!

Finalemment je dis un énorme merci à mes deux filles, Sév et Léonie, qui m’ont toujours encouragé, et que j’aime très fort.