

Freie Universität



Berlin

A dissipative finite volume scheme for reaction-diffusion systems in heterogeneous materials

Dissertation zur Erlangung des Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)
am Fachbereich Mathematik und Informatik
der Freien Universität Berlin

von

André Fiebach

Berlin

18. Dezember 2013

1. Gutachter: Prof. Dr. Volker John, *Freie Universität Berlin und Weierstraß-Institut, Berlin*
2. Gutachter: Prof. Dr. Danielle Hilhorst, *Université de Paris-Sud, Paris*

Tag der Disputation: 12 Juni 2014
Erscheinungsjahr: 2014

Preface

The thesis collects the fruits of my work in the research group "Numerical Mathematics and Scientific Computing" at the Weierstrass Institute in Berlin during the last years. In a project of the European Commission within the *Seventh Framework Programme (FP7) MD³ "Material Development for Double Exposure and Double Patterning"* the first results [39–41] were achieved together with Dr. Jürgen Fuhrmann. After that time the main results of Chapter 3 were attained in collaboration with Priv.-Doz. Dr. Annegret Glitzky and Dr. Alexander Linke. Also during my time at the Weierstrass Institute, two software modules were created. The first one, the pdelib2 based module WiasPeb was used to obtain the results in the MD₃ project, see Section 1.3. The second one, building the base for the results in Chapter 5, was created in cooperation with Dr. Klaus Gärtner.

Finally, I would like to express my gratitude to all my colleagues from the Weierstrass Institute for providing an excellent and friendly working environment.

My special thanks go to three persons. Without their engagement this work would not have been possible at all. Thanks to Priv.-Doz. Dr. Annegret Glitzky I gained valuable experience in the analytical treatment of reaction-diffusion systems. Her constructive criticism and her continuous helpfulness contributed to the creation of this thesis. Further, I would like to thank Dr. Klaus Gärtner for his permanent interest in all scientific and nonscientific problems. I could participate in his knowledge in scientific computing and his valuable advices influenced the implementation and the debugging of the program. Moreover, I greatly benefited from the numerous valuable remarks concerning the analytical treatment of the discretized equations of Dr. Alexander Linke. His continuous interest in the topic and his cooperativeness were a great support.

I have to express my gratitude to Dr. Jürgen Fuhrmann and Prof. Dr. Volker John, giving me the opportunity to work under their management on this topic in a project of the European Commission within the *Seventh Framework Programme (FP7) MD³ "Material Development for Double Exposure and Double Patterning"*.

I wish to thank Dr. Jens André Griepentrog for a lot of interesting discussions and interesting remarks. Last but not least, I would like to thank Prof. Dr. Danielle Hilhorst for coexamining this thesis and for helpful comments improving the manuscript.

Parts of the thesis have been previously published in

- [30] A. Erdmann, F. Shao, J. Fuhrmann, A. Fiebach, G. P. Patis, and P. Trefonas. “Modeling of double patterning interactions in litho-cure-litho-etch (LCLE) processes”. In: *Proc. of SPIE* 7640.1 (2010), 76400B.
- [35] A. Fiebach, A. Glitzky, and A. Linke. *Convergence of an implicit Voronoi finite volume method for reaction-diffusion problems*. Preprint 1827. WIAS, 2013.
- [36] A. Fiebach, A. Glitzky, and A. Linke. “Uniform global bounds for solutions of an implicit Voronoi finite volume method for reaction–diffusion problems”. In: *Numer. Math.* 128.1 (2014), pp. 31–72.
- [39] J. Fuhrmann, A. Fiebach, A. Erdmann, and P. Trefonas. “Acid diffusion effects between resists in freezing processes used for contact hole patterning”. In: *Microelectron. Eng.* 87.5–8 (2010), pp. 951–954.
- [40] J. Fuhrmann, A. Fiebach, and G. P. Patis. “Macroscopic and stochastic modeling approaches to pattern doubling by acid catalyzed cross-linking”. In: *Proc. of SPIE* 7639, 76392I (2010), p. 76392I.
- [41] J. Fuhrmann, A. Fiebach, M. Uhle, A. Erdmann, C. R. Szmanda, and C. Truong. “A model of self-limiting residual acid diffusion for pattern doubling”. In: *Microelectron. Eng.* 86.4–6 (2009), pp. 792–795.

The final publication of [36] is available on link.springer.com. The final publications of [39, 41] are available on www.sciencedirect.com. The final publications of [30, 40] are available on proceedings.spiedigitallibrary.org.

Contents

Notation	v
1 Introduction	1
1.1 Zusammenfassung	1
1.2 Résumé	4
1.3 Application in photolithography	6
1.3.1 Double patterning	9
1.3.2 Pattern doubling	11
1.4 Formulation of model equations	15
1.5 Relevant literature	18
2 The continuous problem	20
2.1 General assumptions on the data	20
2.2 Weak formulation	21
2.3 Summary of known results	22
3 The discretized problem	27
3.1 Introduction	27
3.2 Voronoi finite volume discretization	28
3.3 Time discretization	31
3.4 Invariants of the system	33
3.5 Local existence	34
3.6 Physically motivated estimates	37
3.7 Thermodynamic equilibrium	40
3.8 Global upper bounds	42
3.9 Asymptotics	50
3.10 Global lower bounds	59
3.11 Global existence and uniqueness	67

3.12	Convergence	69
3.12.1	Introduction	69
3.12.2	Prolongated quantities	70
3.12.3	A priori estimates	71
3.12.4	Weak solution	74
3.13	Comments on 3D	81
4	Implementation of the scheme	83
4.1	Solution of the nonlinear equations	83
4.2	Solution of linear equations	90
4.3	Adaptive time stepping	90
5	Numerical Example	93
5.1	Introduction	93
5.2	Homogeneous material	95
5.3	Layer behavior	98
5.4	Moving front	102
5.5	Nested Circles – Order of convergence	109
5.5.1	Uniformly refined mesh	109
5.5.2	Aligned mesh region	117
5.6	Tiles	120
5.7	Daisy	123
5.8	Instantaneous reaction	127
5.9	Conclusion	134
A	Appendix	136
A.1	Symmetric reaction coefficients	136
A.2	Discrete Gagliardo-Nirenberg inequality	139
A.2.1	1. Approach – restricted to 2d	139
A.2.2	2. Approach	146
A.3	Technical lemmas	149
A.3.1	Global existence proof	149
A.3.2	Convergence proof	151
A.4	Convex analysis	157
	Bibliography	158

Notation and Conventions

We use the following notations:

Sets:

\emptyset	Empty set
\mathbb{N}	Set of all natural numbers
\mathbb{N}_0	$\mathbb{N} \cup \{0\}$
\mathbb{R}	Set of all real numbers
\mathbb{R}_+	$\{x \in \mathbb{R} \mid x \geq 0\}$
\mathbb{R}_-	$\{x \in \mathbb{R} \mid x \leq 0\}$
$\text{int } M$	Interior of a set $M \subset \mathbb{R}^n$
\overline{M}	Closure of a set $M \subset \mathbb{R}^n$
∂M	Boundary of a set $M \subset \mathbb{R}^n$
$ M $	Lebesgue measure of $M \subset \mathbb{R}^n$
$\text{span } M$	Linear hull of the set $M \subset \mathbb{R}^n$
$\text{diam } M$	Diameter of a set $M \subset \mathbb{R}^n$

Domain:

$\Omega \subset \mathbb{R}^2$	Bounded, polygonal, Lipschitzian domain
Γ	Boundary of Ω
\boldsymbol{n}	Outer normal vector

Functions and mappings:

f^+	Positive part of a function f , $f^+ = \sup\{f, 0\}$
f^-	Negative part of a function f , $f^- = \sup\{-f, 0\}$
$ f $	Absolute value of a function f , $ f = f^+ + f^-$
$f \circ g$	Composition of the functions f and g
$f _M$	Restriction of f to the set M
$\text{dom}(f)$	Domain of a function f

Function spaces:

$C(\Omega, \mathbb{R}^m)$	Space of the continuous \mathbb{R}^m -valued functions on Ω
$X := H^1(\Omega, \mathbb{R}^m)$	Space of all H^1 functions with values in \mathbb{R}^m
$Y := L^2(\Omega, \mathbb{R}^m)$	Space of \mathbb{R}^m -valued square-integrable functions on Ω
$W := X \cap L^\infty(\Omega, \mathbb{R}^m)$	
X^*	Dual of X
\hookrightarrow	Compact embedding
$\langle \cdot, \cdot \rangle_X$	Dual pairing in X

Derivatives:

∇f	Gradient of a function f
$\nabla \cdot h$	Divergence of a vector field h
$\partial_t f$	Partial derivative of a function f with respect to t
$\partial_n f$	Normal derivative
$\partial F(u)$	Subdifferential of the Functional F at u

Quantities in the state equations:

m	Number of species
X_ν	ν -th species
u_ν	Density of the ν -th species
U_ν	Initial density of the ν -th species
\bar{u}_ν	Reference density of the ν -th species
v_ν	Chemical potential of the ν -th species
a_ν	Chemical activity of the ν -th species
D_ν	Diffusion coefficient of the ν -th species
\mathcal{R}	Finite set of reactions
α, β	Stoichiometric coefficients
$R_{(\alpha, \beta)}$	Reaction rates
$k_\alpha, k_\beta, k_{(\alpha, \beta)}$	Kinetic coefficients
\mathcal{S}	Stoichiometric subspace
\mathcal{S}^\perp	orthogonal complement in \mathbb{R}^m of \mathcal{S}
\mathcal{A}	Set of all equilibriums in a suitable parametrization
$F(u(t)), \Psi(u(t))$	Free energy, relative free energy at state $u(t)$
$D(v(t))$	Dissipation rate of the system at state $v(t)$

Discretized quantities (Sec. 3)

$\mathcal{P}, \mathcal{V}, \mathcal{E}, \mathcal{T}$	Family of points, Voronoi volumes, edges, triangles
$\mathcal{M} = (\mathcal{P}, \mathcal{V}, \mathcal{E})$	Voronoi mesh
$\mathcal{N}_\mathcal{V}(K)$	Neighboring volumes of $K \in \mathcal{V}$
\mathcal{E}_{int}	Set of all interior faces
f_K	Value of f in the Voronoi volume $K \in \mathcal{V}$
$ K $	Lebesgue measure of $K \in \mathcal{V}$
$\widehat{F}(\mathbf{u}), \widehat{\Psi}(\mathbf{u})$	Discrete free energy and relative free energy at state \mathbf{u}
$\widehat{D}(\mathbf{u})$	Discrete dissipation rate at state \mathbf{u}
f_h	Piecewise constant functions on \mathcal{M}
f_d	Piecewise constant functions on Donald boxes
f_l	Piecewise linear functions on the triangles

Abbreviations

Fig.	Figure
resp.	respectively
i.e.	id est
f.a.a.	for almost all
w.l.o.g.	without loss of generality
p.	page
e.g.	exempli gratia
a.e.	almost every

Other conventions

We use the following notations and conventions which are not listed before. In some sections we introduce specific symbols which are not listed here. In the work, $\Omega \subset \mathbb{R}^2$ is a bounded polygonal Lipschitz domain, unless otherwise stated. As long as the meaning is clear we shortly write L^p instead of $L^p(\Omega)$ and H^1 instead of $H^1(\Omega)$.

In the treatment of evolution equations we need the following spaces: Let V be a Banach space and S a unbounded resp. bounded interval of \mathbb{R}_+ . We denote by $L_{\text{loc}}^p(S, V)$ resp. $L^p(S, V)$, $p \in [1, \infty]$, the space of all equivalence classes of all Bochner measurable functions $u : S \rightarrow V$ such that $\|u(\cdot)\|_V \in L_{\text{loc}}^p(S)$ resp. $\|u(\cdot)\|_V \in L^p(S)$. The spaces will be equipped with their usual norm resp. seminorm. The space $H_{\text{loc}}^1(S, V)$ is defined as the space of all $u \in L_{\text{loc}}^2(S, V)$ such that $u' \in L_{\text{loc}}^2(S, V)$, where u' denotes the derivative of u in the sense of V -valued distributions. The space $H^1(S, V)$ is defined analogously.

We set $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$ for $x \in \mathbb{R}$ the positive and negative part of x and mention the following relationships

$$\begin{aligned} x &= x^+ - x^-, & |x| &= x^+ + x^-, & x^+ &= \frac{|x| + x}{2}, & x^- &= \frac{|x| - x}{2}, \\ x &\leq x^+, & x &\geq -x^-, & x^+ x^- &= 0. \end{aligned}$$

Concerning vectors $\mathbf{v} \in \mathbb{R}^m$ we use: Writing $\mathbf{v} \geq \mathbf{0}$ resp. $\mathbf{v} > \mathbf{0}$ we mean $v_i \geq 0$ resp. $v_i > 0$ for $i = 1, \dots, m$. By $\ln \mathbf{v}$ we denote $(\ln v_i)_{i=1}^m$, by $\mathbf{e}^{\mathbf{v}}$ the vector $(e^{v_i})_{i=1}^m$ and by \mathbf{v}^α with $\alpha \in \mathbb{Q}$ we understand $(v_i^\alpha)_{i=1}^m$. For $(x_i y_i)_{i=1}^m$ resp. $(x_i / y_i)_{i=1}^m$ we shortly write $\mathbf{x} \mathbf{y}$ resp. \mathbf{x} / \mathbf{y} . The usual scalar product in \mathbb{R}^m is denoted by $\mathbf{x}^T \mathbf{y}$. Together with $\alpha \in \mathbb{Z}_+^m$ the symbol \mathbf{v}^α means $\prod_{i=1}^m v_i^{\alpha_i}$.

The zero matrix in $\mathbb{R}^{m \times m}$ is denoted by 0_m and the identity matrix by I_m . Under $[\mathbf{v}]$ for $\mathbf{v} \in \mathbb{R}^m$ we understand the diagonal matrix $D \in \mathbb{R}^{m \times m}$ with elements $d_{i,i} = v_i$, $i = 1, \dots, m$, on the main diagonal. The rank of a matrix $A \in \mathbb{R}^{m \times m}$ is denoted by $\text{rk } A$. The transpose and inverse of a matrix

$A \in \mathbb{R}^{m \times m}$ is denoted with A^T and A^{-1} resp. The Kronecker product and the direct sum of two matrices $A, B \in \mathbb{R}^{m \times m}$ is denoted by $A \otimes B$ and $A \oplus B$ resp.

In all proofs, real constants c with different meaning are numbered consecutively. The constants only depend on the data and not on the discretization, unless otherwise stated. Moreover the symbols S_1, S_2 , and S_3 have a local meaning and differ from time to time. Generally, these terms arise from testing the problem by test functions and discussing the expressions for the time-derivative (S_1), diffusion term (S_2), and reaction term (S_3) separately. As long as the meaning is clear we neglect the actual time step t_n and we shortly write $u_{\nu K}$ instead of $u_{\nu K}(t_n)$.

Introduction

1.1 Zusammenfassung

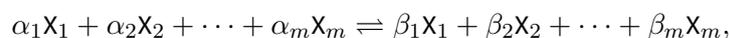
Die Modellierung von Prozessen der chemischen Verfahrenstechnik, wie sie zum Beispiel bei der Herstellung von Halbleitern in der Fotolithografie [95] auftreten, führt auf Reaktions-Diffusionssysteme, die durch im Orte nicht glatte Daten gekennzeichnet sind. Vom besonderem Interesse ist dabei ein chemischer Prozess, der Strukturen für die Halbleiterherstellung ausbildet und in zwei Materialien stattfindet, siehe [41].

Gegenstand der Arbeit sollen Systeme von parabolischen partiellen Differentialgleichungen sein, welche die Umwandlung von chemischen, elektrisch neutralen Spezies in heterogenen Materialien beschreiben, welche durch Diffusion und Reaktionen getrieben werden. Vor diesem Hintergrund sind Schwierigkeiten in der numerischen Analyse zu erwarten, welche durch die starke Kopplung der Gleichungen, die nicht glatte Ortsabhängigkeit der Reaktions- und Diffusionskoeffizienten, sowie das Nichtvorhandensein des Maximumprinzips hervorgerufen werden.

Im Speziellen wird auf einer beschränkten Teilmenge $\Omega \subset \mathbb{R}^2$ ein System von m (elektrisch neutralen) Spezies betrachtet. Die Konzentration einer einzelnen Spezies X_ν sei mit u_ν , ihr chemisches Potential mit v_ν , und ihre Referenzdichte mit \bar{u}_ν , $\nu = 1, \dots, m$, bezeichnet. Der Zusammenhang zwischen der Konzentration einer Spezies und ihrem chemischen Potential sei durch die Boltzmann-Statistik $u_\nu = \bar{u}_\nu e^{v_\nu}$ beschrieben. Das Model besteht aus m Evolutionsgleichungen

$$\mathbf{j}_\nu = -D_\nu u_\nu \nabla v_\nu - \frac{\partial u_\nu}{\partial t} + \nabla \cdot \mathbf{j}_\nu = R_\nu \quad \text{in } \mathbb{R}_+ \times \Omega$$

und den homogenen Neumann-Randbedingungen $\partial_n \mathbf{j}_\nu = 0$ auf $\partial\Omega$, $\nu = 1, \dots, m$. Die Quellterme R_ν resultieren aus Kombinationen der Reaktionsraten von reversiblen chemischen Reaktionen der Form



wobei $\alpha = (\alpha_1, \dots, \alpha_m)$ und $\beta = (\beta_1, \dots, \beta_m)$ die zugehörigen stöchiometrischen Koeffizienten sind. Die endliche Teilmenge $\mathcal{R} \subset \mathbb{Z}_+^m \times \mathbb{Z}_+^m$ enthält $n_{\mathcal{R}} = \#\mathcal{R}$ Paare (α, β) zu den im Gebiet Ω stattfindenden Reaktionen. Entsprechend dem Massenwirkungsgesetz sind die Nettoreaktionsraten durch

$$R_\nu := \sum_{(\alpha, \beta) \in \mathcal{R}} k_{(\alpha, \beta)} \left(e^{\sum_{i=1}^m \alpha_i v_i} - e^{\sum_{i=1}^m \beta_i v_i} \right) (\beta_\nu - \alpha_\nu)$$

definiert. Aufgrund der heterogenen Materialien sind die Referenzdichten \bar{u}_ν , die Diffusionskoeffizienten D_ν , und die Reaktionskoeffizienten $k_{(\alpha, \beta)}$ (möglicherweise in nicht glatter Form) vom Orte abhängig. Zudem können die Koeffizienten D_ν und $k_{(\alpha, \beta)}$ vom Zustand aller Spezies abhängen. Ferner sind Reaktionen zugelassen, die nur auf Teilgebieten existieren. Gegenstand der Arbeit sollen nur Reaktionssysteme sein, die einen positiven Gleichgewichtszustand besitzen, der durch den stöchiometrischen Teilraum und den Anfangswert bestimmt ist. Es stellt sich heraus, dass diese stationäre Lösung das einzige thermodynamische Gleichgewicht des Systems ist.

Der Beitrag dieser Arbeit betrifft die örtliche Voronoi Finite-Volumen und die zeitliche implizite Euler Diskretisierung solcher Reaktions-Diffusionssysteme. Für den Beweis von Aussagen über die Stabilität von Lösungen und von a priori-Abschätzungen für die diskreten Lösungen werden Techniken in die diskrete Welt übertragen, die sich bereits beim Studium des kontinuierlichen Problems als zielführend herausgestellt haben. Der Ausgangspunkt sind dabei Abschätzungen der diskreten freien Energie

$$\hat{F}(\mathbf{u}) = \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| \left(u_{\nu K} \left(\ln \frac{u_{\nu K}}{\bar{u}_{\nu K}} - 1 \right) + \bar{u}_{\nu K} \right),$$

die entlang von Trajektorien des diskreten Problems monoton und exponentiell gegen ihren Gleichgewichtswert $\hat{F}(\mathbf{u}^*)$ fällt. Unter zusätzlichen Annahmen an die Reaktionsordnung (Reaktionen mit maximal quadratischen Quelltermen) können aus dieser Aussage a priori-Abschätzungen für die Konzentrationen gewonnen werden. Ein neues Resultat betrifft die uniforme zeitliche globale Beschränktheit von oben der Konzentrationen, welche mit Moser Iteration gezeigt wurde. Im Beweis wird ein diskretes Analogon der Gagliardo–Nirenberg Ungleichung benutzt, das im Abschnitt A.2 bereitgestellt wird.

Mittels des exponentiellen Fallens der freien Energie konnte auf die zeitliche globale Beschränktheit der chemischen Potentiale in $L^1(\Omega)$ geschlossen werden. Wiederum mit Hilfe einer Moser Iteration konnte die uniforme globale Beschränktheit der Konzentrationen von Null weg gezeigt werden. Ausdrücklich sei darauf hingewiesen, dass in beiden Fällen die auftretenden Konstanten unabhängig von der Güte des Voronoi-Gitters und vom Zeitschritt sind, womit auch anisotrope Gitter nicht ausgeschlossen sind.

Die lokale Existenz von Lösungen des diskreten Problems konnte unter den natürlichen Voraussetzungen der Quasipositivität der Reaktionsterme und der Erhaltung der Atomzahl bewiesen werden. Ein Einzigkeitsresultat konnte nur für vom Zustand unabhängige Diffusionskoeffizienten D_ν und

für kleine Zeitschritte erzielt werden. Zusammenfassend konnten alle qualitativen Eigenschaften des kontinuierlichen Systems (thermodynamisches Gleichgewicht, monotonen und exponentielles Fallen der freien Energie, globale obere und untere Schranken) auch für das diskretisierte Problem nachgewiesen werden.

Ein weiteres neues Resultat betrifft die Konvergenz des Schemas. Unter Zuhilfenahme der globalen Beschränktheit der Konzentrationen von unten und oben konnten Kompaktheitsaussagen für die Konzentrationen und die chemischen Aktivitäten für eine Folge von Diskretisierungen erzielt werden. Damit kann auf konvergente Teilfolgen von Konzentrationen und Aktivitäten geschlossen werden, deren Grenzwerte eine schwache Lösung der kontinuierlichen Aufgabe sind.

Eine prototypische Implementierung des Schemas für das Beispiel der Michaelis-Menten-Henry-Kinetik wird im Abschnitt 4 aufgezeigt. Hierbei wird großer Wert auf die Erhaltung der theoretisch erzielten Eigenschaften über große Zeitintervalle gelegt.

Um die Stabilität und Anwendbarkeit der Methode auf reale Probleme zu demonstrieren, werden im Abschnitt 5 verschiedene Beispiele mit Michaelis-Menten-Henry-Kinetik betrachtet. Sie widmen sich speziellen Eigenschaften der implementierten Methode. In allen Beispielen sind während der Rechnungen die physikalischen Eigenschaften des Schemas bis auf Rundungsfehler auch über lange Zeitintervalle erhalten. Darüber hinaus löst das analysierte Verfahren die verschiedenen Zeitskalen des Systems auf, wenn auch wandernde Reaktionsfronten oder starke Gradienten in der Nachbarschaft von Materialübergängen nur grob durch das Gitter approximiert sind.

1.2 Résumé

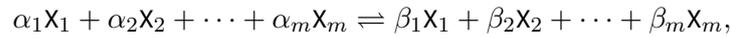
The modeling of a process in photolithography [95] leads to reaction-diffusion systems with spatially dependent (nonsmooth) data. Of particular interest is a chemical process [41], which creates structures in semiconductor manufacturing and takes place in two materials.

We consider systems of parabolic partial differential equations which model the relocation of chemical, electrically neutral species in heterogeneous materials induced by diffusion processes and chemical reactions. In this setting we expect difficulties in the numerical analysis caused by strongly coupled systems, nonsmooth spatially dependent reaction and diffusion coefficients, and the absence of the maximum principle.

In detail, on a bounded subset $\Omega \subset \mathbb{R}^2$ a system consisting of m electrically neutral species is examined. The concentration of the species X_ν , $\nu = 1, \dots, m$, is denoted by u_ν , their reference density by \bar{u}_ν , and their chemical potential by v_ν . The connection between their concentration and their chemical potential is given by Boltzmann statistics $u_\nu = \bar{u}_\nu e^{v_\nu}$. The model consists of m continuity equations

$$\mathbf{j}_\nu = -D_\nu u_\nu \nabla v_\nu - \frac{\partial u_\nu}{\partial t} + \nabla \cdot \mathbf{j}_\nu = R_\nu \quad \text{in } \mathbb{R}_+ \times \Omega$$

with Neumann boundary conditions $\partial_n \mathbf{j}_\nu = 0$ on $\partial\Omega$. The reaction rates R_ν describe reversible chemical reactions of the form



where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)$ and $\boldsymbol{\beta} = (\beta_1, \dots, \beta_m)$ denote vectors of stoichiometric coefficients. According to the mass action law, the reaction rates are given by

$$R_\nu := \sum_{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{R}} k_{(\boldsymbol{\alpha}, \boldsymbol{\beta})} \left(e^{\sum_{i=1}^m \alpha_i v_i} - e^{\sum_{i=1}^m \beta_i v_i} \right) (\beta_\nu - \alpha_\nu).$$

Due to the heterogeneous materials, the reference densities \bar{u}_ν , the diffusion coefficients D_ν , as well as the reaction coefficients $k_{(\boldsymbol{\alpha}, \boldsymbol{\beta})}$ are spatially dependent in a nonsmooth way. The coefficients D_ν and $k_{(\boldsymbol{\alpha}, \boldsymbol{\beta})}$ may also depend on the state of all species. Moreover we allow that reactions can vanish on subdomains. In this work we consider only the case, where the reaction system possesses a positive stationary state, determined by the stoichiometric subspace and the initial state. This stationary state turned out to be the only thermodynamic equilibrium of the system.

Our investigations are related to the Voronoi finite volume discretization in space and backward Euler discretization in time of such reaction-diffusion systems. In order to obtain a priori bounds and stability results of the discrete problem we apply techniques which are well established in the analysis

of the continuous problem. The origin of studying the discrete evolution problem are physically inspired estimates of the discrete free energy

$$\widehat{F}(\mathbf{u}) = \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| \left(u_{\nu K} \left(\ln \frac{u_{\nu K}}{\bar{u}_{\nu K}} - 1 \right) + \bar{u}_{\nu K} \right),$$

which decays along trajectories of the discrete evolution system monotonously and exponentially to its equilibrium value $\widehat{F}(\mathbf{u}^*)$. These statements deliver a priori estimates on the concentrations under additional assumptions on the reaction order (reactions with at most quadratic source terms). In a first new result we conclude from the global boundedness of the free energy, the global in time boundedness of the species concentrations from above by Moser iteration. During the proof, a discrete version of the Gagliardo–Nirenberg inequality is required, which is provided in Section A.2.

Exploiting the exponential decay of the free energy, we obtain the global in time boundedness of the chemical potentials in L^1 . Again using Moser iteration we prove the global boundedness of the species concentrations from below by some positive constant only depending on the data and not on the mesh.

Local existence of solutions of the discrete system is obtained under the natural assumption of quasi-positivity and conservation of number of atoms. A uniqueness result was achieved under the assumption that the diffusion coefficients are independent of the state variables and under the restriction of small time steps.

In an additional new result, the convergence of the scheme is proven in Section 3.12. Using the global bounds from above and below we conclude compactness of the concentrations and chemical potentials for a sequence of discretizations. Hence there exist converging subsequences such that the limit is a weak solution of the continuous problem. At the end of the chapter, all qualitative properties relating to the continuous problem are also true for the discrete problem (steady states, thermodynamic equilibrium, monotonous exponentially decay of the free energy, global upper and lower bounds).

In Chapter 4 we present a prototypical implementation of the method for the Michaelis-Menten-Henry reaction kinetic. Our aim is to preserve the physical properties of the calculated solution up to rounding errors over large time intervals.

The stability of the method is illustrated by means of examples in the subsequent Chapter 5. The examples are devoted to different phenomena of the scheme. During the calculations the properties of the scheme are valid up to rounding errors. Moreover, all timescales are resolved by the method, even if the front movement or the strong gradients in the direct neighborhood of interfaces are only roughly resolved.

1.3 Application in photolithography

The investigation of stable discretizations for reaction-diffusion systems in heterogeneous materials is motivated by two processes for the manufacturing of structures in optical lithography.

Higher integration of electronic components provides a mean to reduce the cost per component and to increase the performance in the electronic industry. The development of integrated circuits is driven by the Moore's law [76, 95]. Every 24 months the number of transistors on a chip is doubled. Modern circuits from Intel are manufactured at 22 nm structure size and the fabrication of 15 nm structure size is under development, [25, 76]. In Figure 1.1, the forecast of the structure size (half pitch¹) is depicted. Optical lithography is a standard technology to transfer a pattern from a mask

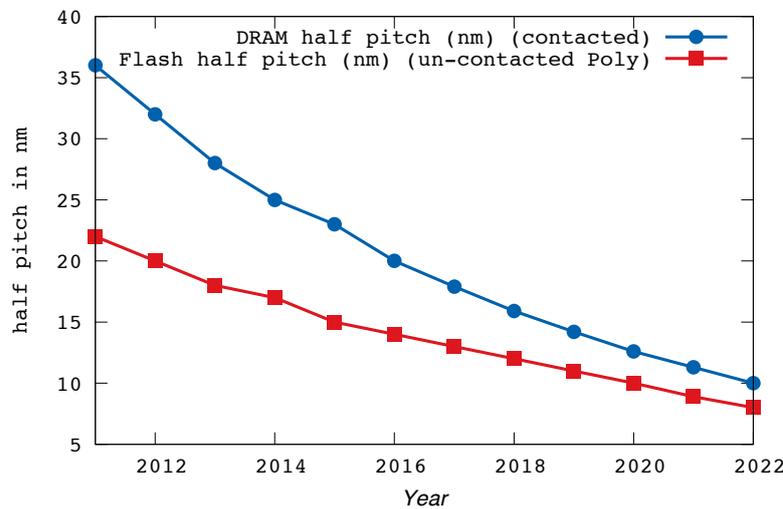


Figure 1.1: Forecast of the structure size (half pitch) of DRAM's and Flash memory. The data are listed in [76]

onto the resist-coated wafer² and continue onto a integrated circuit, see Figure 1.2 and [95, 118]. The standard process consist of nine steps, see [95, 118]:

- Since an absolute clean wafer is essential for the lithography process the wafer surface is cleaned by mechanical scrubbing, high pressure water stream, or nitrogen blow-off.
- Afterwards the photoresist, which is a suspension of polymers and additives in a solvent, is spin-coated onto the wafer. For the following exposure step it is necessary that the thin film of photoresist is defect free and uniform.

¹ This denotes the half of the distance of two neighboring lines.

² Two types are possible: Using a positive resist the soluble increases during exposure. Using a negative resist the solubility decreases during exposure.

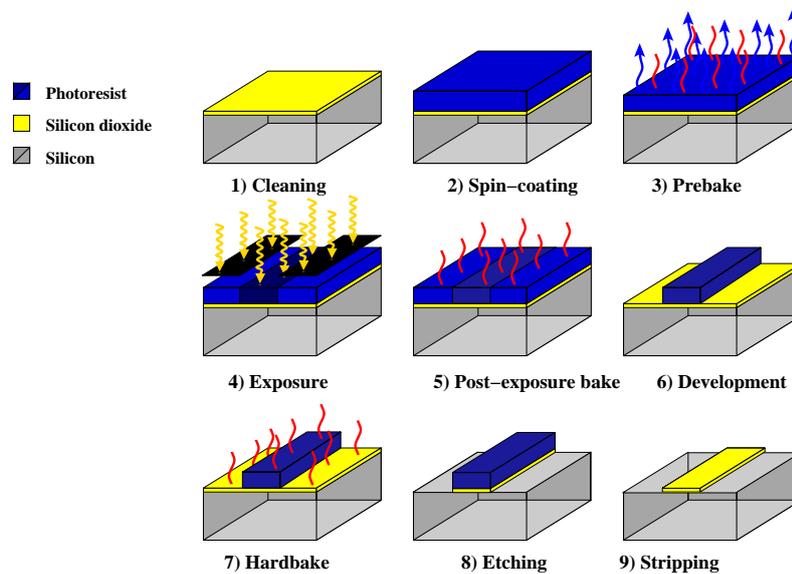


Figure 1.2: The optical lithography process steps to transfer a structure onto a integrated circuit

- During the pre-brake the waver is heated up on a hot plate. Most of the solvent is diffused out and the photoresist becomes a mechanically stable film. The remaining solvent has an influence on the optical properties and the diffusion constant of the acid and base during post exposure bake.
- One of the main steps is the exposure, where the image of a photomask is transferred into the photoresist and causes the spatially selective generation of acid from photo acid generator molecules. The resolution of the projection system is essential for the performance of the process and is characterized by two quality criteria. The first criteria is the resolution R defining the minimal distance between two neighboring features that can still be resolved. The resolution R of the method is limited by the Rayleigh quotient

$$R = k_1 \frac{\lambda}{A_N}.$$

In this equation λ is the wavelength of the light, A_N denotes the numerical aperture and k_1 denotes a process constant of the used lithography devices. A common value of k_1 is 0.3 and the theoretical minimum is 0.25. The numerical aperture is given by $A_N = n \sin \theta$, where n is the refractive index of the medium between the photoresist and projection lens and θ denotes the half maximum illumination angle between the exit lens of the projection system and the image plane. The second criteria is the depth of focus DOF which characterizes the tolerable

defocus for printing the features within the process specifications. The *DOF* is approximately given by the second Rayleigh criterion

$$DOF = k_2 \frac{n\lambda}{A_N^2},$$

where k_2 is a technology parameter. Obviously reducing the R with a shorter wavelength λ or by an higher A_N is in contradiction to obtaining a good *DOF*.

- During the post exposure (PEB) step, the photo acid molecules, generated during exposure, catalyze the decomposition of the polymer dissolution inhibitor groups. In unexposed regions small amounts of acid are neutralized by base molecules, which are added to the photoresist. This fast neutralization prevents from significant cumulative inhibitor deprotection in stray light regions. Additional reactions of the byproducts can be neglected, since most of them are gaseous and evaporate from the wafer. The process carried out on a hot plate by a temperature of 120°C for a time of 90s . Therefore the acid and base molecules begin to diffuse, which reduces the amplitudes of standing wave patterns in the latent acid profiles, generated during exposure. The main purpose of the PEB step is to transfer the exposed image of the photo acid generator with some small modifications due to the diffusion and the acid-base neutralization into the deprotected, highly soluble, polymer chains. Unexposed regions of the photoresist stay insoluble, due to the fast neutralization of acid and base.
- The propose of the development step is to transform the latent photomask image, represented after PEB by a varying degree of inhibitor deprotection, into a three-dimensional relief structure. This means regions with a low concentration of dissolution inhibitor and high development rates are dissolved. The main chemical reaction is a acid-base neutralization, which ionized Acidic polymer sites, deblocked during PEB, by the alkalic developer and dissolve in the developer liquid. Usually the developer liquid is sprayed on the waver and after a certain time the dissolved photoresist and the developer liquid is removed by spin-drying.
- The propose of the hardbake step is similar to the prebake step. The waver is heated up on an plate and remaining solvents or polymers are diffuses out, which improves the adhesion and the mechanical resilience for the following etching step.
- Using dry etching the pattern is transferred from the photoresist into the substrate. Using gasses the uncovered parts of the silicon dioxide layer are removed. Silicon dioxide regions with photoresist on top are protected against the etching gasses. In the last step, the remaining photoresist on top of the silicon dioxide is removed by dry or wet chemical stripping.

New challenges arise by the projection of smaller and smaller structures with optical lithography and from rapid developments in this business. A modification of the values k_1 and λ has a big impact

on the process performance. Using modern immersion lithography³ with a wavelength of 193 nm even makes it possible to shrink the structures to a size of 45 nm half-pitch, see [127].

By reducing the wave length λ or the process constant k_1 it is possible to achieve smaller structures of a size less than 45 nm. The following two processes decrease the process constant k_1 .

1.3.1 Double patterning

Using double exposure technologies it is possible to reduce the factor k_1 . The idea is to superimpose two or more images on the wafer, see Figure 1.3. The process consists of many complex substeps

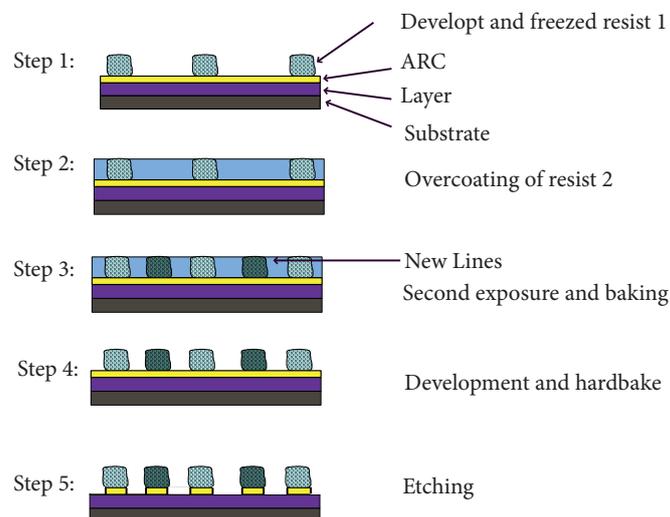


Figure 1.3: Substeps of the double exposure process: 1.) L1 structure + freezing, 2.) L2 resist over coated 3.) L2 exposure and baking, 4.) L2 development, 5.) Etch

which are extremely hard to control. The advantage is that only common materials are used which are also applicable in the standard procedure. If all substeps are under control it is possible to create structures of 32 nm size and below.

After the development step of the standard lithography sequence a second standard lithography sequence is done to create lines with half the distance of the first sequence, see Fig 1.3. During the exposure and development step a so called post exposure bake [95] takes place to smooth the exposed image using diffusion and to stop the catalytic reaction, which was started during the exposure step.

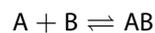
³ Water droplets are inserted between the silicon wafer and the exposure lens.

The post exposure bake during the first lithography step (L1) can be modeled by the *Meta-model* [74, 96, 118], i.e., the following reactions take place



The dissolved parts of the resist M being catalytically decomposed into a product P using an acid A . The symbol X denotes an intermediate with a short life time. Using a second reaction the acid A is neutralized by a base B . It is assumed that the species can diffuse accordingly to Fick's law and the diffusion coefficients are assumed to be constant.

Starting from the Meta-model, we developed a new variant including resist interactions between the first structure created by L_1 and the new second structure L_2 , see [30, 39]. In detail, we assume that in the first structure of L_1 an additional acid base neutralization



with species A and B takes place. The symbol AB denotes a decomposition product. Moreover it is allowed that the species can diffuse from the L_2 into the L_1 region. Using this models it is possible to study resist interactions between different lithography steps. For example we investigate the footing effect which was experimentally observed for different resist formulations, see Figure 1.4. One explanation of the footing is a loss of acid A caused by the existence of a standing lightwave

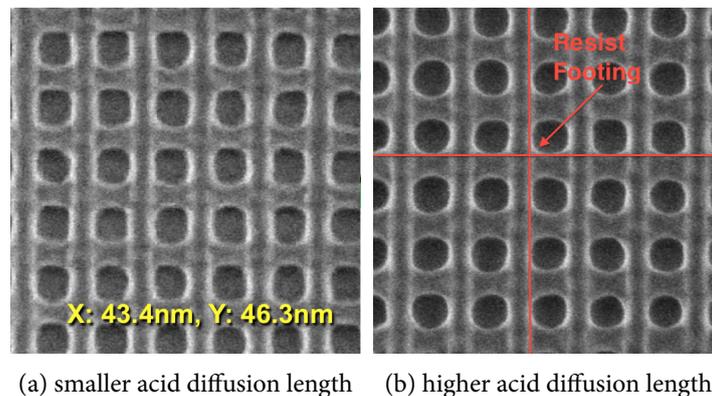


Figure 1.4: Scanning electron microscope (SEM) picture of the observed footing effect for different acid diffusion coefficients. Images are from [30, Talk at SPIE]

created during the exposure step. Using the new variant of the Meta-model another explanation of the footing effect was found. In numerical simulations we could show that the diffusion length in L_1 of the involved species has an impact on the resist footing, i.e., the acid A diffuses from the

L2 structure into the L1 structure and is not available for the catalytic reaction in L2, see [30, 39] and Figure 1.5. The acid loss in the L2 region and the associated footing effect can be reduced by

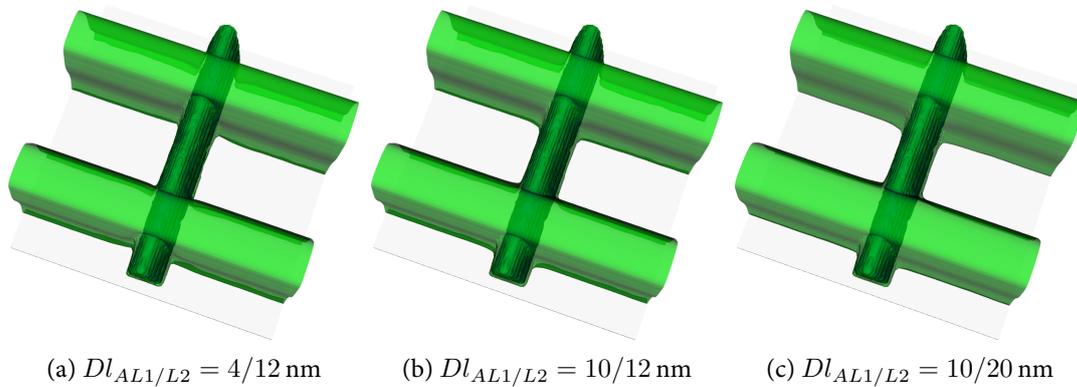


Figure 1.5: Footing-effect of the structure after development for different L1/L2 acid diffusion constants, see [30]

a certain amount of acid inside L1 or a protective layer with small diffusion length on the top of the L1 structure, see [30, 39]. The simulations were done with the lithography simulator Dr.LiTHO⁴ in combination with the module WiasPeb⁵. The code Dr.LiTHO calculates the exposure step from which the module WiasPeb calculates the post exposure bake step. The result of WiasPeb serves as input for the calculation of the development step in Dr.LiTHO. The coupling of the two codes was done using the scripting language Python [114] and the utility Swig [8]. The problem is characterized by three-dimensional geometries and heterogeneous materials. In every subregion of L1, L2, the reaction rates and diffusion coefficients may vary.

1.3.2 Pattern doubling

A new possibility to create structures less than 45 nm is the so called self-limiting residual acid diffusion process, see [40, 41]. The advantage of this approach is that it is relatively easy to control. It does not require multiple shifted exposure and development steps compared to the process of Subsection 1.3.1. The starting point of the process is a primary structure containing a deposit of acid A. In a second step (2), the primary structure is over-coated by a new non-attacking solvent. During a bake process, the acid diffuses from the primary structure into the over-coated region and starts a catalytic cross-linking reaction with precursor C. The end time of the cross-linking reaction can be controlled by the neutralization of A and a base B. After the cross-linking reaction is finished, the non cross-linked material is developed using standard developer, and etched away together with the original pattern see step (4) and (5). A schematic of the process together with SEM pictures of the step (1) and (3) is shown in Figure 1.6.

⁴ Developed at the Fraunhofer IISB [38].

⁵ The module based on the toolbox pdelib2 [123] of the WIAS.

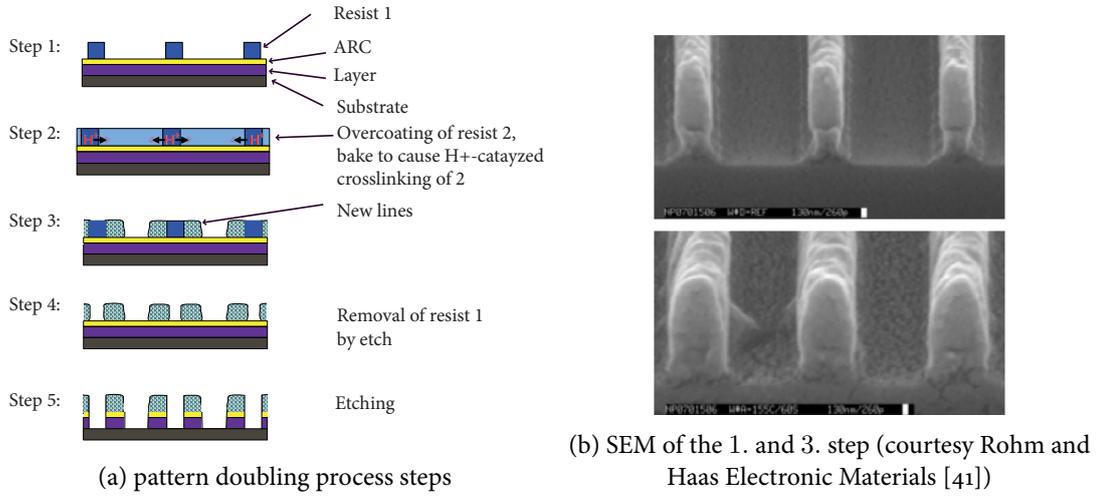
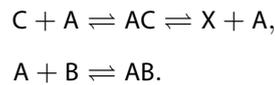
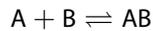


Figure 1.6: Qualitative description of a pattern doubling process using deprotection and cross-link chemistry resists

In [40, 41] the process is modeled using reaction-diffusion equations. In the over-coated region, a two-step catalytic reaction with an acid A as a catalyst is assumed. During a first reaction step it is assumed that a intermediate AC is formed by combining the acid A and the precursor C . In a second step, the intermediate AC attaches to a cross-linked polymer X , releasing the acid A . This catalytic cycle leads to a moving reaction front from the primary region into the over-coated region, i.e. the thickness of the reaction front of the cross-linked polymer X grows in the over-coated region. In order to control the spread of the cross-linked area, a small amount of base B is added, which neutralizes a part of the free acid A released during the catalysis process. If the base concentration is chosen in such a way that after a finite interval of front moving, all the catalyzing acid A is neutralized, the reaction front might stop, leaving a certain part of the spacer unaffected. Formally this can be expressed by



In the primary region, the base B can diffuse into this region, and can neutralize the acid A , i.e. the neutralization reaction



takes place. The acid A and the base B are assumed to be mobile. All other species are assumed to be immobile. The reaction rates may depend on the space variable, and the diffusion coefficients of A and B depend on the amount of C and X, i.e., the diffusion coefficient of A is of the form

$$D_A(u_C, u_X) = D_0 \exp\left(-\gamma \frac{\alpha u_C + (1 - \alpha)u_X}{\beta u_C + (1 - \beta)u_X}\right),$$

where α, β, γ are constants and u_C, u_X denote the concentration of C and X, see [40]. The diffusion coefficient of B is modeled analogously. The process strongly depends on the ratio of the reaction

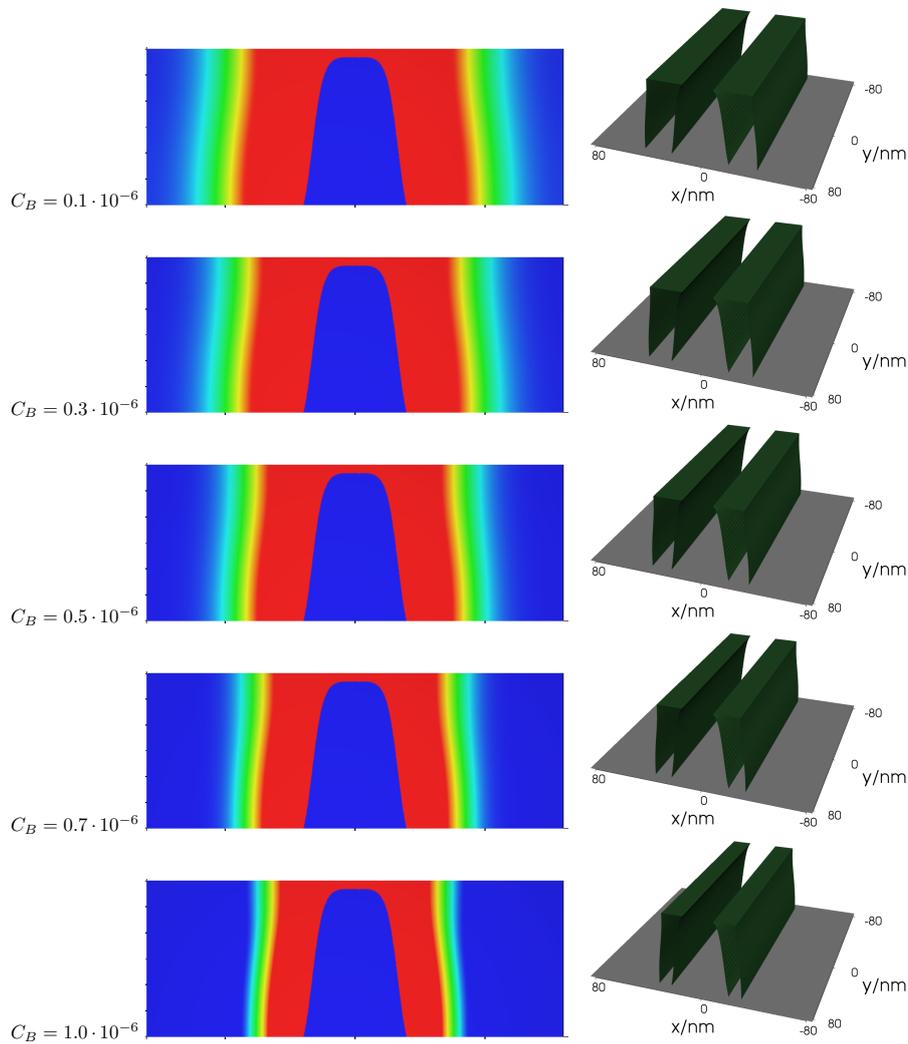


Figure 1.7: Dependence of the concentration of species X (left) and the profile of the L2 line (right) after development on the concentration of species B

rates of the two steps of the catalytic cross-linking reaction in the primary region. Assuming that the

backward reactions are very slow we observe a reaction front if the first forward reaction is much greater than the second cross-linking reaction step. This means that within the over-coated zone, a certain amount of intermediate AC is built up before it is used for cross-linking (building of P) in the second step. The acid base neutralization drops out the acid of the catalytic reaction and can not be "recycled" for the first step. If the base concentration is chosen appropriately, it is possible to stop the reaction front after a finite time.

With the aid of numerical simulations we can show that the concentration of B has an influence on the profile of the cross-linked polymer X and causes a bigger line after development, see Figure 1.7. The module WiasPeb calculates the cross-linking step 2 and the program Dr.LiTHO uses the profile of the cross-linked polymer X to compute the development of the line. The difficulty in the numerical treatment lies in the existence of heterogeneous materials and the resolution of the dynamic reaction fronts by the mesh. A not well resolved front leads to a direct impact on the profile of the new line after development, see [40].

Remark 1.3.1. *Applying in Subsection 1.3.1 an additional exposure, bake and development sequence after step 4 (see Figure 1.3), it is possible to create structures of 10 nm size. The resulting sequence is called triple patterning or multi patterning, see [50, 93, 94]. Such processes may lead to complex interactions between the different patterns during the exposure and post exposure bake step.*

Remark 1.3.2. *The modeling in the papers [30, 39–41] uses the concentration as primary variable. Due to the different materials of the L1 and L2 structure this choice is not appropriate and leads to an thermodynamically inconsistent model, see Remark 1.4.4. The proposed model of Section 1.4 takes this into account by using the chemical potential as primary variable. Except vanishing diffusion coefficients, the model also includes all difficulties of the application: nonsmooth spatially dependent coefficient functions, dependents of the reaction rate and diffusion coefficients on the state of all species, and vanishing reactions on subdomains. The different examples provided in Chapter 5 deal with all these dependences.*

Remark 1.3.3. *Using local grid refinement it might be able to resolve the dynamic reaction fronts. For this purpose stable error estimators and interpolation algorithms are required which preserve the physical properties of the equations. The study of the usability of these methods for reaction-diffusion equations with heterogeneous materials is not subject of this work. We refer to the literature of local grid refinement for reaction-diffusion systems, see [37, 73, 75, 88, 89].*

1.4 Formulation of model equations

In this section, a model for reaction-diffusion systems including heterogeneous materials is introduced, see [26, 51, 53, 64]. We consider m species X_ν with initial densities U_ν which undergo chemical reactions and underlying diffusion processes in a domain $\Omega \subset \mathbb{R}^2$ with boundary $\Gamma = \partial\Omega$. The relation between the densities $u_\nu : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}_+$ of the species X_ν and the corresponding chemical potentials $v_\nu : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$ is assumed to be given by Boltzmann statistics,

$$u_\nu = \bar{u}_\nu e^{v_\nu}, \quad \nu = 1, \dots, m, \quad (1.4.3)$$

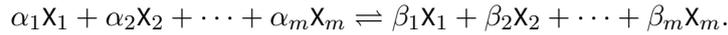
where the reference densities $\bar{u}_\nu : \Omega \rightarrow \mathbb{R}_+$ may depend on the spatial position and express the possible heterogeneity of the system under consideration. For a more elaborated description of the chemical potentials we refer to [68, 86].

For the fluxes j_ν we make the ansatz

$$j_\nu = -D_\nu u_\nu \nabla v_\nu, = -D_\nu \bar{u}_\nu e^{v_\nu} \nabla v_\nu, \quad \nu = 1, \dots, m, \quad (1.4.4)$$

with diffusion coefficients $D_\nu : \Omega \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ which may depend on the space variable and on the state variables. If the reference densities \bar{u}_ν are constant, the relation $j_\nu = -D_\nu \nabla u_\nu$ holds.

To describe chemical reactions we assume that $\mathcal{R} \subset \mathbb{Z}_+^m \times \mathbb{Z}_+^m$ is a finite subset. A pair $(\alpha, \beta) \in \mathcal{R}$ represents the vectors of stoichiometric coefficients of reversible reactions, usually written in the form



According to the mass action law, the *reaction rate* for a pair $(\alpha, \beta) \in \mathcal{R}$ is given by

$$R_{(\alpha, \beta)} = R_{(\alpha, \beta)}(\cdot, a_1, \dots, a_m) = k_{(\alpha, \beta)}(a^\alpha - a^\beta), \quad (1.4.5)$$

where $k_{(\alpha, \beta)} : \Omega \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ is the reaction coefficient, $a_\nu := \exp(v_\nu)$ is the chemical activity of X_ν , and $a^\alpha := \prod_{\nu=1}^m a_\nu^{\alpha_\nu}$. The net production rate of species X_ν corresponding to all reaction rates is

$$R_\nu := \sum_{(\alpha, \beta) \in \mathcal{R}} k_{(\alpha, \beta)}(a^\alpha - a^\beta)(\beta_\nu - \alpha_\nu). \quad (1.4.6)$$

The stoichiometric subspace \mathcal{S} is defined by

$$\mathcal{S} = \text{span}\{\alpha - \beta : (\alpha, \beta) \in \mathcal{R}\} \quad (1.4.7)$$

and its orthogonal complement in \mathbb{R}^m is denoted by \mathcal{S}^\perp .

We consider a closed isothermal process, i.e., the temperature and the pressure are treated as constant parameters. In this notation our reaction-diffusion system consists of m continuity equations with Neumann boundary conditions:

$$\begin{aligned} \frac{\partial u_\nu}{\partial t} + \nabla \cdot j_\nu &= R_\nu \text{ in } \mathbb{R}_+ \times \Omega, & \mathbf{n} \cdot j_\nu &= 0 \text{ on } \mathbb{R}_+ \times \Gamma, \\ u_\nu(0) &= U_\nu \text{ in } \Omega, & \nu &= 1, \dots, m. \end{aligned} \quad (1.4.8)$$

In Remark 2.1.1 some examples illustrate the included dependencies.

Example 1.4.1. *In order to illustrate some basic facts of the model in the thesis, the Michaelis-Menten-Henri reaction mechanism (MMH) [17] is introduced. The reaction step is well known in chemistry and biology and is a model for many fundamental reactions, see [15, 106]. The mechanism is also present in the modeling of a pattern doubling process, see Section 1.3.*

It involves a substrate X_1 binding to a catalyst X_2 to form an intermediate X_3 which decomposes into X_2 and a product X_4 . The mechanism can be symbolically expressed by



Correspondingly, the set \mathcal{R} consists of two pairs of vectors, namely $\alpha_1 = (1, 1, 0, 0)$ and $\beta_1 = (0, 0, 1, 0)$ for the first reaction, and $\alpha_2 = (0, 0, 1, 0)$ and $\beta_2 = (0, 1, 0, 1)$ for the second reaction.

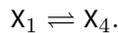
Due to the definition of the reaction term (1.4.6), the net production rates of the species are given by

$$\begin{aligned} R_1(a) &= -k_{(\alpha_1, \beta_1)}(a_1 a_2 - a_3), \\ R_2(a) &= -k_{(\alpha_1, \beta_1)}(a_1 a_2 - a_3) && + k_{(\alpha_2, \beta_2)}(a_3 - a_2 a_4), \\ R_3(a) &= +k_{(\alpha_1, \beta_1)}(a_1 a_2 - a_3) && - k_{(\alpha_2, \beta_2)}(a_3 - a_2 a_4), \\ R_4(a) &= && + k_{(\alpha_2, \beta_2)}(a_3 - a_2 a_4). \end{aligned}$$

The stoichiometric subspace \mathcal{S} and its orthogonal complement \mathcal{S}^\perp are spanned by

$$\mathcal{S} = \text{span}\{(1, 1, -1, 0), (0, -1, 1, -1)\}, \quad \mathcal{S}^\perp = \text{span}\{(1, 0, 1, 1), (0, 1, 1, 0)\}.$$

Remark 1.4.2. *The MMH mechanism (1.4.9) is a catalytic reaction and has a higher reaction rate than an uncatalyzed reaction*



at same environmental conditions. Catalytic reactions start at a lower activation energy than the corresponding uncatalyzed reactions. The MMH reaction step was first deduced for the splitting of sucrose into glucose and fructose, see [17]. In the year 1913, Michaelis and Menten assumed that the intermediate X_3 is in equilibrium and gave an explicit formula for the concentration of the intermediate

χ_3 , see [16, 99]. An interesting review of MMH kinetics in the context of ordinary differential equations is given in [119]. Steady-state approximations for different relations between reaction and diffusion time scales in one space dimension are formulated in [80].

Remark 1.4.3. *If the temperature of an ideal gas (at given pressure) is sufficiently small, Boltzmann statistic becomes invalid. In this special situations, it is common to replace Boltzmann statistics by Fermi-Dirac statistics, i.e.,*

$$u_\nu = \bar{u}_\nu \mathcal{F}_{1/2}(v_\nu), \quad \mathcal{F}_{1/2}(s) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{r}}{1 + e^{r-s}} dr,$$

in order to calculate the particle distribution from the energy, see [86].

Remark 1.4.4. *In contrast to the reactions without heterogeneous materials the concentrations may jump at the interface between two materials, see Chapter 5. Therefore the chemical potentials or activities are the primary variables to describe a system with a thermodynamic equilibrium. Using the potentials as variables it is also possible to include other effects, like pressure, temperature, or the electric and magnetic field, see [46, 68, 86].*

Remark 1.4.5. *We are primarily interested in the study of elementary chemical reactions. Reactions of this kind, which model real life molecular events, are always reversible. Since the model must be able to describe fully mixed steady states, it is necessary to impose the full reversibility of the reactions. We refer to [51, Sec. 2.4.] for a more detailed discussion about that topic. The simplification of elementary chemical reaction networks may lead to irreversible reactions, which do not represent real molecular events.*

1.5 Relevant literature

The literature about reaction-diffusion systems is very comprehensive. Therefore we only discuss the papers containing results or methods which are relevant for proceeding in this thesis.

Reaction-diffusion systems in homogeneous materials The question of global existence of strong and weak solutions of special reaction-diffusion systems without heterogeneous materials is treated, e.g., in [11, 12, 22, 23, 82, 111, 113]. The papers [111, 113] give an interesting review about the topic of global existence under the assumptions that the mass of the system is preserved and the reaction terms are quasi-positive and fulfilling a growth condition. The authors provide a global existence result of strong solutions for reaction-diffusion systems with quadratic reactions, constant diffusivities, and bounded initial data in space dimension $N \leq 5$. The results strongly rely on an L^2 estimate which uses maximal regularity theory on the dual equation and requires the linearity of the diffusion operators. In [113], the approach is extended to smooth, two times differentiable diffusion coefficients. Moreover the global existence of strong solutions for reaction-diffusion systems including advection-migration is studied. The fast reaction limit (rate coefficient tends to infinity) of concrete systems and the resulting limit system is studied in [12, 13, 113]. Especially for a reaction system close to the MMH-kinetic, the limit for a short lifespan of the intermediate X_3 is studied in [12].

Reaction-diffusion systems in heterogeneous materials In the case of reaction-diffusion systems including heterogeneous materials, electrically charged species, and drift terms we mention the papers [46, 47, 53, 63, 65, 67] and references therein. The authors explore and study properties of an energy functional (free energy) which turns out to be a Lyapunov functional of the system. Under an assumption on the source terms of the reaction (at most quadratic source terms in 2d), they provide global well-posedness and stability, i.e., all bounded solutions converge to a uniquely determined thermodynamic equilibrium solution as time goes to infinity. Moreover the free energy decays monotonously and exponentially along trajectories. The work of [46, 47, 61] additionally deals with more general state equations than Boltzmann statistics (1.4.3). In Chapter 3 we adapt these techniques in order to prove global well-posedness of the discretized problem.

There is a new direction – considering reaction-diffusion systems as a gradient-flow problem. The gradient-structure was initially mentioned in [90, 100, 101]. In [102] this structure was exploited in order to provide explicit bounds for the exponential decay of the relative free energy. The authors allow for some vanishing species diffusion coefficients and consider different energy functionals.

Finite volume methods for reaction-diffusion systems Finite volume schemes were introduced by engineers for the study of equations with conservation laws and have been studied mathematically for over 20 years now. In the recent years finite volume schemes for linear diffusion problems involving

the Laplace equation have been proven to be convergent on admissible meshes, that are meshes such that the line segment joining the centers of neighboring cells is orthogonal to the edge between the two cells, see [32, 33].

Finite volume methods for nonlinear parabolic systems, especially global existence of solutions and convergence of such schemes, are an ongoing research topic, see [7, 9, 33, 81]. Obtaining uniform global L^∞ -bounds on the concentration, concluding compactness and then passing to the limit is the standard approach for proving convergence of a scheme to a weak solution of the continuous system, see [7, 34]. In the case of nonlinear systems, where maximum principles and upper and lower solutions method are not applicable, the proof of L^∞ bounds is difficult.

In the setting of finite volume methods it is common to adapt continuous functional analysis techniques in order to prove known properties of the continuous problem also for the discretized problem, see [10, 32, 60]. For example in the papers [54, 56, 57] energy estimates for reaction-diffusion systems with quadratic source terms with or without electrically charged species are established using a discrete Poincaré inequality.

Van Roosbroeck-equation We mention the special case of a reaction-drift-diffusion system with two electrically charged species. The Van Roosbroeck system models the carrier transport in semiconductor devices and consists of two continuity equation for the electron density p and the hole density n coupled with a Poisson equation combining the carrier density with the electrostatic potential. In the context of semiconductor devices more general boundary conditions are of interest in order to model applied voltages and other phenomena. The Van Roosbroeck system is well analyzed and we refer to the articles [45, 115]. In Chapter 4 we pick up ideas coming from the field of semiconductor devices simulation for the assembly of the equations and the adaptive time stepping, see [48].

2

The continuous problem

Abstract This chapter is devoted to the precise statement of the mathematical problem. We rephrase general assumptions concerning the data of the continuous problem and present a weak formulation, which is suitable when dealing with heterogeneous materials. A subsequent part gives a summary on results obtained so far for the continuous problem. The main ideas are physically motivated estimates, invariants, and a bootstrapping technique introduced by Moser [104]. Such techniques are common in the analysis of nonlinear parabolic PDEs, see [78, 79, 115, 125].

2.1 General assumptions on the data

In this section we formulate basic assumptions with respect to the data of the problem, see [46, 47, 53, 56, 64]. The problem is studied under the following assumptions:

Assumption A1. $\Omega \subset \mathbb{R}^2$ is a bounded polygonal domain, $\Gamma := \partial\Omega$.

Let $m \in \mathbb{N}$ be given and \mathcal{R} be a finite subset of $\mathbb{Z}_+^m \times \mathbb{Z}_+^m$. For all $(\alpha, \beta) \in \mathcal{R}$ the reaction rate coefficients $k_{(\alpha, \beta)} : \Omega \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ satisfy the Carathéodory condition¹ and $k_{(\alpha, \beta)}(x, \cdot)$ are locally Lipschitz continuous (uniform in x). There exist a real constant $\bar{c}_k < \infty$ and a function $b_{(\alpha, \beta)} \in L^\infty(\Omega)$ with $b_{(\alpha, \beta)} \geq 0$ and $\|b_{(\alpha, \beta)}\|_{L^1(\Omega)} > 0$ such that $b_{(\alpha, \beta)}(\mathbf{x}) \leq k_{(\alpha, \beta)}(\mathbf{x}, \mathbf{y}) \leq \bar{c}_k$ f.a.a. $\mathbf{x} \in \Omega, \forall \mathbf{y} \in \mathbb{R}^m$.

The diffusion coefficients $D_\nu : \Omega \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ satisfy the Carathéodory condition and there exist constants $0 < \underline{c}_D, \bar{c}_D < \infty$ such that $\underline{c}_D \leq D_\nu(\mathbf{x}, \mathbf{y}) \leq \bar{c}_D$, f.a.a. $\mathbf{x} \in \Omega, \forall \mathbf{y} \in \mathbb{R}^m$ and $\nu = 1, \dots, m$.

Finally, $\bar{u}_\nu, U_\nu \in L^\infty(\Omega)$ and there exist constants $0 < \underline{c}_\bar{u}, \bar{c}_\bar{u} < \infty$ and $0 < \underline{c}_U, \bar{c}_U < \infty$ such that $\underline{c}_\bar{u} \leq \bar{u}_\nu(\mathbf{x}) \leq \bar{c}_\bar{u}$, and $\underline{c}_U \leq U_\nu(\mathbf{x}) \leq \bar{c}_U$, resp., f.a.a. $\mathbf{x} \in \Omega$ and $\nu = 1, \dots, m$.

In some results we need an assumption on the reaction order.

¹ A function $f : \Omega \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ is called *Carathéodory function* iff $z \mapsto f(x, z)$ is continuous in \mathbb{R}^m f.a.a. $x \in \Omega$ and $x \mapsto f(x, z)$ is measurable for all $z \in \mathbb{R}^m$.

Assumption A2 (Reaction order, cf. [46]). A source term of a reaction $(\alpha, \beta) \in \mathcal{R}$ is of order n , iff there exists a smallest number $n \in \mathbb{N}$ such that there exists a constant $c > 0$ with

$$\max_{\nu=1, \dots, m} \left\{ (\beta_\nu - \alpha_\nu) (a^\alpha - a^\beta) \right\} \leq c \left(1 + \sum_{\nu=1}^m a_\nu^n \right) \quad (2.1.1)$$

$$\forall a \in \mathbb{R}_+^m.$$

We assume that the source terms of all reactions are at most quadratic.

Remark 2.1.1. These technical assumptions allow us to handle a general class of reaction-diffusion systems, including heterogeneous materials and nonlinear diffusion processes. Heterogeneous materials can be found quite often in the modeling of biological or chemical processes involving different phases (see [19]). Therefore, we assume the dependence of the diffusion coefficients and the reaction rate coefficients on the spatial variable. For example, a different state of matter or a different background material leads to coefficients which are spatially dependent in a maybe nonsmooth way.

The dependence of the diffusion coefficients on the state variable is motivated by problems like those considered in Section 1.3. For example, recombination reactions of Shockley-Read-Hall, Auger type, and mole fractions, see [46, 51, 72, 121], contain reaction rate coefficients depending on the state variable. Sometimes enzymes have more than one binding site where the reactivity of a docking place is influenced by the number of free bindings (in biochemistry the behavior is called allosteric regulation). From the modeling point of view this leads to reaction coefficients which depend on the concentration of a regulating molecule, see [106]. Especially with regard to the application in Section 1.3 we assume that some reactions can vanish on subdomains. This means that the reaction rate constant becomes zero.

The assumptions on the space dimension and on the reaction order are technical to obtain existence and boundedness results in a general class of problems as in [46] for the continuous problem. Note, that only the source terms of the reaction terms are restricted, the sink terms may be large.

2.2 Weak formulation

The weak formulation of (1.4.8) is written in the variables

$$\begin{aligned} v &= (v_1, \dots, v_m) : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}^m, & \text{(potentials),} \\ a &= (a_1, \dots, a_m) : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}_+^m, & \text{(activities),} \\ u &= (u_1, \dots, u_m) : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}_+^m, & \text{(densities).} \end{aligned}$$

The (chemical) potentials and (chemical) activities lie in a space of sufficiently smooth functions. Since the reference densities may have jumps in the space the corresponding densities are elements of

the dual space, see also Section 5.3. Therefore, we work with the Gelfand triple $X \Subset Y \cong Y^* \Subset X^*$, where

$$X := H^1(\Omega, \mathbb{R}^m), \quad Y := L^2(\Omega, \mathbb{R}^m), \quad W := X \cap L^\infty(\Omega, \mathbb{R}^m). \quad (2.2.2)$$

Moreover, we define the operators $A : W \rightarrow X^*$, $E : X \rightarrow X^*$ by

$$\begin{aligned} \langle Av, w \rangle &:= \int_{\Omega} \left(\sum_{\nu=1}^m D_{\nu}(\cdot, \mathbf{e}^{\nu}) \bar{u}_{\nu} \mathbf{e}^{v_{\nu}} \nabla v_{\nu} \cdot \nabla w_{\nu} - R_{\nu}(\cdot, \mathbf{e}^{\nu}) w_{\nu} \right) dx \quad \forall w \in X, \\ \langle Ev, w \rangle &:= \sum_{\nu=1}^m \int_{\Omega} (\bar{u}_{\nu} \mathbf{e}^{v_{\nu}} w) dx \quad \forall w \in X. \end{aligned} \quad (2.2.3)$$

The operator A contains the reaction and diffusion terms of (1.4.8) and the operator E incorporates the statistical relation (1.4.3). In the setting of Assumption A1, a weak formulation of (1.4.8) can be stated as follows: Find (u, v) such that:

$$\left. \begin{aligned} u'(t) + Av(t) &= 0, \quad u(t) = Ev(t) \text{ f.a.a. } t \in \mathbb{R}_+, \quad u(0) = U, \\ u &\in H_{\text{loc}}^1(\mathbb{R}_+, X^*), \\ v &\in L_{\text{loc}}^2(\mathbb{R}_+, X) \cap L_{\text{loc}}^\infty(\mathbb{R}_+, L^\infty(\Omega, \mathbb{R}^m)). \end{aligned} \right\} \quad (2.2.P)$$

2.3 Summary of known results

The problem (2.2.P) has been investigated in various papers, see e.g. [46, 53, 62, 67]. We shortly summarize results for the continuous reaction-diffusion system (1.4.8) obtained in the cited papers. For an extended summary of continuous results obtained so far we refer to [46, 53]. The starting point in proving the existence of solutions to (2.2.P) are physical motivated a priori estimates. The stoichiometric subspace implies some invariance property of the solutions to (2.2.P). These invariants play a role during the investigation of the long time behavior of the solution. First we introduce the two convex functionals $\Phi : X \rightarrow \mathbb{R}$ and $F : X^* \rightarrow \bar{\mathbb{R}}$, which are known from chemical thermodynamics (see e.g. [46, 110]), by

$$\Phi(v) := \sum_{\nu=1}^m \int_{\Omega} \bar{u}_{\nu} (\mathbf{e}^{v_{\nu}} - 1) dx \quad \text{and} \quad F(u) := \sup_{v \in X} \{ \langle u, v \rangle - \Phi(v) \}.$$

The value $F(u(t))$ can be interpreted as free energy of solutions $(u(t), v(t))$ of the problem (2.2.P) and has the explicit representation

$$F(u) = \int_{\Omega} \sum_{\nu=1}^m \left\{ u_{\nu} \left(\ln \frac{u_{\nu}}{\bar{u}_{\nu}} - 1 \right) + \bar{u}_{\nu} \right\} dx.$$

Since we consider isothermal processes we expect the monotonous decay of the free energy along trajectories. This means for two times $0 \leq t_1 \leq t_2$ the estimate $F(u(t_2)) \leq F(u(t_1)) < \infty$ holds, see [46, Theorem 4.1]. From this explicit representation one can derive global in time a priori estimates in $L^1(\Omega)$.

For the invariance property of the solutions we define the set

$$\mathcal{U} = \{u \in X^* : (\langle u_1, 1 \rangle_{H^1}, \dots, \langle u_m, 1 \rangle_{H^1}) \in \mathcal{S}\},$$

with $\langle u_{\nu}, 1 \rangle_{H^1} = \int_{\Omega} u_{\nu} dx$ for $u \in X^* \cap L^1(\Omega, \mathbb{R}^m)$. It is well known that solutions (u, v) to (2.2.P) stay in the affine subspace $\mathcal{U} + U$, meaning

$$u(t) \in \mathcal{U} + U \quad \forall t \in \mathbb{R}_+, \quad (2.3.4)$$

see [46, 62]. Therefore, if $u^* := \lim_{t \rightarrow \infty} u(t)$ exists, then necessarily $u^* \in U + \mathcal{U}$. According to [46, 61] there exists a unique stationary solution (u^*, v^*) to

$$Av^* = 0, \quad u^* = Ev^*, \quad u^* \in U + \mathcal{U}, \quad v^* \in W. \quad (2.3.S)$$

We are only interested in problems having no steady states with zero density components.

Assumption A3. *Let*

$$\mathcal{A} := \left\{ a \in \mathbb{R}_+^m : a^{\alpha} = a^{\beta} \quad \forall (\alpha, \beta) \in \mathcal{R}, \quad u = \bar{u}a \in \mathcal{U} + U \right\} \quad (2.3.5)$$

be defined and let

$$\mathcal{A} \cap \partial \mathbb{R}_+^m = \emptyset$$

be fulfilled.

On the one hand, if (u, v) is a solution to (2.3.S) then $a = (e^{v_{\nu}})_{\nu=1}^m \in \mathcal{A}$. On the other hand if $a \in \mathcal{A}$ and $a > 0$ then (u, v) defined by $v = \ln u$, $u = \bar{u} e^v$ is a steady state of (2.2.P), that is a solution to (2.3.S). Thus Assumption A3 implies $\mathcal{A} = \{a^*\}$ and therefore (u^*, v^*) is a thermodynamic equilibrium of the system.

The free energy along solutions to (2.2.P) decays monotonously and exponentially to its equilibrium value $F(u^*)$ [57, 67]. This means there exist a constant $\lambda > 0$ such that the *relative free energy* $\Psi(u(t)) := F(u(t)) - F(u^*)$ decays exponentially

$$\Psi(u(t)) \leq e^{-\lambda t} \Psi(U) \quad \forall t \geq 0.$$

Considering regularized problems, finding a priori estimates which do not depend on the regularization level the existence of solutions to (2.2.P) is shown in [46, 63]. In detail, let $q \in \mathbb{R}_+$ such that $\bar{u}_\nu e^{-q} \leq U_\nu \leq \bar{u}_\nu e^q$ for $\nu = 1, \dots, m$. Introducing $P_q : \mathbb{R} \rightarrow [-q, q]$ by

$$P_q(y) := \begin{cases} q, & \text{if } y > q, \\ y, & \text{if } y \in [-q, q], \\ -q & \text{if } y < -q \end{cases} \quad (2.3.6)$$

the regularized versions of the operators E and A are defined by

$$\begin{aligned} \langle E_q v, w \rangle &:= \sum_{\nu=1}^m \int_{\Omega} \bar{u}_\nu e^{P_q v_\nu} w_\nu dx & \forall w \in X, \\ \langle A_q v, w \rangle &:= \sum_{\nu=1}^m \int_{\Omega} (D(\cdot, e^{P_q v}) \bar{u}_\nu \nabla e^{v_\nu}) \cdot \nabla w_\nu - \lambda_q(v) R_\nu(\cdot, e^v) \cdot w dx & \forall w \in X, \end{aligned}$$

where $\lambda_q \in C(\mathbb{R}^m; [0, 1])$ is a fixed function with the following property

$$\lambda_q(\xi) = \begin{cases} 0, & \text{if } |\xi|_\infty \geq q, \\ 1, & \text{if } |\xi|_\infty \leq q/2, \end{cases} \quad , |\xi|_\infty := \max_{\nu=1, \dots, m} |\xi_\nu|.$$

The regularized problem of (2.2.P) reads as follows: Find a tuple (u, v) such that

$$\left. \begin{aligned} u'(t) + A_q v(t) &= 0, \quad u(t) = E_q v(t) \text{ f.a.a. } t \in \mathbb{R}_+, \quad u(0) = U, \\ u &\in H_{\text{loc}}^1(\mathbb{R}_+, X^*) \text{ and } v \in L_{\text{loc}}^2(\mathbb{R}_+, X) \cap L_{\text{loc}}^\infty(\mathbb{R}_+, L^\infty(\Omega, \mathbb{R}^m)). \end{aligned} \right\} \quad (2.3.P_q)$$

The solvability of (2.3.P_q) can be proven by the help of Rothe method (discretizing (2.3.P_q) in time) and a result of operators of variational type in the sense of Lions [91, Ch. 2., Sec. 2.5], see [46, 62]. The local existence result and a priori estimates which do not depend on the regularization level imply global existence. Then a solution to (2.3.P_q) is a solution to (2.2.P) if the regularization level

is large enough. Under Assumption A2 the uniform a priori estimates are obtained by means of a technique introduced by [104]. Testing problem (2.3.P_q) with test functions

$$p e^t (z_1^{p-1}, \dots, z_m^{p-1}), \quad z_\nu = (a_\nu - \kappa)^+, \quad \kappa = \max_{\nu=1, \dots, m} \|U_\nu / \bar{u}_\nu\|_{L^\infty}$$

for $p = 2^k$, $k \geq 1$, one can obtain bounds in $\|u\|_{L^\infty(\mathbb{R}_+, L^{2^k}(\Omega, \mathbb{R}^m))}$ which are independent of the regularization level. Uniform lower bounds on the densities or the negative part of the chemical potentials can be proven by testing (2.3.P_q) with test functions

$$-p e^t (z_1^{p-1}/a_1, \dots, z_m^{p-1}/a_m), \quad z_\nu = (v_\nu + \kappa)^-, \quad \kappa = \max_{\nu=1, \dots, m} \|(\ln U_\nu / \bar{u}_\nu)^-\|_{L^\infty}$$

and using the exponential decay of the free energy obtained under Assumption A3. For the very technical details we refer to [62, 63] and in the discrete setting to Section 3.8 and Section 3.10.

Uniqueness of solutions to problem (2.2.P) can be proven by standard arguments, if the diffusion coefficients do only depend on the space variable and not on the state variables, see [62, 63]. In cases where the diffusion coefficients depend on the state variable we refer to [46]. In this work the flux term has the form $j_\nu = -d_\nu(\cdot, v_\nu, \nabla v_\nu)$, in order to include other physical effects, e.g., an electric or magnetic field. Uniqueness is then proven under the assumption that the diffusion function d_ν can be written in two parts $d_\nu(\cdot, y, w) = e^y \gamma_\nu(\cdot, w)$, where γ_ν is a strongly monotone, Lipschitz continuous Carathéodory function. The proof exploits a special distance ϱ between two solutions (u^1, v^1) and (u^2, v^2) of (2.2.P), i.e.,

$$\varrho(u^1(t), u^2(t)) := F(u^1(t)) + F(u^2(t)) - 2F\left(\frac{u^1(t) + u^2(t)}{2}\right) \quad (2.3.7)$$

and shows that $\|u^1 - u^2\|_{L^2(S, Y)}^2$ can be estimated by (2.3.7) and this distance can be bounded from above by $\|u^1 - u^2\|_{L^2(S, Y)}^2$. Hence, from Gronwall's lemma the result follows.

Remark 2.3.1.

- For other results about global existence of global weak and classic solution of reaction-diffusion systems in the setting of linear diffusion or reactions with at most quadratic source terms in higher space dimensions we refer to [111, 113] and the references therein. The starting point is an a priori estimate in L^1 obtained by the invariance property (2.3.4) of the solution. Using this estimate one applies bootstrapping arguments by duality to get estimates in higher L^p (with $p > 1$) spaces. This approach strongly relies on the smoothness and linearity of the diffusion operators. Such assumptions exclude practically relevant geometries and heterogeneous materials.
- In special cases, using the concrete structure of the underlying reaction system also some systems not fulfilling the general formulated condition of source terms of maximal order 2 can be handled,

e.g., under the 'intermediate sum condition' (see [103]), where a priori estimates for positive linear combinations of densities are obtained or in the case of cluster reactions of higher order (see [53, 65]) where in the proof of the a priori estimates simultaneously different powers of the chemical activities of the different species are used as test functions.

- *For special systems the explicit rates λ of the exponential decay toward equilibrium is exploited in [23]. The papers [44, 46, 61–63] treat also electrically charged species, such that the flux terms additionally contain drift contributions and a Poisson equation for the self consistent calculation of the electrostatic potential has to be added. The papers [46, 47, 61] additionally deal with more general state equations than (1.4.3).*
- *The problem of initial values with zero densities is intensively discussed in [53]. Moreover, the work considered couplings with ordinary differential reaction equations and reactions living only on the boundary of Ω .*

3

The discretized problem

Abstract The numerical simulation of reaction-diffusion systems is a very difficult task. The nonlinear nature and the different timescales lead to phenomena like reaction fronts, strong gradients and concentrations which are nearly zero. Taking heterogeneous materials into account increases these difficulties. An essential feature of the continuous problem is the fact that it allows a thermodynamical equilibrium, where all driving forces vanish. As a consequence, a stable discretization of the equations is necessary, which preserves the dissipative structure and the properties of the continuous problem (2.2.P) independent of the mesh and the time step. At the end of the chapter, we provide the same results which are stated in Section 2.3 for the continuous problem also for the discretized problem. All restrictions (reaction order), which are known for the continuous problem are present on the discrete level (except the time step limitation in the uniqueness result).

3.1 Introduction

In this chapter, we study the properties of a finite volume discretization on Voronoi meshes for the continuous problems.

We first introduce an implicit Euler in time and Voronoi finite volume in space discretization and define the discrete version of the continuous problem (2.2.P) in Section 3.2 and 3.3. Then, we present the first results concerning the invariance property of the solutions (2.3.4) and the asymptotic behavior of solutions, see [56].

In the analysis of finite volume discretizations it is common to adapt techniques which are well suited for investigating the continuous problem. In Section 3.4, we explore some invariance property of the solutions reflecting the stoichiometric subspace of the set of all considered reactions. In contrast to the continuous setting the local existence in Section 3.5 rest on the conservation of the number of atoms. For elementary reversible chemical reactions this property is naturally fulfilled.

Moreover the invariance property plays an important role in the proof of the existence of a thermodynamic equilibrium in Section 3.7, and also in the study of the long time behavior of the solution in Section 3.9. In addition in Section 3.6, we show that physically motivated arguments lead to a priori estimates which are important in the proof of global upper and lower bounds in the Sections 3.8 and 3.10.

In Section 3.8, we present a Moser iteration for the discrete problem in order to achieve (independent of the mesh quality and the time step) uniform global upper bounds. For this purpose we establish a discrete version of a Gagliardo-Nirenberg inequality in Appendix A.2. The proof of the lower bounds in Section 3.10 rests heavily on results of the thermodynamic equilibrium and the asymptotic behavior as time goes to infinity in Section 3.9. The lower bounds for the concentrations are obtained by the Moser iteration technique and are independent of the mesh quality and the time step. In Section 3.11 we provide a uniqueness result in the case of only spatially dependent diffusion coefficients and a restriction on the time step. Starting from the results on the global stability of the discretization, we prove strong convergence of the discretized nonlinear evolution problem (3.3. P_D) to a weak solution of the continuous problem (2.2.P).

3.2 Voronoi finite volume discretization

In the following, we work with boundary conforming Delaunay-Voronoi meshes (see [122]), which represent one class of admissible finite volume meshes [32]. The Voronoi mesh \mathcal{M} is derived as dual grid of the boundary conforming Delaunay triangulation \mathcal{T} .

Our notation is basically taken from [56] and visualized in Figure 3.1.

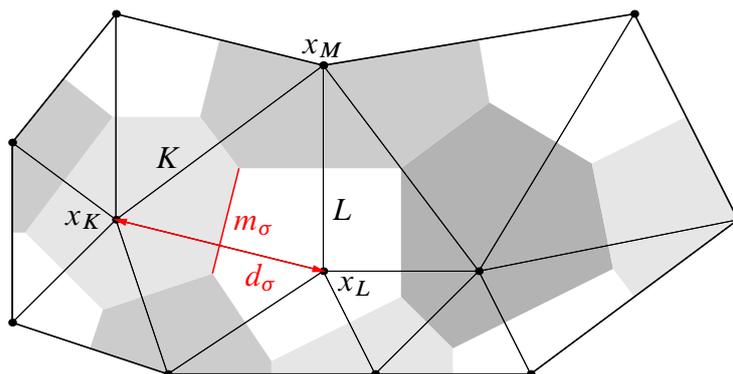


Figure 3.1: Notation of Voronoi meshes $\mathcal{M} = (\mathcal{P}, \mathcal{V}, \mathcal{E})$. The lines represent the edges of the triangles of the Delaunay mesh and the gray areas represents the Voronoi control volumes

Let Ω be an open, bounded, polygonal subset of \mathbb{R}^2 . A Voronoi mesh is defined as triple $\mathcal{M} = (\mathcal{P}, \mathcal{V}, \mathcal{E})$. Here, \mathcal{P} denotes a family of *grid points* in $\bar{\Omega}$, \mathcal{V} denotes a family of *Voronoi control volumes*

and \mathcal{E} denotes a family of *faces*. A face is a interval of \mathbb{R} . The number of grid points is denoted by $M = \#\mathcal{P}$.

The corresponding control volume K of each grid point $x_K \in \mathcal{P}$ is defined by

$$K = \{x \in \Omega : |x - x_K| < |x - x_L| \quad \forall x_L \in \mathcal{P}, x_L \neq x_K\}.$$

The set of all neighboring control volumes of K is denoted by $\mathcal{N}_{\mathcal{V}}(K)$. The Lebesgue measure of each control volume K is denoted by $|K|$ and the mesh size of \mathcal{M} by

$$\text{size}(\mathcal{M}) = \sup_{K \in \mathcal{V}} \text{diam}(K).$$

For two different $K, L \in \mathcal{V}$ the one-dimensional Lebesgue measure of $\overline{K} \cap \overline{L}$ is either zero or $\overline{K} \cap \overline{L} = \overline{\sigma}$ for one $\sigma \in \mathcal{E}$. Here the symbol $\sigma = K|L$ denotes the one-dimensional face between the control volumes K and L and m_{σ} is its Lebesgue measure.

We introduce the subset $\mathcal{E}_{int} \subset \mathcal{E}$ containing all interior faces. Further, we introduce for all $K \in \mathcal{V}$ the subset $\mathcal{E}_K \subset \mathcal{E}_{int}$, such that $\forall \sigma \in \mathcal{E}_K \exists L \in \mathcal{N}_{\mathcal{V}}(K) : \sigma = \overline{L} \cap \overline{K}$.

The Euclidian distance between two neighboring grid points $x_K, x_L \in \mathcal{P}$ over the face $\sigma = K|L \in \mathcal{E}_{int}$ is denoted by d_{σ} .

Remark 3.2.1 (see Definition 1. and Remark 1. [48]). *Let $\mathcal{M} = (\mathcal{P}, \mathcal{V}, \mathcal{E})$ be a Voronoi mesh. The dual $(\mathcal{P}, \mathcal{T})$ of a Voronoi mesh consists of a family \mathcal{P} of grid points in $\overline{\Omega}$ and a family \mathcal{T} of triangles $T = (x_K, x_L, x_M) \in \mathcal{P}^3$ and is called Delaunay triangulation if $\Omega = \cup_{T \in \mathcal{T}} T$ and iff the circumcircle of each triangle does not contain any other vertices $x_N \in \mathcal{P}$ with $x_N \in T$ in its interior. If in addition each triangle $T \in \overline{\Omega}$ has its circumcenter in $\overline{\Omega}$, then a Delaunay triangulation is called boundary conforming Delaunay.*

A Voronoi mesh is usually constructed from a Delaunay triangulation by the intersection of half spaces in \mathbb{R}^2 and hence each Voronoi cell is convex. If the domain consists of more than one subdomain (material) and the interfaces are aligned to the triangle edges of the Delaunay mesh, then all subdomain triangulations have to be boundary conforming, i.e. the circumcenter of a triangle consisting to one subdomain (material) have to be inside the subdomain, see [48].

Definition 3.2.2 (Discrete Norms, see [56]).

Let Ω be an open bounded, polygonal subset of \mathbb{R}^2 and $\mathcal{M} = (\mathcal{P}, \mathcal{V}, \mathcal{E})$ a Voronoi mesh.

- *The symbol $X_{\mathcal{V}}(\mathcal{M})$ denotes the set of all piecewise constant functions from Ω to \mathbb{R} which are constant on every Voronoi control volume $K \in \mathcal{V}$. The constant value of $w_h \in X_{\mathcal{V}}(\mathcal{M})$ on the control volume $K \in \mathcal{V}$ is denoted by w_K .*

- Let $p \geq 1$. The discrete L^p - norm of $w_h \in X_{\mathcal{V}}(\mathcal{M})$ is defined by

$$\|w_h\|_{L^p} = \left(\sum_{K \in \mathcal{V}} |K| |w_K|^p \right)^{1/p}.$$

- The discrete H^1 -seminorm of $w_h \in X_{\mathcal{V}}(\mathcal{M})$ is defined by

$$|w_h|_{H^1, \mathcal{M}}^2 = \sum_{\sigma=K|L \in \mathcal{E}_{int}} T_{\sigma} |w_K - w_L|^2, \quad T_{\sigma} := \frac{m_{\sigma}}{d_{\sigma}}.$$

Here w_K and w_L are the constant values of w_h in the control volumes K and L . The term T_{σ} is the so called transmissibility across the face $\sigma = K|L \in \mathcal{E}_{int}$. The discrete H^1 - norm is given by $\|w_h\|_{H^1, \mathcal{M}}^2 = |w_h|_{H^1, \mathcal{M}}^2 + \|w_h\|_{L^2}^2$.

We prescribe the approximation of a function $f : \Omega \times \mathbb{R}^m \rightarrow \mathbb{R}$ by

$$f_K(\cdot) := \frac{1}{|K|} \int_K f(x, \cdot) dx,$$

where $K \in \mathcal{V}$. In this context we introduce the approximation of the diffusion coefficients and the reaction terms on a Voronoi cell $K \in \mathcal{V}$ by

$$D_{\nu K}(\cdot) = \frac{1}{|K|} \int_K D_{\nu}(x, \cdot) dx, \quad R_{\nu K}(\cdot) = \frac{1}{|K|} \int_K R_{\nu}(x, \cdot) dx, \quad (3.2.1)$$

and analogously the approximation of $k_{(\alpha, \beta)}$.

The corresponding piecewise constant function can be estimated from above and below by the upper and lower bound of the continuous function. We define the reference density $\bar{u}_{\nu K}$ and the density $u_{\nu K}$ being constant on a control volume $K \in \mathcal{V}$, and the mass $u_{\nu}^{(K)}$ of the ν - th species in $K \in \mathcal{V}$ by

$$\bar{u}_{\nu K} = \frac{1}{|K|} \int_K \bar{u}_{\nu}(x) dx, \quad u_{\nu K} = \bar{u}_{\nu K} e^{v_{\nu K}} \text{ and } u_{\nu}^{(K)} = |K| u_{\nu K}.$$

For every species X_{ν} , $\nu = 1, \dots, m$, we introduce the discrete initial values by

$$U_{\nu}^{(K)} := \int_K U_{\nu}(x) dx, \quad K \in \mathcal{V}.$$

The space-discrete version of the continuous problem (2.2.P) is formally obtained by testing with the characteristic function of K . Using Gauss theorem, we derive the approximated flux term

$$\int_K \nabla \cdot \mathbf{j}_\nu dx = \int_{\partial K} \mathbf{j}_\nu \cdot \mathbf{n}_K d\Gamma \approx \sum_{\sigma=K|L \in \mathcal{E}_K} -T_\sigma Y_\nu^\sigma Z(v_{\nu L}, v_{\nu K})(v_{\nu L} - v_{\nu K}),$$

where

$$Z(x, y) = \begin{cases} \frac{e^x - e^y}{x - y}, & \text{for } x \neq y, \\ e^x, & \text{for } x = y \end{cases}, \quad x, y \in \mathbb{R}, \quad (3.2.2)$$

represents the logarithmic mean value of e^x in the interval $[x, y]$. In the following we write $Z_\nu^\sigma = Z(v_{\nu L}, v_{\nu K})$ for $\sigma = K|L \in \mathcal{E}_{int}$ and $D_{\nu K} = D_{\nu K}(e^{v_{1K}}, \dots, e^{v_{mK}})$. With this definition of Z_ν^σ it is possible to switch between a gradient in potentials and activities, i.e., the discrete version of $\nabla a_\nu = a_\nu \nabla v_\nu$ holds. The symbol Y_ν^σ defines some averaging of $D_{\nu K} \bar{u}_\nu$ over the edge $\sigma = K|L$, which is symmetric in K and L . Possible averagings are, e.g.,

$$Y_\nu^\sigma = \frac{D_{\nu K} \bar{u}_{\nu K} + D_{\nu L} \bar{u}_{\nu L}}{2}, \quad Y_\nu^\sigma = \frac{D_{\nu K} + D_{\nu L}}{2} \frac{\bar{u}_{\nu K} + \bar{u}_{\nu L}}{2}, \quad \sigma = K|L.$$

For another averaging which is exact along an aligned face we refer to [31]. In the sequel all results are independent of the particular chosen Y_ν^σ .

Following [54] we use the notation

$$\begin{aligned} \mathbf{u}_\nu &= (u_\nu^{(K)})_{K \in \mathcal{V}}, & \mathbf{u} &= (u_1, \dots, u_m), & \mathbf{u}_K &= (u_{\nu K})_{\nu=1}^m, \\ \mathbf{v}_\nu &= (v_{\nu K})_{K \in \mathcal{V}}, & \mathbf{v} &= (v_1, \dots, v_m), & \mathbf{v}_K &= (v_{\nu K})_{\nu=1}^m, \\ \mathbf{U}_\nu &= (U_\nu^{(K)})_{K \in \mathcal{V}}, & \mathbf{U} &= (U_1, \dots, U_m), \\ \mathbf{a}_K &= (e^{v_{\nu K}})_{\nu=1}^m, & \mathbf{a}_\nu &= (e^{v_{\nu K}})_{K \in \mathcal{V}}, & \nu &= 1, \dots, m. \end{aligned}$$

Furthermore, we define the scalar products

$$\langle \mathbf{u}_\nu, \mathbf{v}_\nu \rangle_{\mathbb{R}^M} = \sum_{K \in \mathcal{V}} u_\nu^{(K)} v_{\nu K}, \quad \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^{Mm}} = \sum_{\nu=1}^m \langle \mathbf{u}_\nu, \mathbf{v}_\nu \rangle_{\mathbb{R}^M}.$$

3.3 Time discretization

Since we deal with an evolution problem, we need a discretization of the time interval. By doing this, we can formulate a fully discretized version of the continuous problem (2.2.P). We use the implicit Euler method, which respects the dissipative structure of parabolic equations [28]. It is well known,

that the implicit Euler method is A-stable¹ therefore there exist no limitations on the time step size to ensure numerical stability.

Definition 3.3.1 (Time discretization).

A time discretization of \mathbb{R}_+ is defined as a strictly increasing sequence of real numbers $(t_n)_{n \in \mathbb{N}} \subset \mathbb{R}_+$ with $t_0 = 0$ and $t_n \rightarrow \infty$ for $n \rightarrow \infty$. The time step is defined by

$$t_\delta^{(n)} = t_n - t_{n-1} < \infty, \quad \text{for } n \in \mathbb{N}.$$

A discretization of the whole domain $Q = \mathbb{R}_+ \times \Omega$ is defined by the tuple $\mathcal{D} = (\mathcal{M}, (t_n)_{n \in \mathbb{N}})$. To indicate the relation to the continuous problem (2.2.P) we give a variational form of the discrete problem. We introduce the operator $\widehat{E} : \mathbb{R}^{Mm} \rightarrow \mathbb{R}_+^{Mm}$ by

$$\widehat{E}\mathbf{v} = \left((\bar{u}_{\nu K} e^{v_{\nu K}} |K|)_{\nu=1, \dots, m} \right)_{K \in \mathcal{V}},$$

which maps in every control volume the chemical potential of the species to its mass. Furthermore we define $\widehat{A} : \mathbb{R}^{Mm} \rightarrow \mathbb{R}^{Mm}$ by

$$\widehat{A}\mathbf{v} = \left(\sum_{\sigma=K|L \in \mathcal{E}_K} -T_\sigma Y_\nu^\sigma Z_\nu^\sigma (v_{\nu L} - v_{\nu K}) - |K| R_{\nu K}(\mathbf{e}^{\mathbf{v}_K}) \right)_{\substack{K \in \mathcal{V}, \\ \nu=1, \dots, m}}. \quad (3.3.3)$$

Using these definitions we can state the discrete version of (2.2.P) by: Find a pair (\mathbf{u}, \mathbf{v}) such that

$$\left. \begin{aligned} \frac{\mathbf{u}(t_n) - \mathbf{u}(t_{n-1})}{t_\delta^{(n)}} + \widehat{A}\mathbf{v}(t_n) &= \mathbf{0}, \quad \mathbf{u}(t_n) = \widehat{E}\mathbf{v}(t_n), \quad n \geq 1 \\ \mathbf{u}(0) &= \mathbf{U}. \end{aligned} \right\} \quad (3.3.PD)$$

As long as the meaning is clear, we will neglect the current time step t_n and shortly write $u_{\nu K}$ instead of $u_{\nu K}(t_n)$. Using this notation, for every time step, in every control volume $K \in \mathcal{V}$ and $\nu = 1, \dots, m$ we have to solve the following fully implicit system of nonlinear equations

$$|K| \frac{u_{\nu K}(t_n) - u_{\nu K}(t_{n-1})}{t_\delta^{(n)}} = \sum_{\sigma=K|L \in \mathcal{E}_K} -T_\sigma Y_\nu^\sigma Z_\nu^\sigma (v_{\nu L} - v_{\nu K}) - |K| R_{\nu K}(\mathbf{e}^{\mathbf{v}_K})$$

¹ A numerical method for the solution of an initial value problem is called A-stable iff the method applied to the Dahlquist's test equation $\dot{y} = \lambda y$, $y(0) = 1$ creates for all $\lambda \in \mathbb{C}$ with negative real part a monotone decreasing sequence of approximated solutions independent of the time step, i.e., the stability region of the method covers the complete left half-plane, see [24, 71, 124].

with $u_{\nu K}(t_n) = \bar{u}_{\nu K} e^{v_{\nu K}}$. The discrete variational form of \hat{A} is given by

$$\begin{aligned} \left\langle \hat{A}\mathbf{v}, \mathbf{w} \right\rangle_{\mathbb{R}^{Mm}} &= \sum_{\nu=1}^m \sum_{\sigma=K|L \in \mathcal{E}_{int}} T_{\sigma} Y_{\nu}^{\sigma} Z_{\nu}^{\sigma} (v_{\nu L} - v_{\nu K})(w_{\nu L} - w_{\nu K}) \\ &\quad - \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| R_{\nu K} (e^{v_{\nu K}}) w_{\nu K} \quad \forall \mathbf{v}, \mathbf{w} \in \mathbb{R}^{Mm}. \end{aligned} \quad (3.3.4)$$

For $n \in \mathbb{N}$ we associate the discrete (vectorial) solutions $(\mathbf{u}(t_n), \mathbf{v}(t_n))$ to (3.3. P_D) to piecewise constant functions $(u_h(t_n), v_h(t_n))$ and call them also solutions to (3.3. P_D).

3.4 Invariants of the system

In analogy to the continuous setting we define the subspaces

$$\hat{\mathcal{U}} = \{ \mathbf{u} \in \mathbb{R}^{Mm} : (\langle \mathbf{u}_{\nu}, \mathbf{1} \rangle_{\mathbb{R}^M})_{\nu=1}^m \in \mathcal{S} \} \quad (3.4.5)$$

and

$$\hat{\mathcal{U}}^{\perp} = \{ \mathbf{v} \in \mathbb{R}^{Mm} : \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^{Mm}} = 0 \forall \mathbf{u} \in \hat{\mathcal{U}} \}. \quad (3.4.6)$$

An equivalent characterization of $\hat{\mathcal{U}}^{\perp}$ is given by

$$\hat{\mathcal{U}}^{\perp} = \{ \mathbf{v} \in \mathbb{R}^{Mm} : v_{\nu K} = \hat{v}_{\nu} \forall K \in \mathcal{V}, \nu = 1, \dots, m, (\hat{v}_{\nu})_{\nu=1}^m \in \mathcal{S}^{\perp} \}, \quad (3.4.7)$$

i.e., $\hat{\mathcal{U}}^{\perp}$ consists of vectors that are constant on all control volumes for every species. Every element of \mathcal{S}^{\perp} generates an invariant of the reaction-diffusion system. From the definition of \hat{A} it follows immediately that

$$\left\langle \hat{A}\mathbf{v}, \mathbf{w}^{\perp} \right\rangle_{\mathbb{R}^{Mm}} = 0 \quad \forall \mathbf{w}^{\perp} \in \hat{\mathcal{U}}^{\perp} \text{ and } \mathbf{v} \in \mathbb{R}^{Mm}. \quad (3.4.8)$$

Lemma 3.4.1 (Invariance property, see [56, Lemma 3.2]).

We assume A1. Any solution (\mathbf{u}, \mathbf{v}) of the discrete Problem (3.3. P_D) fulfills

$$\mathbf{u}(t_n) \in \hat{\mathcal{U}} + \mathbf{U} \quad \forall n \geq 1.$$

Proof. Let (\mathbf{u}, \mathbf{v}) be a solution of (3.3. P_D). Testing (3.3. P_D) with an arbitrary $\mathbf{w}^\perp \in \widehat{\mathcal{U}}^\perp$ and using (3.4.8) we find

$$\begin{aligned} \left\langle \mathbf{u}(t_N) - \mathbf{U}, \mathbf{w}^\perp \right\rangle_{\mathbb{R}^{Mm}} &= \sum_{n=1}^N \left\langle \mathbf{u}(t_n) - \mathbf{u}(t_{n-1}), \mathbf{w}^\perp \right\rangle_{\mathbb{R}^{Mm}} \\ &= - \sum_{n=1}^N t_\delta^{(n)} \left\langle \widehat{A}\mathbf{v}, \mathbf{w}^\perp \right\rangle_{\mathbb{R}^{Mm}} = 0. \end{aligned}$$

Since (3.4.8) holds the desired result follows. \square

3.5 Local existence

Next we define the operator $\widehat{B} : \mathbb{R}^{Mm} \rightarrow \mathbb{R}^{Mm}$, by

$$\widehat{B}\mathbf{u} = \left(\sum_{\sigma=K|L \in \mathcal{E}_K} -T_\sigma Y_\nu^\sigma \left(\frac{u_{\nu L}}{\bar{u}_{\nu L}} - \frac{u_{\nu K}}{\bar{u}_{\nu K}} \right) - |K| R_{\nu K} \left(\frac{u_{\nu K}}{\bar{u}_{\nu K}} \right) \right)_{K \in \mathcal{V}, \nu=1, \dots, m},$$

for all $\mathbf{u} \in \mathbb{R}^{Mm}$. The solvability of (3.3. P_D) can be proved by the investigation of the solvability of the following problem: Find a positive $\mathbf{u} \in \mathbb{R}^{Mm}$ such that

$$\left. \begin{aligned} \frac{\mathbf{u}(t_n) - \mathbf{u}(t_{n-1})}{t_\delta^{(n)}} + \widehat{B}\mathbf{u}(t_n) &= \mathbf{0}, \quad n \geq 1, \\ \mathbf{u}(t_0) &= \mathbf{U} \end{aligned} \right\}. \quad (3.5.P_{\widehat{D}})$$

The relation between \widehat{A} and \widehat{B} is given by

$$\widehat{B}\mathbf{u} = \widehat{A}(\ln(\mathbf{u}/\bar{\mathbf{u}})) \quad \forall \mathbf{0} < \mathbf{u} \in \mathbb{R}^{Mm}. \quad (3.5.9)$$

In particular, it follows from (3.4.8) that

$$\left\langle \widehat{B}\mathbf{u}, \mathbf{w}^\perp \right\rangle_{\mathbb{R}^{Mm}} = 0 \quad \forall \mathbf{w}^\perp \in \widehat{\mathcal{U}}^\perp \text{ and } \mathbf{u} \in \mathbb{R}^{Mm}. \quad (3.5.10)$$

Existence of local solutions is proven under additional assumptions with respect to the reaction terms. We assume that:

Assumption A4. We assume the quasi positivity of the reaction terms, i.e., for all positive $\mathbf{a} \in \mathbb{R}^m$

$$R_{\nu K}(\cdot, (a_1, \dots, a_{\nu-1}, 0, a_{\nu+1}, \dots, a_m)) \geq 0 \quad \forall \nu = 1, \dots, m \text{ and } K \in \mathcal{V} \quad (3.5.11)$$

holds, see [12, 111]. Furthermore we assume conservation of the number of atoms

$$\exists \mathbf{s}^\perp \in S^\perp : \mathbf{s}^\perp > \mathbf{0}, \quad (3.5.12)$$

see [51, (Th₂)].

From the quasi positivity we deduce if $u_{\nu K} = 0$ for one $K \in \mathcal{V}$ and one $\nu = 1, \dots, m$ that

$$\begin{aligned} (\widehat{B}\mathbf{u})_{\nu K} &= \sum_{\sigma=K|L \in \mathcal{E}_K} -T_\sigma Y_\nu^\sigma \left(\frac{u_{\nu L}}{\bar{u}_{\nu L}} - 0 \right) \\ &- |K| R_{\nu K} \left(\cdot, \left(\frac{u_{1K}}{\bar{u}_{1K}}, \dots, \frac{u_{\nu-1K}}{\bar{u}_{\nu-1K}}, 0, \frac{u_{\nu+1K}}{\bar{u}_{\nu+1K}}, \dots, \frac{u_{mK}}{\bar{u}_{mK}} \right) \right) \leq 0. \end{aligned} \quad (3.5.13)$$

Therefore from (3.5.P_D) we find $u_{\nu K}(t_n) - u_{\nu K}(t_{n-1}) \geq 0$, which means that zero concentrations are increased by the system to positive concentrations.

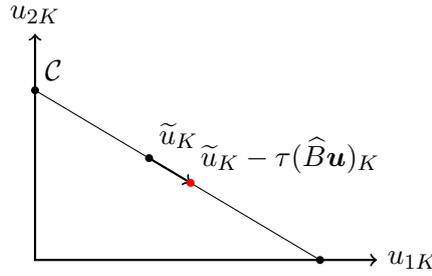


Figure 3.2: Illustration of the definition (3.5.15) and $\theta(\mathbf{u}, s)$ for the reaction $X_1 \rightleftharpoons X_2$ in a cell $K \in \mathcal{V}$. If we are on the boundary one of the components U_{1K} or U_{2K} must be zero

Remark 3.5.1. The quasi positivity assumption of the reaction terms is naturally fulfilled by reversible mass action type reactions as defined in (1.4.6). Since R_ν is symmetric in α and β , we assume without loss of generality $\alpha_\nu \geq \beta_\nu \geq 0$. Then $\beta_\nu - \alpha_\nu \leq 0$ and

$$R_{(\alpha, \beta)}(\mathbf{a}) = k_{(\alpha, \beta)} \left(\prod_{j=1}^m a_j^{\alpha_j} - \prod_{j=1}^m a_j^{\beta_j} \right) \leq 0,$$

holds if $a_\nu = 0$ and $\alpha_\nu \geq \beta_\nu > 0$. Hence $R_\nu(\mathbf{a}) \geq 0$ holds since the reaction rate coefficients $k_{(\alpha, \beta)}$ are nonnegative.

Lemma 3.5.2. We assume A1 and A4. Let $\tilde{\mathbf{u}} \in \widehat{U} + \mathbf{U}$ with $\tilde{\mathbf{u}} > \mathbf{0}$. Then for all $s > 0$, there exists an $\mathbf{u} \in \mathbb{R}^{Mm}$ such that

$$\mathbf{u} = \tilde{\mathbf{u}} - s\widehat{B}\mathbf{u}, \quad (3.5.14)$$

and $\mathbf{u} > \mathbf{0}$. Furthermore $\mathbf{u} \in \widehat{\mathcal{U}} + \mathbf{U}$.

Proof. In the following, we use Brouwer's fixed point theorem to deduce the existence of a solution. We define the set of all densities which fulfill the same invariants as the initial concentration \mathbf{U}

$$\mathcal{C} := \left\{ \mathbf{u} \in \mathbb{R}^{Mm} : \mathbf{u} \geq \mathbf{0} \wedge \mathbf{u} \in \widehat{\mathcal{U}} + \mathbf{U} \right\}, \quad (3.5.15)$$

see Figure 3.2. The main point of the proof is to show that the fixed point is positive. Since \mathcal{C} is the intersection of an affine space with Mm half spaces of nonnegative densities, the set \mathcal{C} is convex. By construction, \mathcal{C} is also a closed set. Because of the nonnegativity of the elements of \mathcal{C} and the fulfillment of the invariants, it follows the boundedness of \mathcal{C} . Also by construction, the boundary of \mathcal{C} possesses at least one component that is zero. From (3.5.12), we deduce the existence of a vector $\mathbf{s}^\perp \in \mathcal{S}^\perp$, with only positive entries. As a consequence of the definition of $\widehat{\mathcal{U}}^\perp$, we conclude the existence of $\mathbf{w}^\perp \in \widehat{\mathcal{U}}^\perp$ with $w_{\nu K}^\perp = s_\nu^\perp$ for all $K \in \mathcal{V}, \nu = 1, \dots, m$. Since $\tilde{\mathbf{u}} \in \widehat{\mathcal{U}} + \mathbf{U}$ we obtain from (3.5.10) that $\langle \mathbf{u} - \mathbf{U}, \mathbf{w}^\perp \rangle_{\mathbb{R}^{Mm}} = 0$ for all $\mathbf{u} \in \mathcal{C}$ and particular with (3.5.10) that

$$\left\langle \tilde{\mathbf{u}} - \tau \widehat{B}\mathbf{u}, \mathbf{w}^\perp \right\rangle_{\mathbb{R}^{Mm}} = \left\langle \mathbf{U}, \mathbf{w}^\perp \right\rangle_{\mathbb{R}^{Mm}} \quad \forall \mathbf{w}^\perp \in \widehat{\mathcal{U}}^\perp.$$

We define $\theta : \mathcal{C} \times [0, \infty) \rightarrow \mathbb{R}$ by

$$\theta(\mathbf{u}, s) := \sup_{\tau \in [0, s]} \left\{ \tau : \tilde{\mathbf{u}} - \tau \widehat{B}\mathbf{u} \in \mathcal{C} \right\}.$$

Since $\tilde{\mathbf{u}} \in \mathcal{C}$ holds, and \mathcal{C} is bounded, the function θ is well defined. Using the convexity of \mathcal{C} , we deduce the continuity of θ . Hence, the function $\varphi_s : \mathcal{C} \rightarrow \mathcal{C}$ with

$$\varphi_s(\mathbf{u}) = \tilde{\mathbf{u}} - \theta(\mathbf{u}, s) \widehat{B}\mathbf{u} \quad (3.5.16)$$

is continuous for every $s > 0$, and the function $\theta(\mathbf{u}, s)$ ensures that all components of $\varphi_s(\mathbf{u})$ are nonnegative, i.e., $\varphi_s(\mathbf{u}) \in \mathcal{C}$ holds. Using Brouwer's fixed point theorem, we conclude the existence of a nonnegative fixed point \mathbf{u} of φ_s for all $s > 0$. Assuming one or more components of \mathbf{u} are zero, then by (3.5.13) we find that these components of $-\theta(\mathbf{u}, s) \widehat{B}\mathbf{u}$ are nonnegative, which leads to a contradiction together with $\tilde{\mathbf{u}} > \mathbf{0}$. Therefore, the fixed point is not only nonnegative, but positive. But then, the fixed point is not on the boundary of \mathcal{C} and from that it follows that $\theta(\mathbf{u}, s) = s$ must hold. This means that the fixed point of (3.5.16) is also fixed point of (3.5.14). \square

By induction we conclude:

Theorem 3.5.3. *We assume A1 and A4. Let $\mathbf{U} = \mathbf{u}(t_0) > \mathbf{0}$. Then for all $t_n > 0$ there exists at least one solution $\mathbf{u}(t_n) > \mathbf{0}$ with $\mathbf{u}(t_n) \in \widehat{\mathcal{U}} + \mathbf{U}$ of the nonlinear equation (3.5.P \tilde{D}).*

This implies:

Theorem 3.5.4 (Local existence).

Let the Assumptions A1 and A4 be fulfilled. Then for every initial value $\mathbf{U} = \mathbf{u}(t_0) > \mathbf{0}$, there exists a solution $(\mathbf{u}(t_n), \mathbf{v}(t_n))$, $n > 0$, of the discrete Problem (3.3. P_D).

Proof. From Theorem (3.5.3) we conclude the existence of a $\mathbf{0} < \mathbf{u}(t_n) \in \mathbb{R}^{Mm}$ which is a solution of (3.5. $P_{\bar{D}}$). Since $\mathbf{u}(t_n) > 0$ and $\mathbf{v}(t_n) = \ln(\mathbf{u}(t_n)/\bar{\mathbf{u}})$ this delivers by (3.5.9) a solution to (3.3. P_D), \square

Remark 3.5.5. Local existence results for reaction-diffusion systems not fulfilling (3.5.11) and (3.5.12) can be proven by investigating a "regularized" problem which arises from (3.3. P_D) by cutting off the nonlinearities in a suitable way at a certain level and using the theory of pseudomonotone operators, see [46, 63, 128]. Similarly, in the proof of Lemma 3.5.2 we cut off the time step s by using the function $\theta(\cdot, \cdot)$.

3.6 Physically motivated estimates

We show that physical motivated arguments provide a priori estimates for solutions to (3.3. P_D). In analogy to the continuous problem the (discrete) *free energy* is a Lyapunov function of the discrete problem and the *dissipation* of (3.3. P_D) is introduced. The considered process is isolated therefore the decay of the free energy along trajectories is expected. The dissipative structure is the starting point in establishing a priori estimates in L^1 which provide the first step in the proof of global upper bounds on the solutions to (3.3. P_D) in Section 3.8.

In the literature the term free energy is often denoted as entropy [22, 23]. All results are based on the articles [56, 57]. In Section A.4 we rephrase basic definitions and properties from convex analysis which we used in this part. We also refer to the classical textbooks [27, 77, 115, 128] for basic notation and results.

First we define the discrete potential $\hat{\Phi} : \mathbb{R}^{Mm} \rightarrow \bar{\mathbb{R}}$ by

$$\hat{\Phi}(\mathbf{v}) = \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} \bar{u}_{\nu K} (e^{v_{\nu K}} - 1) |K|. \quad (3.6.17)$$

From [56] we rephrase that $\hat{\Phi}$ is proper, continuous, and strictly convex, hence subdifferentiable and $\{\hat{E}\mathbf{v}\} = \partial\hat{\Phi}(\mathbf{v})$ holds.

The conjugate functional of $\hat{\Phi}$ is defined by $\hat{F} : \mathbb{R}^{Mm} \rightarrow \bar{\mathbb{R}}$,

$$\hat{F}(\mathbf{u}) := \sup_{\mathbf{w} \in \mathbb{R}^{Mm}} \left\{ \langle \mathbf{u}, \mathbf{w} \rangle_{\mathbb{R}^{Mm}} - \hat{\Phi}(\mathbf{w}) \right\}. \quad (3.6.18)$$

It is proper, lower semicontinuous, and convex (in \mathbf{u}) and $\mathbf{u} = \widehat{E}\mathbf{v} \in \partial\widehat{\Phi}(\mathbf{v})$ holds iff $\mathbf{v} \in \partial\widehat{F}(\mathbf{u})$. Together with $\mathbf{u} = \widehat{E}\mathbf{v}$ the value of \widehat{F} at \mathbf{u} can be explicitly calculated by

$$\widehat{F}(\mathbf{u}) = \left\langle \widehat{E}\mathbf{v}, \mathbf{v} \right\rangle_{\mathbb{R}^{Mm}} - \widehat{\Phi}(\mathbf{v}) = \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} (u_{\nu K}(v_{\nu K} - 1) + \bar{u}_{\nu K}) |K|.$$

For a given argument $\mathbf{u} \in \mathbb{R}^{Mm}$ the value of $\widehat{F}(\mathbf{u})$ can be interpreted as the *free energy* of the state \mathbf{u} . Using the elementary inequality $\ln s \geq 1 - 1/s$, $s > 0$ we observe the nonnegativity of the free energy

$$\widehat{F}(\mathbf{u}) = \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} \left(u_{\nu K} \left(\ln \frac{u_{\nu K}}{\bar{u}_{\nu K}} - 1 \right) + \bar{u}_{\nu K} \right) |K| \geq 0. \quad (3.6.19)$$

Finally the discrete dissipation as functional $\widehat{D} : \mathbb{R}^{Mm} \rightarrow \mathbb{R}$ is introduced by

$$\widehat{D}(\mathbf{v}) := \left\langle \widehat{A}\mathbf{v}, \mathbf{v} \right\rangle_{\mathbb{R}^{Mm}}, \quad \mathbf{v} \in \mathbb{R}^{Mm}. \quad (3.6.20)$$

As a consequence of (3.3.4), the definition of the reaction terms (1.4.6), the nonnegativity of Y_ν^σ and Z^σ (see also Assumption A1) we observe by $(e^x - e^y)(x - y) \geq 0$ for $x, y \in \mathbb{R}$ the nonnegativity of the dissipation, i.e., for all $\mathbf{v} \in \mathbb{R}^{Mm}$ holds

$$\begin{aligned} \widehat{D}(\mathbf{v}) &= \left\langle \widehat{A}\mathbf{v}, \mathbf{v} \right\rangle_{\mathbb{R}^{Mm}} \\ &= \sum_{\nu=1}^m \sum_{\sigma=K|L \in \mathcal{E}_{int}} T_\sigma Y_\nu^\sigma Z^\sigma (v_{\nu L} - v_{\nu K})^2 \\ &\quad + \sum_{K \in \mathcal{V}} |K| \sum_{(\alpha, \beta) \in \mathcal{R}} k_{(\alpha, \beta)K} \left(e^{\alpha \cdot \mathbf{v}_K} - e^{\beta \cdot \mathbf{v}_K} \right) (\alpha \cdot \mathbf{v}_K - \beta \cdot \mathbf{v}_K) \geq 0, \end{aligned} \quad (3.6.21)$$

see [56, 57]. Together with the nonnegativity of the dissipation the following result can be stated:

Lemma 3.6.1 (Monotonicity of the free energy, see [56, Lemma 3.1]).

We assume A1. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n \in \mathbb{N}})$ be a discretization of $\Omega \times \mathbb{R}_+$ and (\mathbf{u}, \mathbf{v}) be a solution to (3.3.P_D). Then

$$\widehat{F}(\mathbf{u}(t_{n_2})) + \sum_{n=n_1+1}^{n_2} t_\delta^{(n)} \widehat{D}(\mathbf{v}(t_n)) \leq \widehat{F}(\mathbf{u}(t_{n_1})) \leq \widehat{F}(\mathbf{U}) \text{ for } 0 \leq t_{n_1} < t_{n_2}$$

holds, i.e., the free energy decays monotonously along all solutions to (3.3.P_D). Moreover, it holds

$$\sum_{\nu=1}^m \|u_{\nu, h}(t_n)\|_{L^1} \leq 2 \left(\widehat{F}(\mathbf{U}) + \sum_{\nu=1}^m \|\bar{u}_{\nu, h}\|_{L^1} \right) \quad \forall n \geq 1.$$

Proof. For $\mathbf{u} = \widehat{E}\mathbf{v}$ with $\mathbf{v} \in \mathbb{R}^{Mm}$ we have $\mathbf{v} \in \partial\widehat{F}(\mathbf{u})$ and $\mathbf{u} \in \partial\widehat{\Phi}(\mathbf{v})$ hold. Therefore, using the subdifferential property of \widehat{F} it holds

$$\widehat{F}(\mathbf{w}) - \widehat{F}(\mathbf{u}) \geq \left\langle \partial\widehat{F}(\mathbf{u}), \mathbf{w} - \mathbf{u} \right\rangle_{\mathbb{R}^{Mm}} \quad \forall \mathbf{w} \in \mathbb{R}^{Mm}, \quad (3.6.22)$$

see [54, Eq. (3.9)]. By (3.3. P_D) and the nonnegativity of the dissipation (3.6.21) we obtain

$$\begin{aligned} \widehat{F}(\mathbf{u}(t_{n_2})) - \widehat{F}(\mathbf{u}(t_{n_1})) &= \sum_{n=n_1+1}^{n_2} \widehat{F}(\mathbf{u}(t_n)) - \widehat{F}(\mathbf{u}(t_{n-1})) \\ &\leq \sum_{n=n_1+1}^{n_2} \langle \mathbf{u}(t_n) - \mathbf{u}(t_{n-1}), \mathbf{v}(t_n) \rangle_{\mathbb{R}^{Mm}} \\ &= - \sum_{n=n_1+1}^{n_2} t_\delta^{(n)} \left\langle \widehat{A}\mathbf{v}(t_n), \mathbf{v}(t_n) \right\rangle_{\mathbb{R}^{Mm}} \leq 0, \end{aligned}$$

which means that the free energy decays monotonously along trajectories. By using the elementary inequalities

$$(x/2 - y) \leq (\sqrt{x} - \sqrt{y})^2 \leq x \ln(x/y) - x + y \quad \forall x \geq 0, y > 0, \quad (3.6.23)$$

(The value of $x \ln(x/y)$ at $x = 0$ is understood as $\lim_{x \rightarrow 0} x \ln(x/y) = 0$) we get

$$\begin{aligned} \widehat{F}(\mathbf{U}) \geq \widehat{F}(\mathbf{u}(t_n)) &= \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| \left\{ u_{\nu K}(t_n) \left(\ln \frac{u_{\nu K}(t_n)}{\bar{u}_{\nu K}} - 1 \right) + \bar{u}_{\nu K} \right\} \\ &\geq \sum_{\nu=1}^m \left\{ \frac{1}{2} \|u_{\nu, h}(t_n)\|_{L^1} - \|\bar{u}_{\nu, h}\|_{L^1} \right\}. \end{aligned}$$

□

Remark 3.6.2. Conservation of atoms (3.5.12) implies another possibility to obtain a priori bounds in L^1 without using the free energy of the system. From (3.5.12), we deduce the existence of a vector $\mathbf{s}^\perp \in \mathcal{S}^\perp$, with only positive entries. From the definition (3.4.7) of \widehat{U}^\perp we conclude the existence of $\mathbf{w}^\perp \in \widehat{U}^\perp$ with $w_{\nu K}^\perp = s_\nu^\perp > 0$ for all $K \in \mathcal{V}$ and $\nu = 1, \dots, m$. By (3.4.8) and (3.3. P_D) we find

$$\begin{aligned} 0 &= - \sum_{n=1}^N t_\delta^{(n)} \left\langle \widehat{A}\mathbf{v}, \mathbf{w}^\perp \right\rangle_{\mathbb{R}^{Mm}} = \sum_{n=1}^N \left\langle \mathbf{u}(t_n) - \mathbf{u}(t_{n-1}), \mathbf{w}^\perp \right\rangle_{\mathbb{R}^{Mm}} \\ &= \sum_{\nu=1}^m s_\nu^\perp \sum_{K \in \mathcal{V}} |K| (u_{\nu K}(t_N) - U_{\nu K}) = \sum_{\nu=1}^m s_\nu^\perp (\|u_{\nu, h}(t_N)\|_{L^1} - \|U_{\nu, h}\|_{L^1}). \end{aligned}$$

For the example in Subsection 1.4.1 conservation of atoms hold with $\mathbf{s}^\perp = (1, 1, 2, 1)$. We refer to [111] for more examples of systems which fulfill the property (3.5.12) and for generalizations.

3.7 Thermodynamic equilibrium

As a consequence of Lemma 3.4.1 and Lemma 3.6.1 we want to answer the question if there exists a unique (thermodynamic) equilibrium solution (\mathbf{u}, \mathbf{v}) to the discrete Problem (3.3. P_D) which fulfills the invariance property of Lemma 3.4.1. This is equivalent to solving the problem

$$\hat{A}\mathbf{v} = 0, \quad \mathbf{u} = \hat{E}\mathbf{v}, \quad \mathbf{u} \in \hat{\mathcal{U}} + \mathbf{U}, \quad \mathbf{v} \in \mathbb{R}^{Mm}. \quad (3.7.S_D)$$

Remark 3.7.1. *The proofs of the following results are a special case of [54], if no anisotropies and charged species are taken into account, and [56] if the coefficients are not depending on the state variable. But due Assumption A1, our diffusion coefficients, reaction rate coefficients, and reference densities are $L^\infty(\Omega)$ functions and may depend on the state variable. Therefore for the sake of completeness we modify the proofs here.*

Lemma 3.7.2 (see [54, Lemma 3.3]). *We assume A1. Let $\hat{\Phi}_0 : \mathbb{R}^{Mm} \rightarrow \bar{\mathbb{R}}$ be the functional, defined by*

$$\hat{\Phi}_0(\mathbf{v}) = \hat{\Phi}(\mathbf{v}) + I_{\hat{\mathcal{U}}^\perp}(\mathbf{v}) - \langle \mathbf{U}, \mathbf{v} \rangle_{\mathbb{R}^{Mm}}, \quad \mathbf{v} \in \mathbb{R}^{Mm},$$

where $I_{\hat{\mathcal{U}}^\perp}(\mathbf{v})$ is the characteristic function of the set $\hat{\mathcal{U}}^\perp$, i.e.,

$$I_{\hat{\mathcal{U}}^\perp}(\mathbf{v}) = \begin{cases} 0 & \text{for } \mathbf{v} \in \hat{\mathcal{U}}^\perp, \\ +\infty & \text{for } \mathbf{v} \notin \hat{\mathcal{U}}^\perp. \end{cases}$$

If (\mathbf{u}, \mathbf{v}) is a solution to (3.7. S_D) then \mathbf{v} is the unique minimizer of $\hat{\Phi}_0$. On the other hand if \mathbf{v} is a minimizer of $\hat{\Phi}_0$ then $(\hat{E}\mathbf{v}, \mathbf{v})$ is a solution to (3.7. S_D).

Proof. Following [54], the functional $\hat{\Phi}_0$ is proper, lower semicontinuous and strictly convex and by the Moreau-Rockafeller theorem (see [27])

$$\partial\hat{\Phi}_0(\mathbf{v}) = \hat{E}\mathbf{v} + \partial I_{\hat{\mathcal{U}}^\perp}(\mathbf{v}) - \mathbf{U}, \quad \mathbf{v} \in \mathbb{R}^{Mm}.$$

Let (\mathbf{u}, \mathbf{v}) be a solution to (3.7. S_D) then $\hat{D}(\mathbf{v}) = 0$ hence $\mathbf{v} \in \hat{\mathcal{U}}^\perp$. Therefore $\hat{\Phi}_0(\mathbf{v}) < \infty$ and $\partial\hat{\Phi}_0(\mathbf{v}) = \hat{\mathcal{U}}$. Additionally $\mathbf{u} = \hat{E}\mathbf{v}$ and $\tilde{\mathbf{u}} = \mathbf{u} - \mathbf{U} \in \hat{\mathcal{U}}$ holds and $0 = \mathbf{u} - \tilde{\mathbf{u}} - \mathbf{U} \in \partial\hat{\Phi}_0(\mathbf{v})$ ensures $\hat{\Phi}_0(\mathbf{v}) = \min_{\mathbf{w} \in \mathbb{R}^{Mm}} \hat{\Phi}_0(\mathbf{w})$.

Now assume that \mathbf{v} is a minimizer of $\hat{\Phi}_0$ then $\mathbf{v} \in \hat{\mathcal{U}}^\perp$, $0 \in \partial\hat{\Phi}_0$ and there exists a $\tilde{\mathbf{u}} \in \partial I_{\hat{\mathcal{U}}^\perp}(\mathbf{v}) = \hat{\mathcal{U}}$ such that $\hat{E}\mathbf{v} - \mathbf{U} = \tilde{\mathbf{u}} \in \hat{\mathcal{U}}$. By $\mathbf{v} \in \hat{\mathcal{U}}^\perp$ (see (3.4.7)) we conclude $\hat{D}(\mathbf{v}) = 0$ and $\hat{A}\mathbf{v} = 0$. Hence $(\hat{E}\mathbf{v}, \mathbf{v})$ is a solution to (3.7. S_D). \square

Theorem 3.7.3 (Thermodynamic equilibrium, see [56, Theorem 3.2]).

We assume A1, A3 and A4. Then there exists a unique thermodynamic equilibrium solution $(\mathbf{u}^*, \mathbf{v}^*)$ to the discrete Problem (3.3.P_D). The solution satisfies $\mathbf{v}^* \in \widehat{\mathcal{U}}^\perp$ and $\mathbf{u}^* > \mathbf{0}$.

Proof. In this proof the alphabet c denotes (possibly different) constants and $\|\cdot\|_2$ the 2-norm of \mathbb{R}^{Mm} . According to Lemma 3.7.2 it suffices to show that $\widehat{\Phi}_0(\mathbf{v}) \rightarrow \infty$ if $\|\mathbf{v}\|_2 \rightarrow \infty$. Suppose that this growth condition is violated. Then there exist a $K > 0$ and $\mathbf{v}_n \in \widehat{\mathcal{U}}^\perp$ such that $\|\mathbf{v}_n\|_2 \rightarrow \infty$ and

$$\widehat{\Phi}_0(\mathbf{v}_n) = \widehat{\Phi}(\mathbf{v}_n) - \langle \mathbf{U}, \mathbf{v}_n \rangle_{\mathbb{R}^{Mm}} < K.$$

By the definition of $\widehat{\Phi}$ (see (3.6.17)) we obtain

$$c \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| |(v_{n\nu K})^+|^2 - \langle \mathbf{U}, \mathbf{v} \rangle_{\mathbb{R}^{Mm}} \leq K + c. \quad (3.7.24)$$

Let $\mathbf{w}_n := \mathbf{v}_n / \|\mathbf{v}_n\|_2$, then $\mathbf{w}_n \rightarrow \tilde{\mathbf{w}} \in \mathbb{R}^{Mm}$ at least for a subsequence and

$$c \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| |(w_{n\nu K})^+|^2 \leq \frac{K + c}{\|\mathbf{v}_n\|_2^2} + \frac{\|\mathbf{U}\|_2}{\|\mathbf{v}_n\|_2}.$$

This leads to $(w_{n\nu K})^+ \rightarrow 0$ for $n \rightarrow \infty$ and $w_{n\nu K} = (w_{n\nu K})^+ - (w_{n\nu K})^- \rightarrow \tilde{w}_{n\nu K}$ gives $-\tilde{w}_{n\nu K} \geq 0$, $K \in \mathcal{V}$, $\nu = 1, \dots, m$. Since $\mathbf{w}_n \in \widehat{\mathcal{U}}^\perp$ and $\widehat{\mathcal{U}}^\perp$ is closed it follows that $\tilde{\mathbf{w}} \in \widehat{\mathcal{U}}^\perp$. Therefore $\tilde{w}_{\nu K} = s_\nu^\perp$ with $\mathbf{s}^\perp = (s_\nu^\perp)_{\nu=1}^m \in \mathcal{S}^\perp$ for all $K \in \mathcal{V}$ and $\nu = 1, \dots, m$. From $\|\mathbf{w}_n\| = 1$ we conclude $\mathbf{s}^\perp \neq \mathbf{0}$ and by exploiting (3.7.24) again we obtain

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} \frac{K + c}{\|\mathbf{v}_n\|_2^2} \geq - \lim_{n \rightarrow \infty} \langle \mathbf{U}, \mathbf{w}_n \rangle_{\mathbb{R}^{Mm}} = - \langle \mathbf{U}, \tilde{\mathbf{w}} \rangle_{\mathbb{R}^{Mm}} \\ &= - \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| U_{\nu K} s_\nu^\perp = - \sum_{\nu=1}^m s_\nu^\perp \|U_\nu\|_{L^1(\Omega)} \end{aligned}$$

which gives a contradiction to Assumption A4 together with $\mathbf{U} > \mathbf{0}$. \square

The connection between the discrete and continuous thermodynamic equilibrium is given by the following corollary.

Corollary 3.7.4 (see [56, Corollary 3.3]). We assume A1, A3 and A4. Let $(\mathbf{u}^*, \mathbf{v}^*)$ be the unique thermodynamic equilibrium solution to (3.7.S_D). The corresponding piecewise constant functions $(u_{\nu,h}^*, v_{\nu,h}^*)$ are related to the continuous thermodynamic equilibrium solution (u^*, v^*) of Problem (2.3.S) by

$$u_{\nu,h}^* = \frac{\bar{u}_{\nu,h}}{\bar{u}_\nu} u_\nu^*, \quad v_{\nu,h}^* = v_\nu^*, \quad a_{\nu,h}^* = a_\nu^*, \quad \nu = 1, \dots, m.$$

Proof. In the case of diffusion coefficients only depending on the space variable and not on the state variables the result was proven in [56, Corollary 3.3].

If $(\mathbf{u}^*, \mathbf{v}^*)$ solution to (3.7. S_D) then $\widehat{D}(\mathbf{v}^*) = 0$ and therefore $v_{\nu,K}^* = w_\nu$ with $(w_\nu)_{\nu=1}^m \in \mathcal{S}^\perp$ for all $K \in \mathcal{V}$. Hence $a_{\nu,h}(x) = e^{w_\nu} = e^{v_{\nu,K}^*} = a_{\nu K}$ for all $x \in K \in \mathcal{V}$. Moreover

$$a_h^* \in \mathbb{R}_+^m, (a_h^*)^\alpha = (a_h^*)^\beta \quad \forall (\alpha, \beta) \in \mathcal{R}.$$

Using $\mathbf{u}^* = \widehat{E}\mathbf{v}^* \in \widehat{\mathcal{U}} + \mathbf{U}$ we find $u_h^* - U \in \mathcal{U}$ and for all $\mathbf{s}^\perp \in \mathcal{S}^\perp$ the identity

$$0 = \sum_{\nu=1}^m s_\nu^\perp \int_{\Omega} u_{\nu,h}^* - U_\nu dx = \sum_{\nu=1}^m s_\nu^\perp \int_{\Omega} a_{\nu,h}^* \bar{u}_{\nu,h} - U_\nu dx = \sum_{\nu=1}^m s_\nu^\perp \int_{\Omega} a_{\nu,h}^* \bar{u}_\nu - U_\nu dx$$

holds. Hence $a_h^* \in \mathcal{A}$ (see (2.3.5)) and by Assumption A3 $\mathcal{A} = \{a^*\}$ which ensures that $a_h^* = a^*$, $v_h^* = v^*$, and $u_{\nu,h}^* = \frac{\bar{u}_{\nu,h}}{\bar{u}_\nu} u_\nu^*$ for all $\nu = 1, \dots, m$. \square

Remark 3.7.5. *It is essential to use a dissipative discretization scheme. A non-dissipative scheme may lead to an increase of the free energy. As a consequence, the computed equilibrium solution of the discrete problem could be far apart from the exact minimum of the free energy. A violation of the invariance property of Lemma 3.4.1 behaves in the same way, see Subsection 5.2.*

3.8 Global upper bounds

In this section we want to prove upper bounds for the densities that are uniform in time and space. We start with a simple example:

Example 3.8.1. *Conservation of atoms (3.5.12), implies the existence of a vector $\mathbf{s}^\perp \in \mathcal{S}^\perp$, and from (3.3. P_D) we deduce*

$$\sum_{\nu=1}^m s_\nu^\perp \frac{u_\nu^{(K)}(t_n) - u_\nu^{(K)}(t_{n-1})}{t_\delta^{(n)}} - \sum_{\nu=1}^m s_\nu^\perp \sum_{\sigma=K|L \in \mathcal{E}_K} T_\sigma Y_\nu^\sigma (a_{\nu L} - a_{\nu K}) = 0 \quad \forall K \in \mathcal{V}.$$

Assuming a homogeneous material and constant diffusion coefficients $D_\nu = D$ for $\nu = 1, \dots, m$ implies $Y_\nu^\sigma (a_{\nu L} - a_{\nu K}) = D(u_{\nu L} - u_{\nu K})$ and by introducing $\mathbf{w}(t_n) = \sum_{\nu=1}^m s_\nu^\perp \mathbf{u}_\nu(t_n)$ we obtain

$$\frac{w^{(K)}(t_n) - w^{(K)}(t_{n-1})}{t_\delta^{(n)}} - D \sum_{\sigma=K|L \in \mathcal{E}_K} T_\sigma (w_L - w_K) = 0 \quad \forall K \in \mathcal{V}.$$

This is a discretized homogeneous heat equation and by using discrete maximum principle [32] we obtain the L^∞ estimate

$$\|w_h(t_n)\|_{L^\infty} \leq c \|W_h\|_{L^\infty} \quad \forall n > 0$$

with a constant $c > 0$ and initial value $\mathbf{W} = \sum_{\nu=1}^m s_{\nu}^{\perp} U_{\nu}$. Here $w_h \in X_{\nu}(\mathcal{M})$ denotes the corresponding piecewise constant function of w . Such strong assumptions on the diffusion coefficient exclude practical relevant examples.

The maximum principle is only applicable to systems having a special structure, e.g., systems of triangle form. Therefore we use a technique introduced by Moser [104] to show global boundedness of discrete solutions to (3.3.P_D). For the continuous problem this was done in [45, 64].

Remark 3.8.2. *Conservation of qualitative properties like local and global mass conservation, maximum principles, positivity or more generally, L^{∞} bounds is known to be difficult for finite element discretizations, see e.g. [21, 97]. One important reason for that is that finite element methods do usually not allow to test the discrete equations with the positive part of a finite element function, since it does not lie in the discrete test space, in general. On the contrary, finite volume methods allow such test functions [34, 48], and we will show how the use of appropriate discrete test functions will deliver us the desired conservation of qualitative properties.*

In order to establish the new results in the following sections we need the additional assumption:

Assumption A5. *There exists a constant $\bar{t}_{\delta} < \infty$ such that the largest possible time step is bounded by $\max_{n \in \mathbb{N}} t_{\delta}^{(n)} \leq \bar{t}_{\delta}$ for all considered time discretizations.*

For obtaining the global bounds, we use a bootstrapping technique introduced by Moser, see [104], and apply it to the discretized problem.

Theorem 3.8.3 (Upper bounds).

We assume A1 and A2. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n \in \mathbb{N}})$ be a class of discretizations fulfilling the Assumptions A5. Then there exists a constant $c_1 > 0$ only depending on the data and not on \mathcal{D} such that for any solution (u_h, v_h) to (3.3.P_D)

$$\sum_{\nu=1}^m \|u_{\nu,h}(t_N)/\bar{u}_{\nu,h}\|_{L^2} \leq c_1 \quad \forall N \geq 1$$

holds uniformly for all discretizations \mathcal{D} . Furthermore there exists a second constant $c_2 > 0$ only depending on the data and not on \mathcal{D} such that

$$\|u_{\nu,h}(t_N)/\bar{u}_{\nu,h}\|_{L^{\infty}} \leq c_2 \quad \forall N \geq 1, \quad \nu = 1, \dots, m$$

holds uniformly for all discretizations \mathcal{D} .

Remark 3.8.4. *In the continuous case, the continuous problem (2.2.P) is tested with functions*

$$p e^t (z_1^{p-1}, \dots, z_m^{p-1}), \quad z_{\nu} = (a_{\nu} - \kappa)^+, \quad \kappa = \|U_{\nu}/\bar{u}_{\nu}\|_{L^{\infty}}$$

for $p = 2^k$, $k \geq 1$ in order to obtain global upper bounds of the solutions, see [63, Lemma 4.1 and Theorem 4.2]. The discrete proof follows the continuous counterpart. However, since we have a time discretization, the test functions have to be modified and we have to estimate some error terms coming from the discretization of the diffusion term.

Proof. First we mention that

$$e^{t_{n-1}} \leq \frac{e^{t_n} - e^{t_{n-1}}}{t_\delta^{(n)}} \leq e^{t_n} = e^{t_{n-1} + t_\delta^{(n)}} \leq e^{\bar{t}_\delta} e^{t_{n-1}} \quad (3.8.25)$$

with \bar{t}_δ given in Definition 3.3.1. We introduce $z_{\nu,h} = (e^{v_{\nu,h}} - \kappa)^+$ with

$$\kappa := \max_{\nu=1,\dots,m} \frac{\text{ess sup}_{x \in \Omega} U_\nu(x)}{\text{ess inf}_{x \in \Omega} \bar{u}_\nu(x)}$$

and $w_{\nu,h} = z_{\nu,h}^{p/2}$, $p \geq 2$. The constant κ is chosen in such a way that $z_{\nu,h}(t_0) = (U_{\nu,h}/\bar{u}_{\nu,h} - \kappa)^+ = 0$. Now, we test (3.3.PD) with test functions $pe^{t_{n-1}} z_{\nu,h}^{p-1}(t_n)$, $p \geq 2$, and obtain

$$\begin{aligned} S_1 &:= \sum_{n=1}^N t_\delta^{(n)} p e^{t_{n-1}} \left\langle \frac{\mathbf{u}(t_n) - \mathbf{u}(t_{n-1})}{t_\delta^{(n)}}, \mathbf{z}^{p-1}(t_n) \right\rangle_{\mathbb{R}^{Mm}} \\ &= - \sum_{n=1}^N t_\delta^{(n)} p e^{t_{n-1}} \left\langle \widehat{A}\mathbf{v}(t_n), \mathbf{z}^{p-1}(t_n) \right\rangle_{\mathbb{R}^{Mm}} = S_2 + S_3 \end{aligned}$$

with

$$\begin{aligned} S_2 &:= - \sum_{\nu=1}^m \sum_{n=1}^N p t_\delta^{(n)} e^{t_{n-1}} \sum_{\sigma=K|L \in \mathcal{E}_{int}} T_\sigma Y_\nu^\sigma Z_\nu^\sigma (v_{\nu L} - v_{\nu K}) (z_{\nu L}^{p-1} - z_{\nu K}^{p-1}), \\ S_3 &:= \sum_{n=1}^N t_\delta^{(n)} p e^{t_{n-1}} \sum_{K \in \mathcal{V}} |K| \sum_{\nu=1}^m z_{\nu K}^{p-1}(t_n) R_{\nu K}(e^{\mathbf{v}_K(t_n)}). \end{aligned}$$

Next, the expressions for the time-derivative (S_1), diffusion term (S_2) and reaction term (S_3) are estimated separately.

Time derivative: Straightforward calculations, using that the product of the positive and negative part of a function vanishes, yield

$$\begin{aligned} S_1 &= p \sum_{\nu=1}^m \sum_{n=1}^N e^{t_{n-1}} \sum_{K \in \mathcal{V}} |K| z_{\nu K}^{p-1}(t_n) (u_{\nu K}(t_n) - u_{\nu K}(t_{n-1})) \\ &= p \sum_{\nu=1}^m \sum_{n=1}^N e^{t_{n-1}} \sum_{K \in \mathcal{V}} |K| \bar{u}_{\nu K} z_{\nu K}^{p-1}(t_n) \left((z_{\nu K}(t_n) - z_{\nu K}(t_{n-1})) + (e^{v_{\nu K}(t_{n-1})} - \kappa)^- \right). \end{aligned}$$

Using (C.27) and the fact that $z_{\nu K}^{p-1}(t_n)(e^{v_{\nu K}(t_{n-1})} - \kappa)^- \geq 0$ holds, we get

$$\begin{aligned} S_1 &\geq \sum_{\nu=1}^m \sum_{n=1}^N e^{t_{n-1}} \sum_{K \in \mathcal{V}} |K| \bar{u}_{\nu K} (z_{\nu K}^p(t_n) - z_{\nu K}^p(t_{n-1})) \\ &= \sum_{\nu=1}^m \sum_{n=1}^N \sum_{K \in \mathcal{V}} |K| \bar{u}_{\nu K} \left\{ (e^{t_n} z_{\nu K}^p(t_n) - e^{t_{n-1}} z_{\nu K}^p(t_{n-1})) - (e^{t_n} - e^{t_{n-1}}) z_{\nu K}^p(t_n) \right\} \\ &\geq \sum_{\nu=1}^m \left\{ e^{t_N} \underline{c}_{\bar{u}} \|z_{\nu, h}(t_N)\|_{L^p}^p - \sum_{n=1}^N t_{\delta}^{(n)} e^{\bar{t}_{\delta}} e^{t_{n-1}} \|\bar{u}_{\nu, h}\|_{L^\infty} \|z_{\nu, h}\|_{L^p}^p \right\}. \end{aligned} \tag{3.8.26}$$

In the last line we used $z_{\nu, h}(t_0) = 0$ and (3.8.25).

Diffusion term: Now, we consider the diffusion term S_2 . Applying the definition (3.2.2) to

$$\begin{aligned} Z_{\nu}^{\sigma}(v_{\nu L} - v_{\nu K}) z_{\nu K}^{p-1} &= (z_{\nu L} - z_{\nu K}) z_{\nu K}^{p-1} - ((e^{v_{\nu L}} - \kappa)^- - (e^{v_{\nu K}} - \kappa)^-) z_{\nu K}^{p-1} \\ &= (z_{\nu L} - z_{\nu K}) z_{\nu K}^{p-1} - (e^{v_{\nu L}} - \kappa)^- z_{\nu K}^{p-1} \\ &\leq (z_{\nu L} - z_{\nu K}) z_{\nu K}^{p-1}, \end{aligned}$$

using the lower bound of the reference densities and the diffusion coefficients, inequality (C.26), the notation $w_{\nu,h} = z_{\nu,h}^{p/2}$, and the extension of the H^1 -seminorm to the full H^1 -norm, we find for $p \geq 2$

$$\begin{aligned}
S_2 &= \sum_{\nu=1}^m \sum_{n=1}^N p t_\delta^{(n)} e^{t_{n-1}} \sum_{K \in \mathcal{V}} \sum_{\sigma=K|L \in \mathcal{E}_K} T_\sigma Y_\nu^\sigma Z_\nu^\sigma (v_{\nu L} - v_{\nu K}) z_{\nu K}^{p-1} \\
&\leq - \sum_{\nu=1}^m \sum_{n=1}^N p t_\delta^{(n)} e^{t_{n-1}} \sum_{\sigma=K|L \in \mathcal{E}_{int}} T_\sigma Y_\nu^\sigma (z_{\nu L} - z_{\nu K}) (z_{\nu L}^{p-1} - z_{\nu K}^{p-1}) \\
&\leq \sum_{\nu=1}^m \sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \frac{4(p-1)c_{D\bar{u}}}{p} \left\{ -\|w_{\nu,h}\|_{H^1, \mathcal{M}}^2 + \|z_{\nu,h}\|_{L^p}^p \right\} \\
&\leq \sum_{\nu=1}^m \sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \left\{ -2c_{D\bar{u}} \|w_{\nu,h}\|_{H^1, \mathcal{M}}^2 + p c_{D\bar{u}} \|z_{\nu,h}\|_{L^p}^p \right\}.
\end{aligned}$$

In the last two lines we set $c_{D\bar{u}} = c_D c_{\bar{u}}$ and use $2 \leq 4(p-1)/p \leq p$ for $p \in [2, \infty)$.

Reaction terms: Together with (2.1.1), the inequalities

$$e^{2v_{\nu K}} \leq (z_{\nu K} + \kappa)^2 \leq 2(z_{\nu K}^2 + \kappa^2),$$

Muirhead's inequality (see [105])

$$\sum_{\nu,j=1}^m z_{jK}^2 z_{\nu K}^{p-1} \leq m \sum_{\nu=1}^m z_{\nu K}^{p+1},$$

and $x^p \leq x^{p+1} + 1$ for $x \geq 0$ and $p \geq 1$ we can estimate the reaction terms (1.4.6) with at most quadratic source terms (see Assumption A1 and (2.1.1)) by

$$\begin{aligned}
S_3 &\leq C_1 \sum_{\nu=1}^m \sum_{n=1}^N p t_\delta^{(n)} e^{t_{n-1}} \sum_{K \in \mathcal{V}} |K| z_{\nu K}^{p-1} \left(1 + \sum_{j=1}^m e^{2v_{jK}} \right) \\
&\leq 2C_1 \sum_{\nu=1}^m \sum_{n=1}^N p t_\delta^{(n)} e^{t_{n-1}} \left\{ (1 + m\kappa^2) \|z_{\nu,h}\|_{L^{p-1}}^{p-1} + \sum_{j=1}^m \sum_{K \in \mathcal{V}} |K| z_{jK}^2 z_{\nu K}^{p-1} \right\} \\
&\leq C_2 \sum_{n=1}^N t_\delta^{(n)} p e^{t_{n-1}} \sum_{\nu=1}^m \left(\|z_{\nu,h}\|_{L^{p+1}}^{p+1} + 1 \right)
\end{aligned}$$

with constants $C_1, C_2 > 0$.

The tested equation: Using the obtained estimates of the three terms together with $S_1 = S_2 + S_3$, leads to

$$\begin{aligned} S_4 &:= \sum_{\nu=1}^m e^{t_N} \underline{c}_{\bar{u}} \|z_{\nu,h}(t_N)\|_{L^p}^p \\ &\leq \sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \sum_{\nu=1}^m \left\{ -2\underline{c}_{D\bar{u}} \|w_{\nu,h}\|_{H^1, \mathcal{M}}^2 + pC_3 (\|z_{\nu,h}\|_{L^{p+1}}^{p+1} + 1) \right\} \end{aligned} \quad (3.8.27)$$

with a constant $C_3 > 0$ which can be chosen such that it depends on the largest possible time step (see Definition 3.3.1 and (3.8.26)) and the data, but not on p .

Bounds in L^2 : For obtaining the L^2 bound, we set $p = 2$. The last term in (3.8.27) (the L^3 - norm of $z_{\nu,h}$) can be controlled by the discrete Gagliardo-Nirenberg inequality (B.24), i.e., we find for all $\varepsilon > 0$ a $c_{\varepsilon,3} > 0$ such that

$$\|z_{\nu,h}\|_{L^3}^3 \leq 2\varepsilon \|z_{\nu,h} \ln z_{\nu,h}\|_{L^1} \|z_{\nu,h}\|_{H^1, \mathcal{M}}^2 + c_{\varepsilon,3} \|z_{\nu,h}\|_{L^1}. \quad (3.8.28)$$

We continue (3.8.27) by

$$S_4 \leq \sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \sum_{\nu=1}^m \left\{ g_1(\varepsilon) \|w_{\nu,h}\|_{H^1, \mathcal{M}}^2 + 2c_{\varepsilon,3} C_3 \|z_{\nu,h}\|_{L^1} \right\},$$

where $g_1(\varepsilon) := -2\underline{c}_{D\bar{u}} + 4\varepsilon C_3 \|z_{\nu,h} \ln z_{\nu,h}\|_{L^1}$. To choose the constant ε we have to control the L^1 - norm of $z_{\nu,h} \ln z_{\nu,h}$ and $z_{\nu,h}$. From Lemma 3.6.1 and $z_{\nu,h} \leq a_{\nu,h}$ we deduce the boundedness of $\|z_{\nu,h}\|_{L^1}$. Since $|(x - \mu) \ln(x - \mu)| \leq x \ln x + 1$ holds for $x \geq \mu \geq 0$ and by Lemma 3.6.1 we obtain

$$\begin{aligned} \|z_{\nu,h} \ln z_{\nu,h}\|_{L^1} &= \sum_{\substack{K \in \mathcal{V}_i \\ u_{\nu K} > \kappa \bar{u}_{\nu K}}} |K| |(a_{\nu K} - \kappa) \ln(a_{\nu K} - \kappa)| \\ &\leq \frac{1}{\underline{c}_{\bar{u}}} \sum_{\substack{K \in \mathcal{V}_i \\ u_{\nu K} > \kappa \bar{u}_{\nu K}}} |K| \left(u_{\nu K} \ln \frac{u_{\nu K}}{\bar{u}_{\nu K}} + \bar{u}_{\nu K} - u_{\nu K} + u_{\nu K} \right) \\ &\leq \frac{1}{\underline{c}_{\bar{u}}} \left(\widehat{F}(U) + \|u_{\nu,h}\|_{L^1} \right), \quad \nu = 1, \dots, m. \end{aligned}$$

We fix the constant $\varepsilon > 0$, coming from the Gagliardo-Nirenberg inequality such that $g_1(\varepsilon) \leq 0$ holds. From (3.12.78) we get

$$\sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \leq \sum_{n=1}^N e^{t_n} - e^{t_{n-1}} = e^{t_N} - 1$$

and from (3.8.27) for $p = 2$ we can derive the boundedness of

$$\underline{c}_{\bar{u}} \sum_{\nu=1}^m \|z_{\nu,h}(t_N)\|_{L^2}^2 \leq C_4, \quad N \geq 1 \quad (3.8.29)$$

with a constant $C_4 > 0$. The first result of the theorem follows by using the inequality $u_{\nu,K}/\bar{u}_{\nu,K} \leq z_{\nu,K} + \kappa$.

Moser iteration for $p \geq 4$: For $p \geq 4$ let $r = \frac{2(p+1)}{p}$ be introduced. Note that $r \in (2, 5/2]$ for $p \in [4, \infty)$. Using $w_{\nu,h} = z_{\nu,h}^{p/2}$, the estimate (3.8.27) can be written as

$$\begin{aligned} S_4 &= \sum_{\nu=1}^m e^{t_N} \underline{c}_{\bar{u}} \|w_{\nu,h}(t_N)\|_{L^2}^2 \\ &\leq \sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \sum_{\nu=1}^m \left\{ -2\underline{c}_{D\bar{u}} \|w_{\nu,h}\|_{H^1, \mathcal{M}}^2 + pC_3 (\|w_{\nu,h}\|_{L^r}^r + 1) \right\}. \end{aligned} \quad (3.8.30)$$

By the discrete Gagliardo-Nirenberg inequality (B.19) we obtain

$$\|w_{\nu,h}\|_{L^r}^r \leq c_{gn,r}^r \|w_{\nu,h}\|_{L^1} \|w_{\nu,h}\|_{H^1, \mathcal{M}}^{r-1}, \quad c_{gn,r}^r = 2^{(r-1)/2} C^r.$$

For $r \in (2, 5/2]$ the constants appearing in the Gagliardo-Nirenberg inequality (B.19) can be uniformly bounded by

$$c_{gn,r}^r \leq \bar{c}_r := \max \left\{ 1, \max_{s \in [2, 5/2]} \{c_{gn,s}\} \right\}^{5/2}.$$

Using Young's inequality with $q = \frac{2p}{p+2}$, $q' = \frac{2p}{p-2}$, and $\varepsilon > 0$ we get

$$\|w_{\nu,h}\|_{L^1} \|w_{\nu,h}\|_{H^1, \mathcal{M}}^{r-1} \leq \frac{\varepsilon}{q} \|w_{\nu,h}\|_{H^1, \mathcal{M}}^2 + \frac{\varepsilon^{-q'/q}}{q'} \|w_{\nu,h}\|_{L^1}^{q'}.$$

Inserting this in (3.8.30) we find

$$S_4 \leq \sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \sum_{\nu=1}^m \left\{ g_2(\varepsilon) \|w_{\nu,h}\|_{H^1, \mathcal{M}}^2 + pC_3 \left(\frac{\varepsilon^{-q'/q}}{\bar{c}_r} \|w_{\nu,h}\|_{L^1}^{q'} + 1 \right) \right\}$$

with $g_2(\varepsilon) = -2\underline{c}_{D\bar{u}} + C_3 \bar{c}_r \varepsilon^{\frac{p}{q}}$. The constant ε is fixed such that $g_2(\varepsilon) = 0$, i.e.,

$$\varepsilon = \frac{2\underline{c}_{D\bar{u}} q}{C_3 \bar{c}_r p}.$$

Then the term in front of the L^1 norm can be decomposed into two factors

$$\frac{\bar{c}_r \varepsilon^{-q'/q}}{q'} \leq \left(\bar{c}_r \left(\frac{C_3 \bar{c}_r}{2 \underline{c}_{D\bar{u}}} \right)^{\frac{p+2}{p-2}} \right) \left(\frac{p-2}{2p} \left(\frac{p+2}{2} \right)^{\frac{p+2}{p-2}} \right).$$

The first factor is bounded and for the second factor we find for $p \geq 4$ by monotonicity

$$\frac{p-2}{2p} \left(\frac{p+2}{2} \right)^{\frac{p+2}{p-2}} = p \frac{1-4/p^2}{4} \left(\frac{p+2}{2} \right)^{\frac{4}{p-2}} \leq \frac{27}{16} p.$$

Therefore we define

$$C_4 := C_3 \max \left(1, \frac{27 \bar{c}_r}{16} \left(\max \left(1, \frac{C_3 \bar{c}_r}{2 \underline{c}_{D\bar{u}}} \right) \right)^3 \right)$$

and we proceed with

$$\begin{aligned} e^{-t_N} S_4 &= \sum_{\nu=1}^m \underline{c}_{\bar{u}} \|z_{\nu,h}(t_N)\|_{L^p}^p \\ &\leq \sum_{n=1}^N e^{-t_N} t_\delta^{(n)} e^{t_{n-1}} \sum_{\nu=1}^m p^2 C_4 \left(\|w_{\nu,h}\|_{L^1}^{q'} + 1 \right) \\ &\leq p^2 C_4 \sum_{\nu=1}^m \sup_{n=1,\dots,N} \left(\|z_{\nu,h}(t_n)\|_{L^{p/2}}^{p^2/(p-2)} + 1 \right). \end{aligned}$$

Therefore, with some constant $C_5 > 1$ we get

$$e^{-t_N} S_4 + 1 \leq p^2 C_5 \left\{ \sum_{\nu=1}^m \sup_{n=1,\dots,N} \left(\|z_{\nu,h}(t_n)\|_{L^{p/2}}^{p/2} + 1 \right) \right\}^{2p/(p-2)}.$$

Iteratively using this inequality and setting $p = 2^k$, $k \geq 2$ and

$$b_k := \sum_{\nu=1}^m \sup_{n=1,\dots,N} \|z_{\nu,h}(t_n)\|_{L^{2^k}}^{2^k} + 1, \quad k \geq 1,$$

we find for $k \in \mathbb{N}_+$, $k \geq 2$ the recursion formula $b_k \leq (4)^k C_5 (b_{k-1})^2 \frac{2^{k-1}}{2^{k-1}-1}$. This can be written as

$$b_k \leq \left((4)^\zeta (c_5)^\eta (b_1)^{2^{k-1}} \right) \prod_{j=1}^{k-1} a_j$$

with $q_j = \frac{2^j}{2^{j-1}}$ and

$$\zeta = \left(k + \sum_{i=1}^{k-2} (k-i) 2^i \prod_{l=k-i}^{k-1} q_l \right) \prod_{j=1}^{k-1} q_j^{-1} = k \prod_{j=1}^{k-1} q_j^{-1} + \sum_{i=1}^{k-2} (k-i) 2^i \prod_{j=1}^{k-i-1} q_l^{-1} \leq \sum_{i=0}^{k-2} (k-i) 2^i,$$

$$\eta = \left(1 + \sum_{i=1}^{k-2} 2^i \prod_{l=k-i}^{k-1} q_l \right) \prod_{j=1}^{k-1} q_j^{-1} = \prod_{j=1}^{k-1} q_j^{-1} + \sum_{i=1}^{k-2} 2^i \prod_{j=1}^{k-i-1} q_l^{-1} \leq \sum_{i=0}^{k-2} 2^i$$

because of $q_l^{-1} \leq 1$. By induction follows

$$\sum_{i=0}^{k-2} 2^i \leq 2^{k-1} \leq 2^k, \quad \sum_{i=0}^{k-2} (k-i) 2^i \leq 2^{k+1}, \quad k \geq 2, \quad (3.8.31)$$

see [63, p. 217]. The product $\theta = \prod_{j=1}^{\infty} q_j$ is finite. Therefore we obtain

$$b_k \leq \left[(4)^{2^{k+1}} (C_5)^{2^{k-1}} b_1^{2^{k-1}} \right]^{\prod_{j=1}^{k-1} q_j} \leq (16C_5 b_1)^{\theta 2^k}. \quad (3.8.32)$$

Since b_1 is bounded from above by (3.8.29) we obtain for $k \geq 2$

$$\sum_{\nu=1}^m \|z_{\nu,h}(t_N)\|_{L^2} \leq \sqrt{m} \left\{ 16C_5 \left(\sum_{\nu=1}^m \sup_{n=1,\dots,N} \|z_{\nu,h}(t_n)\|_{L^2}^2 + 1 \right) \right\}^{\theta}$$

and finally with [84, Theorem 2.11.5]

$$\sum_{\nu=1}^m \|z_{\nu,h}(t_N)\|_{L^\infty} \leq \sqrt{m} \left\{ 16C_5 \left(\sum_{\nu=1}^m \sup_{n=1,\dots,N} \|z_{\nu,h}(t_n)\|_{L^2}^2 + 1 \right) \right\}^{\theta}$$

for $k \rightarrow \infty$. From $u_{\nu,h}/\bar{u}_{\nu,h} \leq z_{\nu,h} + \kappa$ the result follows. \square

3.9 Asymptotics

In this section we will extend the result of Lemma 3.6.1. We mention the result of [56] where it is proved that the free energy decays exponentially along trajectories. We also note that in special situations an explicit rate of convergence is proven, see [22].

Definition 3.9.1. Let $(\mathbf{u}^*, \mathbf{v}^*)$ be the unique thermodynamic equilibrium of the discrete Problem (3.3.P_D). For every (\mathbf{u}, \mathbf{v}) with $\mathbf{u} = \hat{E}\mathbf{v}$, the relative free energy of (3.3.P_D) is introduced by

$$\hat{\Psi}(\mathbf{u}) := \hat{F}(\mathbf{u}) - \hat{F}(\mathbf{u}^*).$$

Lemma 3.9.2 (see [56, Lemma 3.4]). *Let the Assumptions of A1, A3, and A4 be fulfilled and let $\mathcal{D} = (\mathcal{M}, (t_n)_{n \in \mathbb{N}})$ be a class of discretizations fulfilling the Assumptions A5. Let $\mathbf{u} = \widehat{E}\mathbf{v} \in \widehat{\mathcal{U}} + \mathbf{U}$ and let $(\mathbf{u}^*, \mathbf{v}^*)$ be the discrete thermodynamic equilibrium according to Theorem 3.7.3. Moreover, let (u_h, v_h) and (u_h^*, v_h^*) the corresponding piecewise constant functions. Then, there exist constants $c_1, c_2 > 0$ not depending on \mathcal{M} such that*

$$c_1 \|\sqrt{u_h} - \sqrt{u_h^*}\|_Y^2 \leq \widehat{\Psi}(\mathbf{u}) \leq c_2 \|u_h - u_h^*\|_Y^2.$$

Proof. From $\langle \mathbf{u} - \mathbf{u}^*, \mathbf{v}^* \rangle_{\mathbb{R}^{Mm}} = 0$ and (3.6.18) we obtain

$$\begin{aligned} \widehat{\Psi}(\mathbf{u}) &= \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^{Mm}} - \widehat{\Phi}(\mathbf{v}) - \langle \mathbf{u}^*, \mathbf{v}^* \rangle_{\mathbb{R}^{Mm}} + \widehat{\Phi}(\mathbf{v}^*) \\ &= \langle \mathbf{u}, \mathbf{v} - \mathbf{v}^* \rangle_{\mathbb{R}^{Mm}} - \widehat{\Phi}(\mathbf{v}) + \widehat{\Phi}(\mathbf{v}^*) \\ &= \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| \bar{u}_{\nu K} \left(e^{v_{\nu K}} (v_{\nu K} - v_{\nu K}^*) - e^{v_{\nu K}} + e^{v_{\nu K}^*} \right) \\ &= \sum_{\nu=1}^m \int_{\Omega} \bar{u}_{\nu, h} \left(e^{v_{\nu, h}} (v_{\nu, h} - v_{\nu, h}^*) - e^{v_{\nu, h}} + e^{v_{\nu, h}^*} \right) dx \\ &= \sum_{\nu=1}^m \int_{\Omega} \left(u_{\nu, h} \ln \frac{u_{\nu, h}}{u_{\nu, h}^*} - u_{\nu, h} + u_{\nu, h}^* \right) dx. \end{aligned}$$

Using the elementary estimates

$$(\sqrt{x} - \sqrt{y})^2 \leq x \ln \frac{x}{y} - x + y \leq \frac{1}{y} (x - y)^2 \quad \forall x, y > 0 \quad (3.9.33)$$

and Corollary 3.7.4 the result of the lemma follows. \square

Theorem 3.9.3 (Estimate of the relative free energy, see [56, Theorem 3.5]).

We assume A1, A3, and A4. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n \in \mathbb{N}})$ be a class of discretizations fulfilling the Assumptions A5. Let $(\mathbf{u}^, \mathbf{v}^*)$ be the unique thermodynamic equilibrium of the discrete Problem. Then for every $\rho > 0$ there exists a constant $c_\rho > 0$ such that for all Voronoi meshes and all*

$$\mathbf{v} \in \widehat{\mathcal{N}}_\rho := \left\{ \mathbf{v} \in \mathbb{R}^{Mm} : \widehat{\Psi}(\widehat{E}\mathbf{v}) \leq \rho, \mathbf{u} = \widehat{E}\mathbf{v} \in \widehat{\mathcal{U}} + \mathbf{U} \right\}$$

the inequality

$$\widehat{\Psi}(\mathbf{u}) \leq c_\rho \widehat{D}(\mathbf{v}) \quad (3.9.34)$$

is fulfilled.

Remark 3.9.4. *The proof is given in [56] in case of strictly positive reaction rate coefficients and for all meshes with size less than a constant determined by the data of the problem. Instead of using the discrete*

Sobolev-Poincare inequality [59, Theorem 2.2] we use Lemma B.4 in order to relax an assumption on the mesh needed by the discrete Sobolev-Poincare inequality [59, Theorem 2.2]. In contrast to [56], in our result nonnegative reaction rate coefficients are allowed (see Assumption A1) and the estimate holds true for meshes with arbitrary mesh size.

Proof. In this proof the alphabet c denotes (possibly different) constants.

1. Let $\rho > 0$ be arbitrary given. Using the elementary inequality

$$(x - y) \ln \frac{x}{y} \geq (\sqrt{x} - \sqrt{y})^2 \quad \forall x, y > 0 \quad (3.9.35)$$

we obtain for $\mathbf{v} = \ln \mathbf{a}$, $\mathbf{a} \in \mathbb{R}_+^m$ and for all $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{R}$ the estimate

$$\left(e^{\mathbf{v} \cdot \boldsymbol{\alpha}} - e^{\mathbf{v} \cdot \boldsymbol{\beta}} \right) (\boldsymbol{\alpha} - \boldsymbol{\beta}) \cdot \mathbf{v} \geq \left(\sqrt{\mathbf{a}^{\boldsymbol{\alpha}}} - \sqrt{\mathbf{a}^{\boldsymbol{\beta}}} \right)^2 =: r_{(\boldsymbol{\alpha}, \boldsymbol{\beta})}(\mathbf{a}). \quad (3.9.36)$$

Therefore with Assumption A1 the discrete dissipation rate can be estimated for $\mathbf{v} \in \mathbb{R}^{Mm}$ by

$$\begin{aligned} \widehat{D}(\mathbf{v}) &\geq c \sum_{\nu=1}^m \sum_{\sigma=K|L \in \mathcal{E}_{int}} T_{\sigma} Z_{\nu}^{\sigma} |v_{\nu L} - v_{\nu K}|^2 + c \sum_{(\boldsymbol{\alpha}, \boldsymbol{\beta})} \int_{\Omega} b_{(\boldsymbol{\alpha}, \boldsymbol{\beta}), h} r_{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{R}}(e^{\mathbf{v}h}) dx \\ &=: D_1(\mathbf{v}) \end{aligned}$$

with

$$b_{(\boldsymbol{\alpha}, \boldsymbol{\beta}), h}(x) = b_{(\boldsymbol{\alpha}, \boldsymbol{\beta})K} = \frac{1}{|K|} \int_K b_{(\boldsymbol{\alpha}, \boldsymbol{\beta})}(y) dy \quad \forall x \in K \in \mathcal{V}.$$

Hence it suffices to prove

$$\widehat{\Psi}(\mathbf{u}) \leq CD_1(\mathbf{v}) \quad \forall \mathbf{v} \in \widehat{\mathcal{N}}_{\rho} \quad (3.9.37)$$

with some positive constant C independent on the mesh \mathcal{M} .

2. Assume (3.9.37) would be false, then there exist a sequence of (admissible) Voronoi meshes \mathcal{M}_n with $\mathbf{v}_n \in \widehat{\mathcal{N}}_{\rho}$, $\mathbf{u}_n = \widehat{E}\mathbf{v}_n \in \widehat{\mathcal{U}} + \mathbf{U}$ such that

$$\widehat{\Psi}(\mathbf{u}_n) \leq C_n D_1(\mathbf{v}_n) > 0, \quad \lim_{n \rightarrow \infty} C_n = +\infty. \quad (3.9.38)$$

For each \mathcal{M}_n we have corresponding quantities a_{ν, h_n} , v_{ν, h_n} , u_{ν, h_n} , \mathbf{U} , \widehat{E} , \widehat{F} , \widehat{D} , ... and sets \mathcal{E}_{int} , $\widehat{\mathcal{U}}$, $\widehat{\mathcal{N}}_{\rho}$ and indicate limits of sequences \mathcal{M}_n by $n \rightarrow \infty$. In order to simplify the notation we don't write them with an index n and call these sequence non-labeled. Let $a_{\nu K} = e^{\mathbf{v} \cdot \boldsymbol{\nu} K}$,

$K \in \mathcal{V}$ and with $a_{\nu,h}$, $v_{\nu,h}$, $u_{\nu,h}$ we denote the corresponding piecewise constant function. Since $a_{\nu} = a_{\nu}^*$ and by Lemma 3.9.2 we find

$$\|\sqrt{a_h} - \sqrt{a_h^*}\|_Y^2 \leq c \|\sqrt{u_h} - \sqrt{u_h^*}\|_Y^2 \leq \frac{c}{c_1} \widehat{\Psi}(\mathbf{u}) \leq c(\rho) \quad (3.9.39)$$

and by Corollary 3.7.4 for all \mathcal{M} the estimate

$$\|\sqrt{a_h}\|_Y \leq c(\rho) \quad (3.9.40)$$

with a constant c only depending on ρ , and on the data.

3. For $\sigma = K|L \in \mathcal{E}_{int}$ we obtain with (3.9.35) the estimate $(\sqrt{a_{\nu L}} - \sqrt{a_{\nu K}})^2 \leq Z_{\nu}^{\sigma}(v_{\nu L} - v_{\nu K})^2$ and therefore $\sum_{\nu=1}^m |\sqrt{a_{\nu,h}}|_{H^1, \mathcal{M}}^2 \leq cD_1(\mathbf{v}) \rightarrow 0$. Applying the discrete Sobolev-Poincaré inequality (B.21) we find for the function $\sqrt{a_{\nu,h}} \in X_{\mathcal{V}}(\mathcal{M})$, $\nu = 1, \dots, m$ that

$$\|\sqrt{a_{\nu,h}} - m_{\Omega}(\sqrt{a_{\nu,h}})\|_{L^2(\Omega)} \rightarrow 0 \text{ with } m_{\Omega}(\sqrt{a_{\nu,h}}) := \frac{1}{|\Omega|} \int_{\Omega} \sqrt{a_{\nu,h}} dx. \quad (3.9.41)$$

Furthermore with the discrete Sobolev-Poincaré inequality (B.21) we obtain

$$\|\sqrt{a_{\nu,h}} - m_{\Omega}(\sqrt{a_{\nu,h}})\|_{L^q(\Omega)} \leq c_q |\sqrt{a_{\nu,h}}|_{H^1, \mathcal{M}} \rightarrow 0 \quad (3.9.42)$$

with a positive constant c_q not depending on \mathcal{M} for $q \in [1, \infty)$.

Since $m_{\Omega}(\sqrt{a_{\nu,h}})|\Omega| = \|\sqrt{a_{\nu,h}}\|_{L^1} \leq c \|\sqrt{a_{\nu,h}}\|_{L^2} \leq c(\rho)$ and by (3.9.40) we obtain for $\widehat{a} \in \mathbb{R}_+^{Mm}$ and a non-labeled subsequence

$$m_{\Omega}(\sqrt{a_{\nu,h}}) \rightarrow \sqrt{\widehat{a}_{\nu}} \quad \text{in } \mathbb{R}. \quad (3.9.43)$$

All further investigations are restricted to this subsequence. Using

$$\left| \sqrt{a_{\nu,h}} - \sqrt{\widehat{a}_{\nu}} \right| \leq \left| \sqrt{a_{\nu,h}} - m_{\Omega}(\sqrt{a_{\nu,h}}) \right| + \left| m_{\Omega}(\sqrt{a_{\nu,h}}) - \sqrt{\widehat{a}_{\nu}} \right|$$

we obtain

$$\sqrt{a_{\nu,h}} \rightarrow \sqrt{\widehat{a}_{\nu}} \text{ in } L^q, \quad q \in [1, \infty), \quad \nu = 1, \dots, m. \quad (3.9.44)$$

Using binomial formula we get

$$\|a_{\nu,h} - \widehat{a}_{\nu}\|_{L^2} \leq \left\| \sqrt{a_{\nu,h}} - \sqrt{\widehat{a}_{\nu}} \right\|_{L^4}^2 + 2\sqrt{\widehat{a}_{\nu}} \left\| \sqrt{a_{\nu,h}} - \sqrt{\widehat{a}_{\nu}} \right\|_{L^2} \rightarrow 0. \quad (3.9.45)$$

4. Let $\widehat{a} = (\widehat{a}_\nu)_{\nu=1}^m$. Thanks to $\|b_{(\alpha,\beta),h}\|_{L^1} = \|b_{(\alpha,\beta)}\|_{L^1}$, (3.9.43), and (3.9.45) we obtain

$$\begin{aligned} 0 &\leq \|b_{(\alpha,\beta)}r_{(\alpha,\beta)}(\widehat{a})\|_{L^1} = \|b_{(\alpha,\beta),h}r_{(\alpha,\beta)}(\widehat{a})\|_{L^1} \\ &\leq \|b_{(\alpha,\beta),h}r_{(\alpha,\beta)}(a_h) - b_{(\alpha,\beta),h}r_{(\alpha,\beta)}(\widehat{a})\|_{L^1} + \|b_{(\alpha,\beta),h}r_{(\alpha,\beta)}(a_h)\|_{L^1} \\ &\leq \|b_{(\alpha,\beta),h}\|_{L^\infty} \|r_{(\alpha,\beta)}(a_h) - r_{(\alpha,\beta)}(\widehat{a})\|_{L^1} + cD_1(\mathbf{v}) \rightarrow 0 \end{aligned} \quad (3.9.46)$$

as $n \rightarrow \infty$ for all $(\alpha, \beta) \in \mathcal{R}$. Thus it results necessarily that

$$\widehat{a}^\alpha = \widehat{a}^\beta \quad \forall (\alpha, \beta) \in \mathcal{R}. \quad (3.9.47)$$

5. Let $\widehat{u}_\nu = \bar{u}_\nu \widehat{a}_\nu$ and $w \in \mathcal{U}^\perp$. Using $u_h \in \mathcal{U} + U$, (3.9.45), and

$$\|\bar{u}_{\nu,h}\|_{L^\infty} \leq \|\bar{u}_\nu\|_{L^\infty}, \quad \nu = 1, \dots, m \quad (3.9.48)$$

we find the estimate

$$\begin{aligned} |\langle \widehat{u} - U, w \rangle_X| &= \left| \sum_{\nu=1}^m \int_{\Omega} (\bar{u}_\nu \widehat{a}_\nu - U_\nu) w_\nu dx \right| \\ &\leq |\langle u_h - U_h, w \rangle_X| + \left| \sum_{\nu=1}^m \int_{\Omega} \bar{u}_{\nu,h} (\widehat{a}_\nu - a_{\nu,h}) w_\nu dx \right| \rightarrow 0 \end{aligned} \quad (3.9.49)$$

as $n \rightarrow \infty$ and thus $\widehat{u} \in \mathcal{U} + U$. From (3.9.47) we conclude $\widehat{a} \in \mathcal{A}$ (see (2.3.5) in Section 2.3) and therefore together with Assumption A3 we get $\widehat{a} = a^*$, hence $\widehat{u} = u^*$.

6. Due to $u_{\nu,h}^* = \bar{u}_{\nu,h} \widehat{a}_\nu$, (3.9.45), and (3.9.48) we obtain

$$\lambda_n^2 := \widehat{\Psi}(\mathbf{u}) \leq c_2 \|u_h - u_h^*\|_Y^2 \leq c_2 \|\bar{u}_h (a_h - \widehat{a})\|_Y^2 \rightarrow 0 \text{ as } n \rightarrow \infty \quad (3.9.50)$$

and by (3.9.38) we get

$$\frac{1}{C_n} = \frac{1}{\lambda_n} D_1(\mathbf{v}) \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (3.9.51)$$

7. For $\nu = 1, \dots, m$ we introduce

$$b_{\nu,h} = \frac{1}{\lambda_n} \left(\sqrt{a_{\nu,h}/\widehat{a}_\nu} - 1 \right) \in X_\nu(\mathcal{M}),$$

then by (3.9.35) we deduce $(b_{\nu,K} - b_{\nu,L})^2 \leq \frac{1}{\lambda_n^2 \widehat{a}_\nu} Z_\nu^\sigma (v_{\nu,K} - v_{\nu,L})^2$ for all $\sigma = K|L \in \mathcal{E}_{int}$ and thus

$$\sum_{\nu=1}^m |b_{\nu,h}|_{H^1, \mathcal{M}}^2 \leq \frac{c}{\lambda^2} D_1(\mathbf{v}) \rightarrow 0.$$

Applying to $b_{\nu,h} \in X_{\mathcal{V}}(\mathcal{M}_n)$ the discrete Poincaré and Sobolev-Poincaré inequality leads to

$$\|b_{\nu,h} - m_{\Omega}(b_{\nu,h})\|_{L^q} \leq c_q |b_{\nu,h}|_{H^1, \mathcal{M}} \rightarrow 0, \quad \nu = 1, \dots, m, \quad (3.9.52)$$

with $c_q > 0$ not depending on \mathcal{M} for $q \in [1, \infty)$ if $N = 2$. Using $\widehat{a}_{\nu} = a_{\nu}^* = a_{\nu,h}^*$, (3.9.39) and (3.9.50) we find

$$\begin{aligned} |m_{\Omega}(b_{\nu,h})| |\Omega| &\leq \frac{1}{\lambda_n \sqrt{\widehat{a}_{\nu}}} \left\| \sqrt{a_{\nu,h}} - \sqrt{\widehat{a}_{\nu}} \right\|_{L^1} \leq \frac{1}{\lambda_n \sqrt{a_{\nu,h}^*}} \left\| \sqrt{a_{\nu,h}} - \sqrt{a_{\nu,h}^*} \right\|_{L^1} \\ &\leq \frac{c}{\lambda_n} \left\| \sqrt{a_{\nu,h}} - \sqrt{a_{\nu,h}^*} \right\|_{L^2} \leq \frac{c}{\lambda_n} \sqrt{\widehat{\Psi}(\mathbf{u})} \leq \frac{c}{\lambda_n} \lambda_n = c. \end{aligned}$$

Thus for $\widehat{b} \in \mathbb{R}_+^{Mm}$ and a non-labeled subsequence, we obtain $m_{\Omega}(b_{\nu,h}) \rightarrow \widehat{b}_{\nu}$ in \mathbb{R} . Using

$$\left| b_{\nu,h} - \widehat{b}_{\nu} \right| \leq |b_{\nu,h} - m_{\Omega}(b_{\nu,h})| + |m_{\Omega}(b_{\nu,h}) - \widehat{b}_{\nu}|$$

we deduce

$$b_{\nu,h} \rightarrow \widehat{b}_{\nu} \quad \text{in } L^q(\Omega), \quad q \in [1, \infty), \quad \nu = 1, \dots, m. \quad (3.9.53)$$

8. For $\nu = 1, \dots, m$ we define $\widehat{y}_{\nu} := 2\widehat{b}_{\nu}u_{\nu}^* = 2\widehat{b}_{\nu}\bar{u}_{\nu}\widehat{a}_{\nu}$. Because of

$$\begin{aligned} 2b_{\nu,h}\widehat{a}_{\nu}\bar{u}_{\nu,h} &= b_{\nu,h}(\sqrt{a_{\nu,h}} + \sqrt{\widehat{a}_{\nu}})\sqrt{\widehat{a}_{\nu}}\bar{u}_{\nu,h} + g_1 \\ &= \frac{\bar{u}_{\nu,h}}{\lambda_n}(\sqrt{a_{\nu,h}} - \sqrt{\widehat{a}_{\nu}})(\sqrt{a_{\nu,h}} - \sqrt{\widehat{a}_{\nu}}) + g_1 \\ &= \frac{u_{\nu,h} - u_{\nu,h}^*}{\lambda_n} + g_1 \end{aligned}$$

with $g_1 = b_{\nu,h}(\sqrt{\widehat{a}_{\nu}} - \sqrt{a_{\nu,h}})\sqrt{\widehat{a}_{\nu}}\bar{u}_{\nu,h}$ we obtain for all $w \in \mathcal{U}^{\perp}$

$$\begin{aligned} |\langle \widehat{y}, w \rangle_X| &= 2 \left| \sum_{\nu=1}^m \int_{\Omega} b_{\nu,h}\widehat{a}_{\nu}\bar{u}_{\nu,h}w_{\nu} \, dx \right| \\ &\leq 2 \left| \sum_{\nu=1}^m \int_{\Omega} b_{\nu,h}\widehat{a}_{\nu}\bar{u}_{\nu,h}w_{\nu} \, dx \right| + 2 \left| \sum_{\nu=1}^m \int_{\Omega} (\widehat{b}_{\nu} - b_{\nu,h})\widehat{a}_{\nu}\bar{u}_{\nu,h}w_{\nu} \, dx \right| \\ &\leq \left| \frac{\langle u_h - u_h^*, w \rangle_X}{\lambda_n} \right| + c \|b_h\|_Y \left\| \sqrt{a_h} - \sqrt{\widehat{a}} \right\|_Y + c \|b_h - \widehat{b}\|_Y \|\widehat{a}\|_Y. \end{aligned}$$

Since $u_h, u_h^* \in \mathcal{U} + U$ the first term converges to zero as $n \rightarrow \infty$. Due to (3.9.44) and $\|b_h\|_Y \leq \frac{c}{\lambda_n^2} \widehat{\Psi}(U)$ the second term converges to zero as $n \rightarrow \infty$. Finally because of (3.9.53) and the boundedness of $\|\sqrt{\widehat{a}}\|_Y$ the third term converges to zero and therefore $\widehat{y} \in \mathcal{U}$ and

$$(\langle \widehat{y}_\nu, 1 \rangle)_{\nu=1}^m \in \mathcal{S}. \quad (3.9.54)$$

By the definition of $b_{\nu,h}$ and \widehat{a} we obtain for all $(\alpha, \beta) \in \mathcal{R}$

$$\begin{aligned} \widehat{a}^{-\alpha} \left(\prod_{\nu=1}^m a_{\nu,h}^{\alpha_\nu/2} - \prod_{\nu=1}^m a_{\nu,h}^{\beta_\nu/2} \right)^2 &= \left(\prod_{\nu=1}^m (\lambda_n b_{\nu,h} + 1)^{\alpha_\nu} - \prod_{\nu=1}^m (\lambda_n b_{\nu,h} + 1)^{\beta_\nu} \right)^2 \\ &= \left(\lambda_n \sum_{\nu=1}^m b_{\nu,h} (\alpha_\nu - \beta_\nu) \right)^2 + Q_n, \end{aligned} \quad (3.9.55)$$

where

$$|Q_n| \leq c \lambda_n^3 (|b_h| + 1)^{p_0}, \quad 0 \leq p_0 \leq 2 \max_{(\alpha, \beta) \in \mathcal{R}} \max \left\{ \sum_{\nu=1}^m \alpha_\nu, \sum_{\nu=1}^m \beta_\nu \right\}. \quad (3.9.56)$$

Using $\lambda_n \rightarrow 0$ as $n \rightarrow \infty$ we conclude

$$\frac{1}{\lambda_n^2} \|Q_n\|_{L^1} \leq c \lambda_n \| |b_h| + 1 \|_{L^{p_0}}^{p_0} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Together with (3.9.51) and (3.9.55) and arguments similar to Step 4 this leads to

$$\lim_{n \rightarrow \infty} \int_{\Omega} b_{(\alpha, \beta), h} \left(\sum_{\nu=1}^m b_{\nu,h} (\alpha_\nu - \beta_\nu) \right)^2 dx = 0 \quad \forall (\alpha, \beta) \in \mathcal{R},$$

and therefore

$$\widehat{b} = (\widehat{b}_\nu)_{\nu=1}^m \in \mathcal{S}^\perp. \quad (3.9.57)$$

Using the definition of \widehat{y}_ν in Step 8 and together with (3.9.54) and (3.9.57) we conclude

$$0 = \langle \widehat{y}, \widehat{b} \rangle_X = 2 \sum_{\nu=1}^m u_\nu^* \widehat{b}_\nu^2$$

and thus $\widehat{b} = 0$, and $\widehat{y} = 0$.

9. From the definition of λ_n (see (3.9.50)) it results

$$\begin{aligned}
1 &= \frac{1}{\lambda_n^2} \widehat{\Psi}(\mathbf{u}) \leq c \sum_{\nu=1}^m \|\bar{u}_{\nu,h}\|_{L^\infty} \left\| \frac{a_{\nu,h} - \widehat{a}_\nu}{\lambda_n} \right\|_{L^2}^2 \\
&\leq c \sum_{\nu=1}^m \int_{\Omega} \frac{(\sqrt{a_{\nu,h}} - \sqrt{\widehat{a}_\nu})^2}{\lambda_n^2} (\sqrt{a_{\nu,h}} + \sqrt{\widehat{a}_\nu})^2 dx \\
&\leq c \sum_{\nu=1}^m \int_{\Omega} b_{\nu,h}^2 \widehat{a}_\nu \left(\widehat{a}_\nu + |\sqrt{a_{\nu,h}} - \sqrt{\widehat{a}_\nu}|^2 \right) dx \\
&\leq c \sum_{\nu=1}^m \|b_{\nu,h}\|_{L^4}^2 \left(1 + \left\| \sqrt{a_{\nu,h}} - \sqrt{\widehat{a}_\nu} \right\|_{L^4}^2 \right).
\end{aligned}$$

Due to (3.9.53) with $\widehat{b} = 0$ and (3.9.44) we find $1 = \frac{1}{\lambda_n^2} \widehat{\Psi}(\mathbf{u}) \rightarrow 0$ as $n \rightarrow \infty$. This gives a contradiction and thus the assumption in the second step of the proof was wrong. Hence (3.9.37) holds and the proof is complete. □

Corollary 3.9.5 (Exponential decay, see [56, Theorem 3.6]).

Let the Assumption A1, A3, and A4 be fulfilled. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n \in \mathbb{N}})$ be a class of discretizations fulfilling the Assumptions A5. Then there exist constants $\lambda > 0$ and $c > 0$ only depending on the data and not on \mathcal{D} such that for every solution (u_h, v_h) to (3.3.P_D) the relative free energy $\widehat{\Psi}(\mathbf{u}(t_N))$ decays exponentially, i.e.,

$$\begin{aligned}
\widehat{\Psi}(\mathbf{u}(t_n)) &\leq e^{-\lambda t_n} \widehat{\Psi}(\mathbf{U}), \quad n \geq 1, \\
\left\| \sqrt{u_{\nu,h}(t_n)} - \sqrt{u_{\nu,h}^*} \right\|_{L^2} &\leq c e^{-\lambda t_n/2}, \quad n \geq 1, \quad \nu = 1, \dots, m,
\end{aligned}$$

holds uniformly for all discretizations \mathcal{D} .

The proof is given in [56]. In consequence of Remark 3.9.4 we rephrase the proof here. The only difference is the use of Theorem 3.9.3 instead of [56, Theorem 3.5].

Proof. From Lemma 3.4.1 we conclude $\mathbf{u}(t_n) - \mathbf{U} \in \widehat{\mathcal{U}}$. Using (3.6.22) we obtain for $n_2 > n_1 \geq 0$ and $\lambda \geq 0$ the estimate

$$\begin{aligned}
 S_1 &= e^{\lambda t_{n_2}} \widehat{\Psi}(\mathbf{u}(t_{n_2})) - e^{\lambda t_{n_1}} \widehat{\Psi}(\mathbf{u}(t_{n_1})) \\
 &= \sum_{r=n_1+1}^{n_2} \left\{ \left(e^{\lambda t_r} - e^{\lambda t_{r-1}} \right) \widehat{\Psi}(\mathbf{u}(t_r)) + e^{\lambda t_{r-1}} \left(\widehat{F}(\mathbf{u}(t_r)) - \widehat{F}(\mathbf{u}(t_{r-1})) \right) \right\} \\
 &\leq \sum_{r=n_1+1}^{n_2} \left\{ e^{\lambda t_{r-1}} \left(e^{\lambda t_\delta^{(r)}} - 1 \right) \widehat{\Psi}(\mathbf{u}(t_r)) + e^{\lambda t_{r-1}} \langle \mathbf{u}(t_r) - \mathbf{u}(t_{r-1}), \mathbf{v}(t_r) \rangle_{\mathbb{R}^{Mm}} \right\} \quad (3.9.58) \\
 &\leq \sum_{r=n_1+1}^{n_2} e^{\lambda t_{r-1}} t_\delta^{(r)} \left(e^{\lambda t_\delta} \lambda \widehat{\Psi}(\mathbf{u}(t_{n_2})) - \widehat{D}(\mathbf{v}(t_r)) \right).
 \end{aligned}$$

Using $\widehat{D}(\mathbf{v}) \geq 0$ for $\mathbf{v} \in \mathbb{R}^{Mm}$ and setting $\lambda = 0$ in (3.9.58) results in

$$\widehat{F}(\mathbf{u}(t_{n_2})) \leq \widehat{F}(\mathbf{u}(t_{n_1})) \leq \widehat{F}(\mathbf{U}) \quad \forall n_2 > n_1 \geq 0. \quad (3.9.59)$$

Using the results of Corollary 3.7.4 and Lemma 3.9.2 we obtain

$$\begin{aligned}
 \widehat{\Psi}(\mathbf{u}) &\leq c_2 \|U_h - u_h^*\|_Y^2 \leq c_2 \sum_{\nu=1}^m \max \left\{ \|U_h\|_{L^\infty}^2, \|u_h^*\|_{L^\infty}^2 \right\} |\Omega| \\
 &\leq c_2 \sum_{\nu=1}^m \max \left\{ \|U\|_{L^\infty}^2, \max_{\nu=1, \dots, m} \frac{\text{ess sup}_{x \in \Omega} \bar{u}_{\nu, h}(x)}{\text{ess inf}_{x \in \Omega} \bar{u}_\nu(x)} \|u_\nu^*\|_{L^\infty}^2 \right\} |\Omega| =: \rho
 \end{aligned}$$

The constant depends on the data but not on \mathcal{M} and therefore $\widehat{\Psi}(\mathbf{u}) \leq \rho$, $\mathbf{u}(t_r) = \widehat{E}\mathbf{v}(t_r) \in \widehat{\mathcal{U}} + \mathbf{U}$ and $\mathbf{v} \in \widehat{\mathcal{N}}_\rho$. Lemma 3.9.3 implies the existence of a constant $c_\rho > 0$ such that (3.9.34) holds for all admissible \mathcal{M} . Now choosing $\lambda > 0$ such that $\lambda e^{\lambda t_\delta} c_\rho < 1$ (see A5) again independent of the mesh and $n_1 = 0$, the estimate (3.9.58) together with Lemma 3.9.2 proves the lemma. \square

Using the L^∞ bounds from Theorem 3.8.3 we can prove the following result.

Corollary 3.9.6 (Asymptotics of the solution).

We assume A1, A2, A3, and A4. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n \in \mathbb{N}})$ be a class of discretizations fulfilling the Assumptions A5, moreover let (u_h^*, u_ν^*) be the thermodynamic equilibrium to (3.3.P_D), see Lemma 3.7.3. Then there exist constants $\lambda_p > 0$ and $c > 0$ only depending on the data and not on \mathcal{D} such that for every solution (u_h, v_h) to (3.3.P_D) and $p \in [1, \infty)$ the estimate

$$\sum_{\nu=1}^m \|u_{\nu, h}(t_n) - u_{\nu, h}^*\|_{L^p} \leq c e^{-\lambda_p t_n}, \quad n \geq 1 \quad (3.9.60)$$

holds uniformly for all discretizations \mathcal{D} .

Proof. Using Hölder's inequality we find

$$\|u_{\nu,h}(t_n) - u_{\nu,h}^*\|_{L^p}^p \leq \|u_{\nu,h}(t_n) - u_{\nu,h}^*\|_{L^1} \|u_{\nu,h}(t_n) - u_{\nu,h}^*\|_{L^\infty}^{p-1}.$$

First, we note that

$$\|u_{\nu,h}(t_n) - u_{\nu,h}^*\|_{L^1} \leq \left\| \sqrt{u_{\nu,h}(t_n)} - \sqrt{u_{\nu,h}^*} \right\|_{L^2} \left\| \sqrt{u_{\nu,h}(t_n)} + \sqrt{u_{\nu,h}^*} \right\|_{L^2}. \quad (3.9.61)$$

As a consequence of Theorem 3.8.3, we obtain with two positive constants c_1, c_2 the boundedness of

$$\|u_{\nu,h}(t_n) - u_{\nu,h}^*\|_{L^\infty}^{p-1} \leq c_1^{p-1}, \quad \left\| \sqrt{u_{\nu,h}(t_n)} + \sqrt{u_{\nu,h}^*} \right\|_{L^2} \leq c_2$$

and find by Corollary 3.9.5 the desired estimate (3.9.60). \square

3.10 Global lower bounds

Now, we intend to show global lower bounds of the densities or in other words upper bounds of the negative part of the chemical potentials. In the continuous setting, this was done in [46] and [63, p. 18]. In a first step we need lower bounds in L^1 which provide a suitable start for the Moser iteration.

Lemma 3.10.1 (Lower bounds in $L^1(\Omega)$).

We assume A1, A2, A3, and A4. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n \in \mathbb{N}})$ be a class of discretizations fulfilling the Assumptions A5. Then there exists a constant $c_1 > 0$ only depending on the data and not on \mathcal{D} such that for every solution (u_h, v_h) to (3.3.PD)

$$\left\| v_{\nu,h}^-(t_N) \right\|_{L^1} \leq c_1 \quad \forall N \geq 1, \quad \nu = 1, \dots, m$$

holds uniformly for all discretizations \mathcal{D} .

Proof. Following [63, p. 18] we define the convex and lower semicontinuous functional $\widehat{\Theta} : \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$ by

$$\widehat{\Theta}(\mathbf{w}) = \sum_{K \in \mathcal{V}} |K| u_{\nu,K}^* \vartheta(w_K), \quad \vartheta(y) := \begin{cases} -\ln(1-y), & \text{for } y \leq 0, \\ \infty & \text{for } y > 0, \end{cases}$$

and its conjugate convex functional

$$\widehat{G}(\mathbf{u}_\nu) = \sup_{\mathbf{w} \in \mathbb{R}^M} \{ \langle \mathbf{u}_\nu, \mathbf{w} \rangle - \widehat{\Theta}(\mathbf{w}) \}, \quad (3.10.62)$$

which may be written in the explicit form

$$\widehat{G}(\mathbf{u}_\nu) = \sum_{K \in \mathcal{V}} |K| \left\{ u_{\nu K}^* \left(\ln \frac{u_{\nu K}}{u_{\nu K}^*} \right)^- - (u_{\nu K} - u_{\nu K}^*)^- \right\}. \quad (3.10.63)$$

Here \mathbf{u}^* is the unique stationary solution to (3.3. P_D) according to Lemma 3.7.3. We introduce $\bar{z}_\nu := ((1 - u_{\nu K}^*/u_{\nu K})^-)_{K \in \mathcal{V}}$ and the corresponding $\bar{z}_{\nu,h} \in X_{\mathcal{V}}(\mathcal{M})$, and observe that $-\bar{z}_\nu \in \partial \widehat{G}(\mathbf{u}_\nu)$. Testing the discrete problem (3.3. P_D) with the test function $(0, \dots, 0, -\bar{z}_\nu, 0, \dots, 0)$ results in $S_1 = \sum_{n=1}^N t_\delta^{(n)} (S_2 + S_3)$ with

$$\begin{aligned} S_1 &:= - \sum_{n=1}^N t_\delta^{(n)} \left\langle \frac{\mathbf{u}_\nu(t_n) - \mathbf{u}_\nu(t_{n-1})}{t_\delta^{(n)}}, \bar{z}_\nu \right\rangle_{\mathbb{R}^M}, \\ S_2 &:= \sum_{\sigma=K|L \in \mathcal{E}_{\text{int}}} T_\sigma Y_\nu^\sigma Z_\nu^\sigma (v_{\nu L} - v_{\nu K}) (\bar{z}_{\nu L} - \bar{z}_{\nu K}), \\ S_3 &:= - \sum_{K \in \mathcal{V}} |K| R_{\nu K} (e^{v_K}) \bar{z}_{\nu K}. \end{aligned}$$

Using the subdifferential property of \widehat{G} we find

$$\widehat{G}(\mathbf{u}_\nu(t_N)) - \widehat{G}(\mathbf{U}_\nu) = \sum_{n=1}^N \widehat{G}(\mathbf{u}_\nu(t_n)) - \widehat{G}(\mathbf{u}_\nu(t_{n-1})) \leq S_1. \quad (3.10.64)$$

Both terms S_2 and S_3 will be estimated on the subsets

$$\Omega_+(t_n) = \{K \in \mathcal{V} : u_{\nu K}(t_n) \geq u_{\nu K}^*\}, \quad \Omega_-(t_n) = \{K \in \mathcal{V} : u_{\nu K}(t_n) < u_{\nu K}^*\}.$$

Diffusion: In a first step we show that the diffusion term S_2 is nonpositive. We remark that $\bar{z}_{\nu K} = 0$ for all $K \in \Omega_+(t_n)$. Using (3.2.2) and $a_{\nu K}^* = a_{\nu L}^* \forall K, L \in \mathcal{V}$, we find $S_2 = S_{21} - S_{22}$ with

$$\begin{aligned} S_{21} &:= \sum_{\substack{\sigma=K|L \in \mathcal{E}_{\text{int}} \\ K, L \in \Omega_-(t_n)}} T_\sigma Y_\nu^\sigma (a_{\nu L} - a_{\nu K}) (1/a_{\nu L} - 1/a_{\nu K}) a_{\nu K}^*, \\ S_{22} &:= \sum_{\substack{\sigma=K|L \in \mathcal{E}_{\text{int}} \\ K \in \Omega_-(t_n), L \in \Omega_+(t_n)}} T_\sigma Y_\nu^\sigma ((a_{\nu L} - a_{\nu L}^*) + (a_{\nu K}^* - a_{\nu K})) \bar{z}_{\nu K}. \end{aligned}$$

Using $(x - y)(1/x - 1/y) \leq 0$ for all $x, y > 0$ we get $S_{21} \leq 0$. If $K \in \Omega_-(t_n)$ holds, the term $a_{\nu K}^* - a_{\nu K}$ is positive. On the other hand if $L \in \Omega_+(t_n)$ we have $a_{\nu L} - a_{\nu L}^* \leq 0$ and therefore $S_{22} \geq 0$ and finally $S_2 \leq 0$.

Reactions: On Ω_+ reaction terms multiplied by the test function vanish. Since $(\alpha - \beta) \cdot \mathbf{v}_K^* = 0$ $\forall K \in \mathcal{V}$, we find on $\Omega_-(t_n)$

$$\begin{aligned} S_3 &= \sum_{K \in \Omega_-} |K| \sum_{(\alpha, \beta) \in \mathcal{R}} k_{(\alpha, \beta)K} e^{\alpha \cdot \mathbf{v}_K^*} \left(\prod_{j=1}^m \left(\frac{u_{jK}}{u_{jK}^*} \right)^{\alpha_j} - \prod_{j=1}^m \left(\frac{u_{jK}}{u_{jK}^*} \right)^{\beta_j} \right) (\alpha_\nu - \beta_\nu) \bar{z}_{\nu K} \\ &= \sum_{K \in \Omega_-} |K| (S_{31K} + S_{32K}) \end{aligned}$$

with

$$\begin{aligned} S_{31K} &= - \sum_{(\alpha, \beta) \in \mathcal{R}} k_{(\alpha, \beta)K} e^{\alpha \cdot \mathbf{v}_K^*} \left(\prod_{j=1}^m \left(\frac{u_{jK}}{u_{jK}^*} \right)^{\alpha_j} - \prod_{j=1}^m \left(\frac{u_{jK}}{u_{jK}^*} \right)^{\beta_j} \right) (\alpha_\nu - \beta_\nu), \\ S_{32K} &= \sum_{(\alpha, \beta) \in \mathcal{R}} k_{(\alpha, \beta)K} e^{\alpha \cdot \mathbf{v}_K^*} \left(\prod_{j=1}^m \left(\frac{u_{jK}}{u_{jK}^*} \right)^{\alpha_j} - \prod_{j=1}^m \left(\frac{u_{jK}}{u_{jK}^*} \right)^{\beta_j} \right) (\alpha_\nu - \beta_\nu) \frac{u_{\nu K}^*}{u_{\nu K}}. \end{aligned}$$

The terms in the parentheses of S_{31K} are Lipschitz continuous in $(\mathbf{u}_K/\mathbf{u}_K^*)$ on $[0, R]^m$, $R > 0$ and have at $(1)_{\nu=1}^m$ the value 0. Together with the global (upper) boundedness of $(\mathbf{u}_K/\mathbf{u}_K^*)$ we obtain

$$S_{31K} \leq C_1 \sum_{\nu=1}^m \left| \frac{u_{\nu K}}{u_{\nu K}^*} - 1 \right|.$$

We consider two cases: Note that $1 < u_{\nu K}^*/u_{\nu K}$ for all $K \in \Omega_-(t_n)$. For $\alpha_\nu > \beta_\nu$ the summands in S_{32K} can be estimated by

$$\left(\frac{u_{\nu K}}{u_{\nu K}^*} \right)^{\alpha_\nu - 1} \prod_{\substack{j=1, \\ j \neq \nu}}^m \left(\frac{u_{jK}}{u_{jK}^*} \right)^{\alpha_j} - \prod_{j=1}^m \left(\frac{u_{jK}}{u_{jK}^*} \right)^{\beta_j}.$$

This term is also Lipschitz continuous in $(\mathbf{u}_K/\mathbf{u}_K^*)$ on $[0, R]^m$, $R > 0$, and has at $(1)_{\nu=1}^m$ the value 0. Together with the global boundedness of $(\mathbf{u}_K/\mathbf{u}_K^*)$ we obtain the bound

$$C_3 \sum_{\nu=1}^m \left| \frac{u_{\nu K}}{u_{\nu K}^*} - 1 \right|.$$

The case $\alpha_\nu < \beta_\nu$ can be handled analogously. Therefore there exists a constant $C_4 > 0$ such that

$$S_3 \leq C_4 \sum_{\nu=1}^m \left\| u_{\nu, h} / u_{\nu, h}^* - 1 \right\|_{L^1}.$$

L^1 -estimate: From (3.10.64) together with $S_1 = \sum_{n=1}^N t_\delta^{(n)} (S_2 + S_3)$ and the estimates of S_2 and S_3 we conclude

$$\widehat{G}(\mathbf{u}_\nu(t_N)) \leq \widehat{G}(\mathbf{U}_\nu) + C_4 \sum_{n=1}^N \sum_{\nu=1}^m t_\delta^{(n)} \|u_{\nu,h}(t_n)/u_{\nu,h}^* - 1\|_{L^1}.$$

From Corollary 3.9.6 and Lemma 3.9.5 we conclude that

$$\left\| \frac{u_{\nu,h}(t_n)}{u_{\nu,h}^*} - 1 \right\|_{L^1} \leq C_5 e^{-\lambda t_n/2} \quad \forall n \geq 0,$$

with the constant $\lambda > 0$ given in Lemma 3.9.5. Hence there exists a constant $C_6 > 0$ such that $\widehat{G}(\mathbf{u}_\nu(t_N)) \leq C_6$. Let $\underline{c}_{u^*} = \min_{\nu=1,\dots,m} \operatorname{ess\,inf}_{x \in \Omega} u_\nu^*(x)$. Together with

$$\|(u_{\nu,h}(t_n) - u_{\nu,h}^*)^-\|_{L^1} \leq \|u_{\nu,h}^*\|_{L^1} \quad \forall n \geq 0$$

and the explicit form of \widehat{G} in (3.10.63) we find for all $N \geq 0$

$$\|(v_{\nu,h}(t_N) - v_{\nu,h}^*)^-\|_{L^1} = \left\| \left(\ln \frac{u_{\nu,h}(t_N)}{u_{\nu,h}^*} \right)^- \right\|_{L^1} \leq \frac{1}{\underline{c}_{u^*}} \left(\widehat{G}(\mathbf{u}_\nu(t_N)) + \|u_{\nu,h}^*\|_{L^1} \right)$$

from which the uniform bound in L^1 for every species $\nu = 1, \dots, m$ and all considered \mathcal{D} follows. \square

Now, we show global lower bounds for the chemical potentials by Moser iteration.

Theorem 3.10.2 (Lower bounds in L^∞).

We assume A1, A2, A3, and A4. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n \in \mathbb{N}})$ be a class of discretizations fulfilling the Assumptions A5. Then there exists a constant $c > 0$ only depending on the data and not on \mathcal{D} such that for every solution (u_h, v_h) to (3.3.PD)

$$\|v_{\nu,h}^-(t_N)\|_{L^\infty} \leq c \quad \forall N \geq 1, \quad \nu = 1, \dots, m$$

holds uniformly for all discretizations \mathcal{D} .

Remark 3.10.3. The proof is based on Moser iteration, too. In [63, Lemma 4.2 and Theorem 4.3] this technique was applied to the continuous case. For $p = 2^k$, $k \geq 1$ one takes the test function which has the ν -th component $-p e^t z_\nu^{p-1} e^{-v_\nu(t)}$ with $z_\nu = (v_\nu + \kappa)^-$, $\kappa = \max_{\nu=1,\dots,m} \|(v_\nu(0))^- \|_{L^\infty}$, the other components are zero. As already mentioned, in the discrete case the test functions have to be modified.

Proof. Let $z_{\nu,h} = (v_{\nu,h} + \kappa)^-$ and $w_{\nu,h} = z_{\nu,h}^{p/2}$. The constant κ is defined by

$$\kappa := \max_{\nu=1,\dots,m} \left(\ln \frac{\text{ess inf}_{x \in \Omega} U_\nu(x)}{\text{ess sup}_{x \in \Omega} \bar{u}_\nu(x)} \right)^-$$

and therefore $z_{\nu,h}(t_0) = 0$ holds. For $p \geq 2$ we test the discrete Problem (3.3. P_D) with test functions which have the ν -th component

$$-p e^{t_{n-1}} z_{\nu,h}^{p-1}(t_n) e^{-v_{\nu,h}(t_n)},$$

the other components are zero. Doing this we derive $S_1 = S_2 + S_3$ with

$$\begin{aligned} S_1 &:= -p \sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \left\langle \frac{\mathbf{u}(t_n) - \mathbf{u}(t_{n-1})}{t_\delta^{(n)}}, \mathbf{z}^{p-1}(t_n) e^{-v_{\nu,h}(t_n)} \right\rangle_{\mathbb{R}^M}, \\ S_2 &:= \sum_{n=1}^N t_\delta^{(n)} p e^{t_{n-1}} \sum_{\sigma=K|L \in \mathcal{E}_{\text{int}}} T_\sigma Y_\nu^\sigma Z_\nu^\sigma (v_{\nu L} - v_{\nu K}) (z_{\nu L}^{p-1} e^{-v_{\nu L}} - z_{\nu K}^{p-1} e^{-v_{\nu K}}), \\ S_3 &:= -p \sum_{K \in \mathcal{V}} |K| \sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} z_{\nu K}^{p-1}(t_n) e^{-v_{\nu K}(t_n)} R_{\nu K}(e^{v_K}). \end{aligned}$$

We estimate the different parts separately:

Time derivative: Since $e^x - e^y = e^\xi(x-y)$ holds for some $\xi \in [x, y] \subset \mathbb{R}$ and $z_{\nu K}^{p-1}(t_n)(v_{\nu K}(t_n) + \kappa)^+ = 0$ we find

$$\begin{aligned} I_{\nu K}(t_n) &:= -p e^{t_{n-1}} z_{\nu K}^{p-1}(t_n) e^{-v_{\nu K}(t_n)} (u_{\nu K}(t_n) - u_{\nu K}(t_{n-1})) \\ &\geq p e^{t_{n-1}} z_{\nu K}^{p-1}(t_n) \bar{u}_{\nu K} e^{\xi_{\nu K} - v_{\nu K}(t_n)} (z_{\nu K}(t_n) - z_{\nu K}(t_{n-1})). \end{aligned}$$

In the following we consider the two cases:

1. From $u_{\nu K}(t_n) > u_{\nu K}(t_{n-1})$ we get $z_{\nu K}(t_n) \leq z_{\nu K}(t_{n-1})$ and $e^{\xi_{\nu K}} < e^{v_{\nu K}(t_n)}$, hence

$$I_{\nu K}(t_n) \geq p e^{t_{n-1}} z_{\nu K}^{p-1}(t_n) \bar{u}_{\nu K} (z_{\nu K}(t_n) - z_{\nu K}(t_{n-1})).$$

2. From $u_{\nu K}(t_n) < u_{\nu K}(t_{n-1})$ we find $z_{\nu K}(t_n) \geq z_{\nu K}(t_{n-1})$ and $e^{\xi_{\nu K}} > e^{v_{\nu K}(t_n)}$, hence

$$I_{\nu K}(t_n) \geq p e^{t_{n-1}} z_{\nu K}^{p-1}(t_n) \bar{u}_{\nu K} (z_{\nu K}(t_n) - z_{\nu K}(t_{n-1})).$$

Together with (C.27) and (3.8.25) we estimate S_1 by

$$\begin{aligned}
 S_1 &= \sum_{n=1}^N \sum_{K \in \mathcal{V}} |K| I_{\nu K}(t_n) \\
 &\geq \sum_{n=1}^N \sum_{K \in \mathcal{V}} |K| \bar{u}_{\nu K} \left\{ (e^{t_n} z_{\nu K}(t_n)^p - e^{t_{n-1}} z_{\nu K}(t_{n-1})^p) - (e^{t_n} - e^{t_{n-1}}) z_{\nu K}(t_n)^p \right\} \\
 &\geq e^{t_N} \underline{c}_{\bar{u}} \|z_{\nu, h}(t_N)\|_{L^p}^p - \sum_{n=1}^N t_\delta^{(n)} e^{\bar{t}_\delta} e^{t_{n-1}} \|\bar{u}_{\nu, h}\|_{L^\infty} \|z_{\nu, h}(t_n)\|_{L^p}^p,
 \end{aligned}$$

where the second line is a telescoping sum. Here \bar{t}_δ denotes the largest time step of all considered discretizations, see Assumption A5.

Diffusion term: A short calculation gives for $\sigma = K|L$ and $x_{\nu KL} := v_{\nu L} - v_{\nu K}$

$$Z_\nu^\sigma(v_{\nu L} - v_{\nu K})(z_{\nu L}^{p-1} e^{-v_{\nu L}} - z_{\nu K}^{p-1} e^{-v_{\nu K}}) = A + B$$

with

$$\begin{aligned}
 A &:= \frac{(e^{x_{\nu KL}} - 1)(e^{-x_{\nu KL}} + 1)}{2x_{\nu KL}} x_{\nu KL} (z_{\nu L}^{p-1} - z_{\nu K}^{p-1}), \\
 B &:= \frac{(e^{x_{\nu KL}} - 1)(e^{-x_{\nu KL}} - 1)}{x_{\nu KL}^2} \frac{x_{\nu KL}^2}{2} (z_{\nu L}^{p-1} + z_{\nu K}^{p-1}).
 \end{aligned}$$

Using Lemma C.2, inequality (C.26) and the following auxiliary calculation (with $x_{\nu L} := (v_{\nu L} + \kappa)^+$ and $x_{\nu K} := (v_{\nu K} + \kappa)^+$)

$$\begin{aligned}
 (v_{\nu L} - v_{\nu K})(z_{\nu L}^{p-1} - z_{\nu K}^{p-1}) &= (x_{\nu L} - x_{\nu K})(z_{\nu L}^{p-1} - z_{\nu K}^{p-1}) - (z_{\nu L} - z_{\nu K})(z_{\nu L}^{p-1} - z_{\nu K}^{p-1}) \\
 &= - (x_{\nu L} z_{\nu K}^{p-1} + x_{\nu K} z_{\nu L}^{p-1}) - (z_{\nu L} - z_{\nu K})(z_{\nu L}^{p-1} - z_{\nu K}^{p-1}) \\
 &\leq - (z_{\nu L} - z_{\nu K})(z_{\nu L}^{p-1} - z_{\nu K}^{p-1}),
 \end{aligned}$$

we can estimate the term A from above by

$$A \leq -\frac{4(p-1)}{p^2} (z_{\nu L}^{p/2} - z_{\nu K}^{p/2})^2.$$

Together with Lemma C.2, inequality (C.28) and the auxiliary calculation

$$\begin{aligned}
 x_{\nu KL}^2 &= (v_{\nu L} - v_{\nu K})^2 = ((x_{\nu L} - x_{\nu K}) - (z_{\nu L} - z_{\nu K}))^2 \\
 &= (x_{\nu L} - x_{\nu K})^2 + 2(x_{\nu L} z_{\nu K} + x_{\nu K} z_{\nu L}) + (z_{\nu L} - z_{\nu K})^2 \\
 &\geq (z_{\nu L} - z_{\nu K})^2,
 \end{aligned}$$

we can bound the term B by

$$B \leq -(z_{\nu L} - z_{\nu K})^2 \frac{(z_{\nu L}^{p-1} + z_{\nu K}^{p-1})}{2} \leq -\frac{1}{(p+1)^2} \left(z_{\nu L}^{\frac{p+1}{2}} - z_{\nu K}^{\frac{p+1}{2}} \right)^2.$$

Therefore we can bound S_2 with some constant $\underline{c}_{D\bar{u}} = \underline{c}_D \underline{c}_{\bar{u}}$ by

$$\begin{aligned} S_2 &\leq -\sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} 4\underline{c}_{D\bar{u}} \left(\frac{p-1}{p} \left| z_{\nu,h}^{p/2} \right|_{H^1, \mathcal{M}}^2 + \frac{p}{4(p+1)^2} \left| z_{\nu,h}^{\frac{p+1}{2}} \right|_{H^1, \mathcal{M}}^2 \right) \\ &\leq \sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \underline{c}_{D\bar{u}} \left\{ -2\|w_{\nu,h}\|_{H^1, \mathcal{M}}^2 + p\|z_{\nu,h}\|_{L^p}^p \right\}. \end{aligned}$$

The last term in the first line of the above inequalities can be neglected, since only the first H^1 -semi-norm is needed in the following estimate. In the last line we extend the H^1 -seminorm to the full H^1 -norm and use $1/2 \leq 1 - 1/p < 1$ and $0 < (p-1)/p^2 \leq 1/4$ for $p \in [2, \infty)$.

Reaction terms: The reaction terms multiplied by the test function can be written as

$$-R_{\nu K}(e^{\mathbf{v}_K} z_{\nu K}^{p-1} e^{-v_{\nu K}}) = \sum_{(\alpha, \beta) \in \mathcal{R}} k_{(\alpha, \beta)K} (\mathbf{a}_K^\alpha - \mathbf{a}_K^\beta) (\alpha_\nu - \beta_\nu) z_{\nu K}^{p-1} e^{-v_{\nu K}}.$$

Using the L^∞ bounds of Theorem 3.8.3 we deduce for $\alpha_\nu > \beta_\nu$ that

$$(\mathbf{a}_K^\alpha - \mathbf{a}_K^\beta) (\alpha_\nu - \beta_\nu) e^{-v_{\nu K}} = (\alpha_\nu - \beta_\nu) \left(a_{\nu K}^{(\alpha_\nu-1)} \prod_{\substack{j=1 \\ j \neq \nu}}^m a_{\nu K}^{\alpha_j} - a_{\nu K}^{(\beta_\nu-1)} \prod_{\substack{j=1 \\ j \neq \nu}}^m a_{\nu K}^{\beta_j} \right) \leq C_1.$$

A corresponding estimate holds for $\alpha_\nu < \beta_\nu$ and therefore we get with $C_2 > 0$

$$S_3 \leq C_2 \sum_{n=1}^N p t_\delta^{(n)} e^{t_{n-1}} \|z_{\nu K}\|_{L^{p-1}}^{p-1}.$$

Moser iteration: From $S_1 = S_2 + S_3$ together with $x^{p-1} \leq x^p + 1$ for $x \geq 0$ and $w_{\nu,h} = z_{\nu,h}^{p/2}$ for $p \geq 2$ we conclude with some constant $C_3 > 0$

$$\begin{aligned} S_4 &:= e^{t_N} \underline{c}_{\bar{u}} \|z_{\nu,h}(t_N)\|_{L^p}^p \\ &\leq \sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \left\{ -2\underline{c}_{D\bar{u}} \|w_{\nu,h}\|_{H^1}^2 + pC_3 (\|w_{\nu,h}\|_{L^2}^2 + 1) \right\}. \end{aligned}$$

Using Gagliardo-Nirenberg's inequality (B.19) and Young's inequality with $(p = 2)$ we can estimate

$$\begin{aligned} \|w_{\nu,h}\|_{L^2}^2 &\leq c_{gn,2}^2 \|w_{\nu,h}\|_{L^1} \|w_{\nu,h}\|_{H^1} \\ &\leq \frac{c_{gn,2}^2}{2} \left(\varepsilon \|w_{\nu,h}\|_{H^1}^2 + \varepsilon^{-1} \|w_{\nu,h}\|_{L^1}^2 \right), \quad c_{gn,2}^2 = \sqrt{2}C^2, \end{aligned}$$

and the estimate of S_4 can be continued by

$$S_4 \leq \sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \left\{ g_1(\varepsilon) \|w_{\nu,h}\|_{H^1}^2 + p^2 C_3 (g_2(\varepsilon) \|w_{\nu,h}\|_{L^1}^2 + 1) \right\}$$

with

$$g_1(\varepsilon) = -2c_{D\bar{u}} + p\varepsilon \frac{C_3 c_{gn,2}^2}{2} \quad \text{and} \quad g_2(\varepsilon) = \frac{c_{gn,2}^2}{2p\varepsilon}.$$

The constant ε is fixed such that $g_1(\varepsilon) = 0$ holds, i.e.,

$$\varepsilon := \frac{4c_{D\bar{u}}}{C_3 c_{gn,2}^2 p} \quad \text{and then} \quad g_2(\varepsilon) = C_3 \frac{c_{gn,2}^4}{8c_{D\bar{u}}}.$$

Setting $C_4 := C_3 \max(g_2(\varepsilon), 1)$ we find

$$e^{-t_N} S_4 = c_{\bar{u}} \|z_{\nu,h}(t_N)\|_{L^p}^p \leq C_4 p^2 e^{-t_N} \left(\sum_{n=1}^N t_\delta^{(n)} e^{t_{n-1}} \|z_{\nu,h}\|_{L^{p/2}}^p + 1 \right).$$

Now we proceed in a similar way as in the proof of Theorem 3.8.3. We set

$$b_k = \sup_{n=1,\dots,N} \|z_{\nu,h}(t_n)\|_{L^{2^k}}^{2^k} + 1, \quad k \geq 0.$$

Moreover we set $p = 2^k$ for $k \geq 1$. Together with

$$\|z_{\nu,h}(t_N)\|_{L^p}^p + 1 \leq C_5 p^2 \sup_{n=1,\dots,N} \left(\|z_{\nu,h}\|_{L^{p/2}}^{p/2} + 1 \right)^2, \quad p \geq 2,$$

where $C_5 > 0$, we find for all $k \geq 1$ the recursion formula

$$b_k \leq 2^{2k} C_5 (b_{k-1})^2 \leq \left\{ (4)^{\sum_{i=0}^{k-1} (k-i)2^i} (C_5)^{\sum_{i=0}^{k-1} 2^i} b_0^{2^k} \right\}.$$

Applying (3.8.31) we conclude $b_k \leq (4^4 C_5 b_0)^{2^k}$ and

$$\|z_{\nu,h}(t_N)\|_{L^{2^k}} \leq 4^4 C_5 b_0, \quad k \geq 1.$$

This estimate holds true for all times t_N , $N \in \mathbb{N}$. The term b_0 is bounded by Lemma 3.10.1. Passing to the limit $k \rightarrow \infty$ we obtain

$$\|z_{\nu,h}(t_N)\|_{L^\infty} \leq 4^4 C_5 b_0 \leq C_8.$$

The procedure can be done for $\nu = 1, \dots, m$ and the result of the theorem follows with $v_{\nu,h}^- = v_{\nu,h}^+ + \kappa - (v_{\nu,h} + \kappa) \leq v_{\nu,h}^+ + \kappa + z_{\nu,h}$ and the bound of $v_{\nu,h}^+$ from Theorem 3.8.3. \square

3.11 Global existence and uniqueness

In this section we prove the uniqueness of the discrete solution. Only for this result we additionally assume that the diffusion coefficients only depend on the space variable and not on the state variable. The results in other sections are independent of this assumption.

Theorem 3.11.1 (Uniqueness of the discrete solution).

Let the Assumption A1, A2, and A4 be fulfilled and let $\mathcal{D} = (\mathcal{M}, (t_n)_{n \in \mathbb{N}})$ be a class of discretizations fulfilling the Assumptions A5. Assuming that the diffusion coefficients only depend on the space variable and not on the state variable and suppose the time step restriction

$$t_\delta^{(n)} < \frac{C_{\bar{u}}}{2C_2} \quad \forall n = 1, \dots, N, \quad (3.11.65)$$

where C_2 depends on the data. Then there exists at most one solution to the discrete problem (3.3.P_D).

Proof. It suffices to prove uniqueness on a time step t_{n-1} to t_n . Assuming that (3.3.P_D) has two solutions $(\mathbf{u}(t_n), \mathbf{v}(t_n))$ and $(\hat{\mathbf{u}}(t_n), \hat{\mathbf{v}}(t_n))$ for the same initial value $(\mathbf{u}(t_{n-1}), \mathbf{v}(t_{n-1}))$. We test (3.3.P_D) by the difference of the two solutions $\tilde{\mathbf{a}} = \mathbf{e}^{\mathbf{v}} - \mathbf{e}^{\hat{\mathbf{v}}}$. From (3.3.4) we deduce $S_1 + S_2 = S_3$ with

$$\begin{aligned} S_1 &:= \left\langle \frac{\mathbf{u}(t_n) - \mathbf{u}(t_{n-1})}{t_\delta^{(n)}}, \tilde{\mathbf{a}}(t_n) \right\rangle_{\mathbb{R}^{Mm}} - \left\langle \frac{\hat{\mathbf{u}}(t_n) - \hat{\mathbf{u}}(t_{n-1})}{t_\delta^{(n)}}, \tilde{\mathbf{a}}(t_n) \right\rangle_{\mathbb{R}^{Mm}} \\ &= \left\langle \frac{\tilde{\mathbf{u}}(t_n) - \tilde{\mathbf{u}}(t_{n-1})}{t_\delta^{(n)}}, \tilde{\mathbf{a}}(t_n) \right\rangle_{\mathbb{R}^{Mm}} \end{aligned}$$

and

$$\begin{aligned} S_2 &:= \sum_{\nu=1}^m \sum_{\sigma=K|L \in \mathcal{E}_{int}} T_\sigma Y_\nu^\sigma ((a_{\nu L} - a_{\nu K}) - (\hat{a}_{\nu L} - \hat{a}_{\nu K})) (\tilde{a}_{\nu L} - \tilde{a}_{\nu K}), \\ S_3 &:= \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| (R_{\nu K}(\mathbf{a}_K) - R_{\nu K}(\hat{\mathbf{a}}_K)) \tilde{a}_{\nu K}. \end{aligned}$$

Since the diffusion coefficients only depend on the space variable, the averaging Y_ν^σ is equal for both solutions. Since $(x - y)x \geq \frac{1}{2}(x^2 - y^2)$ and $\|\tilde{a}_{\nu,h}(t_{n-1})\|_{L^2}^2 = 0$ the term S_1 can be estimated by

$$S_1 \geq \sum_{\nu=1}^m \frac{1}{t_\delta^{(n)}} \left\{ \frac{\underline{c}_{\bar{u}}}{2} \|\tilde{a}_{\nu,h}(t_n)\|_{L^2}^2 - \frac{\bar{c}_{\bar{u}}}{2} \|\tilde{a}_{\nu,h}(t_{n-1})\|_{L^2}^2 \right\} = \frac{\underline{c}_{\bar{u}}}{2t_\delta^{(n)}} \|\tilde{a}_h(t_n)\|_Y^2.$$

By Assumption A1 the flux term S_2 can be estimated with a constant $C_1 > 0$ as follows

$$S_2 = \sum_{\nu=1}^m \sum_{\sigma \in \mathcal{E}_{int}} T_\sigma Y_\nu^\sigma (\tilde{a}_{\nu L} - \tilde{a}_{\nu K})^2 \geq C_1 \underline{c}_{\bar{u}} \sum_{\nu=1}^m |\tilde{a}_{\nu,h}|_{H^1, \mathcal{M}}^2.$$

Due to Theorem 3.8.3 (global bounds for the arguments) and the uniformly local Lipschitz continuity of the reaction terms we find with $C_2 > 0$

$$|S_3| \leq \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| \|R_{\nu K}(\mathbf{a}_K) - R_{\nu K}(\hat{\mathbf{a}}_K)\| |\tilde{\mathbf{a}}| \leq C_2 \sum_{\nu=1}^m \|\tilde{a}_{\nu,h}\|_{L^2}^2.$$

From $S_1 + S_2 = S_3$ we find

$$\sum_{\nu=1}^m \left\{ \underline{c}_{\bar{u}} \left(\frac{1}{2t_\delta^{(n)}} - \frac{C_2}{\underline{c}_{\bar{u}}} \right) \|\tilde{a}_{\nu,h}\|_{L^2}^2 + \underline{c}_{\bar{u}} C_1 |\tilde{a}_{\nu,h}|_{H^1, \mathcal{M}}^2 \right\} \leq 0.$$

The assumption (3.11.65) ensures that the term in front of the L^2 -norm is greater zero. Hence $\tilde{a}_{\nu,h} = 0$ for all $n > 0$ and $\nu = 1, \dots, m$ and the two solutions coincide. \square

Remark 3.11.2. *The uniqueness result, in particular the restriction on the time step, is not satisfying. The time step restriction is due to the use of a discrete Gronwall lemma. To our knowledge, the question of uniqueness of the discrete solution is open in the case of diffusion coefficients depending on the state. Standard arguments are not applicable due to the nonlinearity of the flux term. As already mentioned in [46], additional regularity assumptions, like the boundedness of the gradient in L^∞ , exclude practically relevant geometries and heterogeneous materials. For the continuous problem, the question of uniqueness was discussed in [46, 53] (see also Section 2.3).*

At the end of this section we summarize the results which we have obtained for the discrete Problem (3.3. P_D) under the Assumptions A1 - A4, completed by the Assumption A5 on the time discretization.

Theorem 3.11.3 (Existence of global positive bounded solutions).

We assume A1, A2, A3, and A4. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n \in \mathbb{N}})$ be class of discretizations fulfilling Assumptions A5. Then there is at least one solution of the discrete Problem (3.3. P_D). For this solution global estimates as in Theorem 3.8.3 and 3.10.2 are satisfied. Moreover the asymptotic behavior as in Corollary

3.9.5 and Corollary 3.10.1 are valid uniformly on all considered discretions \mathcal{D} . If in addition the diffusion coefficient are independent of the state variable the solution is unique under the time step restriction (3.11.65).

3.12 Convergence

3.12.1 Introduction

In this section, we investigate the convergence of the scheme. The convergence for linear parabolic problems is studied in [32, 34]. However the convergence of schemes for nonlinear parabolic equations and systems is still an ongoing research topic, see [7, 9, 33].

Convergence of the discrete problem (3.3. P_D) is more involved than the usual parabolic setting due to the Boltzmann statistics, the nonlinear diffusion and the nonsmooth coefficients. These difficulties make it necessary to develop some new tools to handle the statistics and the reference densities. The auxiliary results are collected in Subsection A.3.2.

We first introduce the prolonged quantities, i.e., the piecewise affine functions in space and time and the piecewise linear in space functions for the gradients in Subsection 3.12.2. Since we are in two space dimensions the discrete H^1 -seminorm and the linear finite element H^1 -seminorm are equal. Following the proof of [64, Lemma 6.4], we provide a priori estimates for the gradient and for the time derivative in Subsection 3.12.3. The proof relies on the uniform lower and upper bounds on the concentrations from Section 3.8 and 3.10. Together with a priori estimates, we derive in Subsection 3.12.4 the existence of a converging subsequence and prove that the limit is a weak solution of the continuous problem (2.2.P). Finally we derive strong convergence results.

In contrast to the previous sections, where all results have been proven on the infinite time interval \mathbb{R}_+ , we restrict our investigations to an arbitrary finite time interval $S = [0, T] \subset \mathbb{R}_+$ in order to obtain the new results. We replace the Definition 3.3.1 by

Definition 3.12.1 (Time discretization).

Let $S = [0, T] \subset \mathbb{R}_+$ be a finite time interval. A time discretization of S is defined as a strictly increasing sequence of real numbers $(t_n)_{n=1}^N$ with $t_0 = 0$ and $t_N = T$. The time step is defined by

$$t_\delta^{(n)} = t_n - t_{n-1} < \infty \quad \text{for } n = 1, \dots, N$$

and the largest possible time step is denoted by $\bar{t}_{\delta N} = \sup_{n=1, \dots, N} t_\delta^{(n)}$.

A discretization of the whole domain $Q = S \times \Omega$ is defined by the tuple $\mathcal{D} = (\mathcal{M}, (t_n)_{n=1}^N)$. The size of the discretization is denoted by $\text{size}(\mathcal{D}) = \max\{\text{size}(\mathcal{M}), \bar{t}_{\delta N}\}$. In the convergence proof we consider $\text{size}(\mathcal{D}) \rightarrow 0$.

Additionally in this section, we only consider for simplicity the averaging

$$Y_\nu^\sigma = \frac{\bar{u}_{\nu K} D_{\nu K}(a_{\nu K}) + \bar{u}_{\nu L} D_{\nu L}(a_{\nu L})}{2}, \quad \sigma = K|L.$$

3.12.2 Prolongated quantities

Corresponding to Definition 3.2.2, we define a piecewise constant in time and piecewise constant in space interpolation of $\mathbf{a}_\nu(t_n)$ denoted by $a_{\nu,h}(t, x)$ fulfilling $a_{\nu,h}(t, x) = a_{\nu,h}(t_n, x) = a_{\nu K}(t_n)$ for $t \in (t_{n-1}, t_n]$ and $x \in K$, $K \in \mathcal{V}$.

By $a_{\nu,l}(t, x)$ we denote the piecewise constant in time and piecewise linear in space interpolation of \mathbf{a}_ν fulfilling $a_{\nu,l}(t, x_K) = a_{\nu K}(t_n)$ for $t \in (t_{n-1}, t_n]$ and $K \in \mathcal{V}$. Writing a_l we mean $a_l = (a_{\nu,l})_{\nu=1}^m$ and by a_h we mean $a_h = (a_{\nu,h})_{\nu=1}^m$, see also Figure 3.3 for an illustration. The same

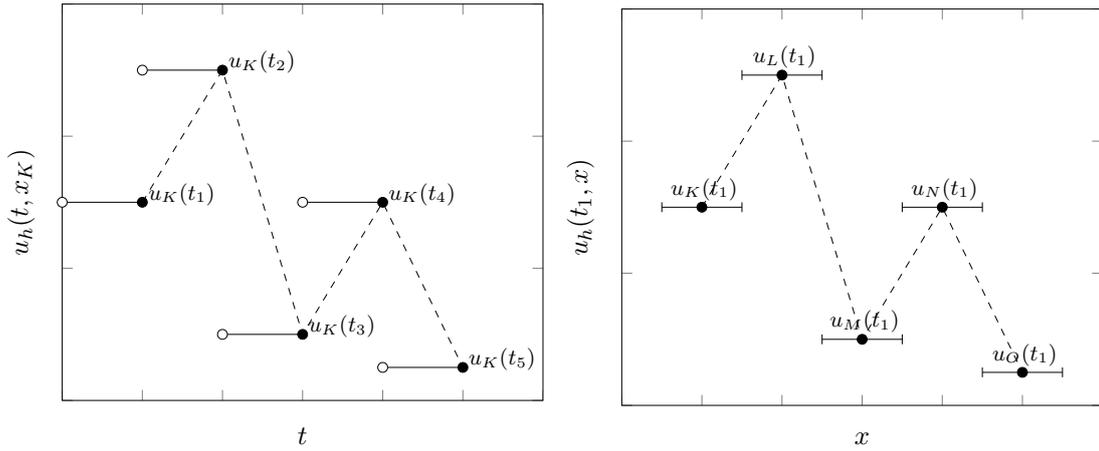


Figure 3.3: Prolongated concentration of \mathbf{u} at note $x_K \in \mathcal{P}$, $K \in \mathcal{V}$ and different times: $u_h(t, x_K)$ piecewise lines, $(K_{\mathcal{D}} u_h)(t, x_K)$ dashed line

notation holds for the chemical potentials \mathbf{v}_ν . Further we introduce an operator $K_{\mathcal{D}}$ mapping the space-time-discrete concentrations into the space $C(S, X^*)$ by

$$(K_{\mathcal{D}} u_h)(t) := \frac{1}{t_\delta^{(n)}} ((t - t_{n-1})u_h(t_n) + (t_n - t)u_h(t_{n-1})) \quad \text{for } t \in (t_{n-1}, t_n]. \quad (3.12.66)$$

We remark that the concentrations u_h are piecewise constant on the control volumes $K \in \mathcal{V}$. Obviously, it holds $(K_{\mathcal{D}} u_h)' = \frac{1}{t_\delta^{(n)}} (u_h(t_n) - u_h(t_{n-1}))$ for all $t \in (t_{n-1}, t_n]$. With \mathbf{e}^{v_l} we denote the vector $(\mathbf{e}^{v_\nu, l})_{\nu=1}^m$.

In the following we work with the half-diamonds

$$\begin{aligned} T_{\sigma K} &= \{tx_K + (1-t)y : t \in (0, 1), y \in \sigma\}, \\ T_{\sigma L} &= \{tx_L + (1-t)y : t \in (0, 1), y \in \sigma\}, \end{aligned}$$

where $\sigma = K|L \in \mathcal{E}_{int}$ denotes the Voronoi surface of two neighboring Voronoi cells K, L with corresponding nodes $x_K, x_L \in \mathcal{P}$, see Figure 3.4. The two half-diamonds $T_{\sigma K}$ and $T_{\sigma L}$ form the

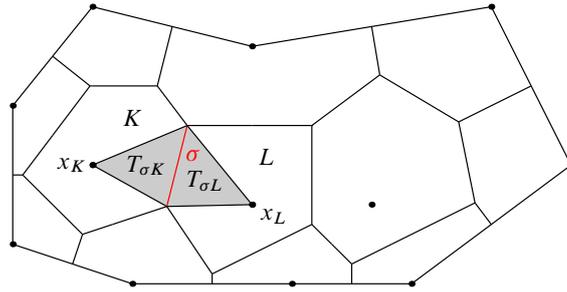


Figure 3.4: Kite $D_\sigma = T_{\sigma K} \cup T_{\sigma L}$ of the edge $\sigma = K|L$

so-called kite $D_\sigma = T_{\sigma K} \cup T_{\sigma L}$ of the edge $\sigma = K|L \in \mathcal{E}_{int}$, the union of all these kites covers the domain Ω . Note that $|T_{\sigma K}| = |T_{\sigma L}| = \frac{1}{2}|D_\sigma|$ holds. The piecewise constant in space interpolation of the reaction rate constant $k_{(\alpha,\beta)K}$ is introduced by $k_{(\alpha,\beta),h}(x, a_h(x)) = k_{(\alpha,\beta)K}(\mathbf{a}_K)$ of the reaction term $R_{\nu K}$ by $R_{\nu,h}(x, a_h(x)) = R_{\nu,K}(\mathbf{a}_K)$, and of the diffusion term $D_{\nu K}$ by $D_{\nu,h}(x, a_h(x)) = D_{\nu,K}(\mathbf{a}_K)$ for all $x \in K, K \in \mathcal{V}$, see (3.2.1). Moreover we identify $\widehat{F}(u)$ with $F_h(u_h)$.

We introduce the weak gradient reconstruction operator $\nabla_w : X_{\mathcal{V}}(\mathcal{M}) \rightarrow L^2(\Omega)^2$ by

$$\nabla_w a_l(\mathbf{x}) := 2 \frac{a_L - a_K}{d_\sigma} \mathbf{n}_\sigma \quad \text{for } \mathbf{x} \in D_\sigma, \sigma = K|L. \quad (3.12.67)$$

The vector \mathbf{n}_σ denotes a unit normal vector of the face $\sigma = K|L \in \mathcal{E}_{int}$, the direction is arbitrary, but fixed.

3.12.3 A priori estimates

In a first step we provide uniform bounds for the solutions to the discrete problem (3.3. P_D) for different discretization levels.

In the following we frequently use for $\nu = 1, \dots, m$

$$0 < \inf_{\mathcal{D}} \operatorname{ess\,inf}_{x \in \Omega} a_{\nu,h}(t), \quad \sup_{\mathcal{D}} \operatorname{ess\,sup}_{x \in \Omega} a_{\nu,h}(t) < \infty \quad \forall t \in S, \quad (3.12.68)$$

which was proven in Theorem 3.8.3 and Theorem 3.10.2.

During the proof of the a priori estimates, we use a result of [32, Lemma 3.4], which holds true under the following assumption.

Assumption A6. *There exists a constant $P \in \mathbb{N}$ such that for all discretizations $\mathcal{D} = (\mathcal{M}, (t_n)_{n=1}^N)$ of $\Omega \times S$ the property $\sup_{\mathcal{M}=(\mathcal{P},\mathcal{V},\mathcal{E})} \max_{K \in \mathcal{V}} \text{card } \mathcal{E}_K \leq P$ holds. Moreover there exist some constant $\xi > 0$ such that $d_\sigma \geq \frac{1}{2}\xi \text{diam}(K)$ for all $K \in \mathcal{V}$ and $\sigma \in \mathcal{E}_K$.*

Lemma 3.12.2 (A priori estimates).

We assume A1, A2, A3, and A4. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n=1}^N)$ be a sequence of discretizations fulfilling Assumptions A5 and A6. For solutions (u_h, v_h) to the discrete Problems (3.3.P_D) the reconstructed quantities a_l and $K_{\mathcal{D}}u_h$ fulfill

$$\sup_{\mathcal{D}} \left\{ \|a_l\|_{L^2(S,X)} + \|K_{\mathcal{D}}u_h\|_{H^1(S,X^*)} + \|K_{\mathcal{D}}u_h\|_{C(S,Y)} \right\} < +\infty.$$

Proof. We test the discrete Problem (3.3.P_D) with $\mathbf{a}(t_n)$. By using the elementary inequality $1/2(x^2 - y^2) \leq (x - y)x$ for $x, y \in \mathbb{R}_+$, the lower and upper bounds on the reference densities, see Assumption A1, we obtain the following estimate

$$\begin{aligned} S_1 &:= \sum_{n=1}^N t_\delta^{(n)} \left\langle \frac{\mathbf{u}(t_n) - \mathbf{u}(t_{n-1})}{t_\delta^{(n)}}, \mathbf{a}(t_n) \right\rangle_{\mathbb{R}^{Mm}} \\ &\geq \sum_{\nu=1}^m \left(\frac{\underline{c}_\nu}{2} \|a_{\nu,h}(t_N)\|_{L^2}^2 - \frac{\bar{c}_\nu}{2} \|a_{\nu,h}(0)\|_{L^2}^2 \right). \end{aligned} \quad (3.12.69)$$

On the other hand by using the definition of \widehat{A} and Z_ν^σ as well as Theorem 3.8.3 we get from (3.3.4) the estimate

$$\begin{aligned} S_1 &= - \sum_{n=1}^N t_\delta^{(n)} \left\langle \widehat{A}\mathbf{v}(t_n), \mathbf{a}(t_n) \right\rangle_{\mathbb{R}^{Mm}} \\ &\leq \sum_{\nu=1}^m \sum_{n=1}^N t_\delta^{(n)} \left\{ -c_1 \|a_{\nu,h}\|_{H^1, \mathcal{M}}^2 + c_1 \|a_{\nu,h}\|_{L^2}^2 + c_2 \right\}. \end{aligned} \quad (3.12.70)$$

Here c_1 is a positive constant depending on the lower bound of the reference densities and the diffusion coefficients. The reaction term was estimated by some constant due to Theorem 3.8.3. The

discrete H^1 -seminorm was completed to the full H^1 -norm. Therefore we deduce from (3.12.69) and (3.12.70) using Theorem 3.8.3 once more with constants $c_3, c_4 > 0$

$$\begin{aligned} \sup_{\mathcal{D}} \|a_l\|_{L^2(S,X)}^2 &\leq \sup_{\mathcal{D}} \left\{ \sum_{n=1}^N t_\delta^{(n)} \sum_{\nu=1}^m \|a_{\nu,h}\|_{H^1,\mathcal{M}}^2 \right\} \\ &\leq \sup_{\mathcal{D}} \sum_{\nu=1}^m \left\{ c_3 \|a_{\nu,h}(0)\|_{L^2}^2 + c_4 \sum_{n=1}^N t_\delta^{(n)} (\|a_{\nu,h}\|_{L^2}^2 + 1) \right\} < \infty. \end{aligned} \quad (3.12.71)$$

From Theorem 3.8.3 we also conclude

$$\sup_{\mathcal{D}} \sup_{t \in S} \|u_h(t)\|_Y = \sup_{\mathcal{D}} \sup_{t \in S} \|K_{\mathcal{D}} u_h(t)\|_Y < \infty$$

and $\sup_{\mathcal{D}} \|K_{\mathcal{D}} u_h\|_{L^2(S,Y)} < \infty$. Since $K_{\mathcal{D}}$ is a continuous interpolant in time we find $K_{\mathcal{D}} u_h \in C(S, Y)$ and $\sup_{\mathcal{D}} \|K_{\mathcal{D}} u_h\|_{C(S,Y)} < \infty$. We prove the boundedness of

$$\sup_{\mathcal{D}} \|(K_{\mathcal{D}} u_h)'\|_{L^2(S,X^*)}^2 = \sup_{\mathcal{D}} \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|(K_{\mathcal{D}} u_h)'\|_{X^*}^2 dt. \quad (3.12.72)$$

First we estimate $\|(K_{\mathcal{D}} u_h)'(t)\|_{X^*}^2$ on the interval (t_{n-1}, t_n) . For this we use an arbitrary test function $w \in X$ and denote by $w_{\nu K} = \frac{1}{|K|} \int_K w_{\nu}(x) dx$ for all $K \in \mathcal{V}$ and $\nu = 1, \dots, m$. By w_h we denote the corresponding piecewise constant function. Then we obtain

$$\begin{aligned} \left| \left\langle \frac{u_h(t_n) - u_h(t_{n-1})}{t_\delta^{(n)}}, w \right\rangle_{X^*} \right| &= \left| \left\langle u_h(t_n) - u_h(t_{n-1}), \frac{w}{t_\delta^{(n)}} \right\rangle_Y \right| \\ &= \left| \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| \frac{u_{\nu K}(t_n) - u_{\nu K}(t_{n-1})}{t_\delta^{(n)}} w_{\nu K} \right| \\ &= \left| \left\langle \frac{\mathbf{u}(t_n) - \mathbf{u}(t_{n-1})}{t_\delta^{(n)}}, \mathbf{w} \right\rangle_{\mathbb{R}^{Mm}} \right| \\ &= \left| \left\langle \widehat{A} \mathbf{v}(t_n), \mathbf{w} \right\rangle_{\mathbb{R}^{Mm}} \right|. \end{aligned}$$

Using (3.3.4), (3.2.2), the boundedness of the reaction terms due to Assumption A1, Theorem 3.8.3 and Hölders inequality we obtain the following estimate

$$\left| \left\langle \widehat{A} \mathbf{v}, \mathbf{w} \right\rangle_{\mathbb{R}^{Mm}} \right| \leq \sum_{\nu=1}^m \left\{ c_5 |a_{\nu,h}|_{H^1,\mathcal{M}} |w_{\nu,h}|_{H^1,\mathcal{M}} + c_6 \|w_{\nu,h}\|_{L^2} \right\}.$$

Using the following arguments

- $|w_{\nu,h}|_{H^1,\mathcal{M}} \leq C \|w_\nu\|_{H^1(\Omega)}$, Assumption A6 is needed, see [32, Lemma 3.4];
- $\|w_{\nu,h}\|_{L^2} \leq C \|w_\nu\|_{H^1} + \|w_\nu - w_{\nu,h}\|_{L^2}$;
- by a discrete Poincaré inequality (see (B.21)) results

$$\begin{aligned} \|w_\nu - w_{\nu,h}\|_{L^2(\Omega)}^2 &= \sum_{K \in \mathcal{V}} \|w_\nu - w_{\nu,K}\|_{L^2(K)}^2 \\ &\leq C \operatorname{diam}(\Omega)^2 \sum_{K \in \mathcal{V}} \|w_\nu\|_{H^1(K)}^2 = C \operatorname{diam}(\Omega)^2 \|w_\nu\|_{H^1(\Omega)}^2, \end{aligned}$$

we get

$$\left| \langle \widehat{A}\mathbf{v}, \mathbf{w} \rangle_{\mathbb{R}^{Mm}} \right| \leq \left(c_7 \sum_{\nu=1}^m |a_{\nu,h}|_{H^1,\mathcal{M}} + c_8 \right) \|w\|_X.$$

Thus we find

$$\|(K_{\mathcal{D}}u_h)'\|_{X^*}^2 \leq \left(c_7 \sum_{\nu=1}^m |a_{\nu,h}|_{H^1,\mathcal{M}} + c_8 \right)^2 \quad \text{a.e. } t \in (t_{n-1}, t_n)$$

and by using $\|a_l\|_{L^2(S,X)} < \infty$ we obtain from (3.12.72) the following estimate

$$\sup_{\mathcal{D}} \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|(K_{\mathcal{D}}u_h)'\|_{X^*}^2 dt \leq \sup_{\mathcal{D}} c_9 \int_S \left(\sum_{\nu=1}^m |a_{\nu,h}|_{H^1,\mathcal{M}}^2 + 1 \right) dt < \infty$$

for all \mathcal{D} . Finally, together with $\sup_{\mathcal{D}} \|K_{\mathcal{D}}u_h\|_{L^2(S,X^*)} \leq c_{10} \sup_{\mathcal{D}} \|K_{\mathcal{D}}u_h\|_{L^2(S,Y)} < \infty$ we obtain $\sup_{\mathcal{D}} \|K_{\mathcal{D}}u_h\|_{H^1(S,X^*)} < \infty$. \square

3.12.4 Weak solution

In this subsection we prove that a subsequence of solutions to the discrete problems (3.3. P_D) converges to a weak solution of the continuous problem (2.2.P). We show the weak convergence of subsequences in different spaces and by using a result of [85] we obtain strong convergence of a subsequence of the chemical activities in $L^2(S, Y)$. The result is proven under an assumption on the discontinuities of the diffusion coefficients, reaction rate coefficients and reference densities.

Assumption A7. Let $f \in L^\infty(\Omega)$ with $f(x) \geq 0$ f.a.a. $x \in \Omega$ and let \mathcal{I} be a finite index set. Furthermore, let $\Omega \subset \mathbb{R}^2$ be a polygonal domain and let $\Omega = \cup_{I \in \mathcal{I}} \Omega_I$ be a finite disjoint union of subdomains such that the discontinuities of f coincide with the subdomain boundaries. Let the over all one-dimensional measure of all internal subdomain boundaries be bounded by θ . There exists some $\gamma \in (0, 1]$ such that $f \in C^{0,\gamma}(\Omega_I) := \{w|_{\Omega_I}, w \in C^{0,\gamma}(\mathbb{R}^2)\}$, $I \in \mathcal{I}$.

In the following we understand under an non-labeled subsequence that a subsequence of $(w_n)_{n \in \mathbb{N}}$ is again denoted by $(w_n)_{n \in \mathbb{N}}$.

Theorem 3.12.3. *We assume A1, A2, A3, and A4. Furthermore we assume that Ω , the reaction rate coefficients $k_{(\alpha, \beta)}(\cdot, \mathbf{y})$, the diffusion coefficients $D_\nu(\cdot, \mathbf{y})$, and the reference densities \bar{u}_ν satisfy Assumption A7 for all $(\alpha, \beta) \in \mathcal{R}$, $\mathbf{y} \in \mathbb{R}^m$ and $\nu = 1, \dots, m$. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n=1}^N)$ be a sequence of discretizations fulfilling Assumptions A5 and A6. Then there exist $\hat{a} \in L^\infty(S, Y) \cap L^2(S, X)$, $\hat{u} \in H^1(S, X^*) \cap L^2(S, Y)$ and non-labeled subsequences a_h, a_l, u_h such that*

$$a_h \rightarrow \hat{a} \text{ in } L^2(S, Y), \quad a_l \rightarrow \hat{a} \text{ in } L^2(S, X), \quad (3.12.73)$$

$$u_h \rightarrow \hat{u} \text{ in } L^2(S, Y), \quad K_{\mathcal{D}} u_h \rightarrow \hat{u} \text{ in } H^1(S, X^*) \quad (3.12.74)$$

as $\text{size}(\mathcal{D}) \rightarrow 0$. And $(\hat{u}, \hat{v}) = (\hat{u}, \ln \hat{a})$ is a weak solution of the continuous problem (2.2.P).

Before we start with the proof we will rephrase the different types of convergence [28, 115]: $a_l \rightarrow \hat{a}$ in $L^2(S, X)$ as $\text{size}(\mathcal{D}) \rightarrow 0$ means

$$\int_0^T \langle a_l - \hat{a}, g \rangle_{X^*} dt \rightarrow 0 \quad \forall g \in L^2(S, X^*).$$

Moreover $a_l \rightarrow \hat{a}$ in $L^q(S, Y)$ as $\text{size}(\mathcal{D}) \rightarrow 0$ iff

$$\int_0^T \|a_l - \hat{a}\|_Y^q dt \rightarrow 0.$$

Finally $(K_{\mathcal{D}} u_h)' \rightarrow \hat{u}'$ in $L^2(S, X^*)$ as $\text{size}(\mathcal{D}) \rightarrow 0$ iff

$$\int_0^T \langle (K_{\mathcal{D}} u_h)' - \hat{u}', g \rangle_X dt \rightarrow 0 \quad \forall g \in L^2(S, X).$$

Have in mind, that the sequence of reference densities is bounded from above and below, i.e. there exist $0 < \underline{c}_{\bar{u}} \leq \bar{c}_{\bar{u}} < \infty$ such that

$$\underline{c}_{\bar{u}} \leq \min_{\nu=1, \dots, m} \text{ess inf}_{x \in \Omega} \bar{u}_{\nu, h}(x) \leq \bar{u}_{\nu, h} \leq \max_{\nu=1, \dots, m} \text{ess sup}_{x \in \Omega} \bar{u}_{\nu, h}(x) \leq \bar{c}_{\bar{u}},$$

see also Assumption A1.

Proof. In order to show the assertions of the theorem we proceed in several steps.

Weak Precompactness: Due to Lemma 3.12.2 and (3.12.68) we deduce the existence of functions

$$\hat{a} \in L^\infty(S, Y) \cap L^2(S, X), \quad \hat{u} \in H^1(S, X^*) \cap L^2(S, Y),$$

such that, at least for non-labeled subsequences,

1. $\bar{u}_h \rightarrow \bar{u}$ in Y due to Lemma C.5,
2. $a_l \rightarrow \hat{a}$ in $L^2(S, X), L^2(S, Y)$,
3. $K_{\mathcal{D}}u_h \rightarrow \hat{u}$ in $H^1(S, X^*), L^2(S, Y)$,
4. $K_{\mathcal{D}}u_h(t) \rightarrow \hat{u}(t)$ in Y , and additionally in $L^p(\Omega)^m$ for $p \in [1, \infty)$ for all $t \in S$

as $\text{size}(\mathcal{D}) \rightarrow 0$. Note that, due to Theorem 3.8.3 we can extract a subsequence such that $K_{\mathcal{D}}u_h(t) \rightarrow \hat{u}(t)$ in $L^p(\Omega)^m$ for $p \in [1, \infty)$ and for all $t \in S$. By the definition of $K_{\mathcal{D}}$ and Lemma C.5 we have $u_h(0) = K_{\mathcal{D}}u_h(0) \rightarrow U$ in Y for $\text{size}(\mathcal{D}) \rightarrow 0$, hence $U = \hat{u}(0)$. Using Definition 3.12.1 we find that

$$\begin{aligned} \|u_h - K_{\mathcal{D}}u_h\|_{L^2(S, X^*)}^2 &= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (t_n - t)^2 \|(K_{\mathcal{D}}u_h)'\|_{X^*}^2 dt \\ &\leq \bar{t}_\delta^2 \|(K_{\mathcal{D}}u_h)'\|_{L^2(S, X^*)}^2 \rightarrow 0 \end{aligned}$$

and therefore for subsequences $K_{\mathcal{D}}u_h(t) - u_h(t) \rightarrow 0$ in X^* and $u_h(t) \rightarrow \hat{u}(t)$ in X^* f.a.a. $t \in S$ as $\text{size}(\mathcal{D}) \rightarrow 0$. Due to Lemma 3.12.2 and Theorem 3.8.3 we get $\sup_{\mathcal{D}} \|a_l(t)\|_Y < \infty$ for all $t \in S$, hence there exists a non-labeled subsequence, such that $a_l(t) \rightarrow \hat{a}(t)$ in Y f.a.a. $t \in S$.

Strong convergence of the activities: Now we use a result of [85]: *Let $(\varphi_j)_{j \in \mathbb{N}}$ be an orthogonal basis of $L^2(\Omega)$. For all $\varepsilon > 0$ there exists a $N_\varepsilon > 0$ such that*

$$\|w\|_{L^2(\Omega)}^2 \leq \sum_{j=1}^{N_\varepsilon} \langle w, \varphi_j \rangle_{L^2(\Omega)}^2 + \varepsilon \|w\|_{H^1(\Omega)}^2 \quad \forall w \in H^1(\Omega). \quad (3.12.75)$$

We apply the result for different discretizations, i.e., $w = a_{\nu,l} - a_{\nu,l'}$, $\nu = 1, \dots, m$, and obtain after integration over S

$$\|a_{\nu,l} - a_{\nu,l'}\|_{L^2(S, L^2(\Omega))}^2 \leq \sum_{j=1}^{N_\varepsilon} \int_0^T \langle a_{\nu,l} - a_{\nu,l'}, \varphi_j \rangle_{L^2(\Omega)}^2 dt + \varepsilon \|a_{\nu,l} - a_{\nu,l'}\|_{L^2(S, H^1(\Omega))}^2.$$

Note the boundedness of $\sup_{\mathcal{D}} \|a_{\nu,l}\|_{L^2(S, H^1(\Omega))}$. Since $a_{\nu,l}$ is bounded in Y for all $t \in S$, by the dominated convergence theorem we conclude that $(a_{\nu,l})_{\mathcal{D}}$ is a Cauchy sequence in $L^2(S, L^2(\Omega))$ and therefore $a_{\nu,l} \rightarrow \hat{a}_\nu$ in $L^2(S, L^2(\Omega))$ for $\nu = 1, \dots, m$. Since $u_h/\bar{u}_h = a_h$ and since $a_h - a_l \rightarrow 0$ in $L^2(S, Y)$ (see (C.29)) we obtain

$$\|u_h/\bar{u}_h - \hat{a}\|_{L^2(S, Y)} \leq \|a_h - a_l\|_{L^2(S, Y)} + \|a_l - \hat{a}\|_{L^2(S, Y)} \rightarrow 0.$$

On the other hand by using $u_h(t) \rightharpoonup \hat{u}(t)$ in X^* f.a.a. $t \in S$ and $\sup_{\mathcal{D}} \|u_h(t)\|_Y < \infty$ for all $t \in S$ we obtain for a non-labeled subsequence $u_h(t) \rightharpoonup \hat{u}(t)$ in Y f.a.a. $t \in S$. Now we apply the strong convergence of $\bar{u}_h \rightarrow \bar{u}$ in Y and the uniform boundedness from above and below of the sequence (\bar{u}_h) to get

$$u_h(t)/\bar{u}_h \rightharpoonup \hat{u}(t)/\bar{u} \text{ in } Y \text{ f.a.a. } t \in S. \quad (3.12.76)$$

Due to $\sup_{\mathcal{D}} \|u_h/\bar{u}_h\|_{L^2(S,Y)} < \infty$ we conclude by the dominated convergence theorem $u_h/\bar{u}_h \rightharpoonup \hat{u}/\bar{u}$ in $L^2(S, Y)$ which gives us $\hat{a} = \hat{u}/\bar{u}$. Since a_h is uniformly bounded from above and below (see (3.12.68)) the bounds also hold for the limit \hat{a} and therefore \hat{a} is positive and $\hat{v} := \ln \hat{a}$. Using the Assumptions A1 and A7 together with Lemma C.6 we conclude with some $c > 0$ that

$$\begin{aligned} \|R_{\nu h}(a_h) - R_{\nu}(\hat{a})\|_{L^2(S,Y)}^2 &\leq c \|R_{\nu h}(a_h) - R_{\nu}(\hat{a})\|_{L^1(S,Y)} \rightarrow 0, \\ \|\bar{u}_{\nu h} D_{\nu h}(a_h) - \bar{u}_{\nu} D_{\nu}(\hat{a})\|_{L^2(S,Y)}^2 &\leq c \|\bar{u}_{\nu h} D_{\nu h}(a_h) - \bar{u}_{\nu} D_{\nu}(\hat{a})\|_{L^1(S,Y)} \rightarrow 0 \end{aligned} \quad (3.12.77)$$

as size $\mathcal{D} \rightarrow 0$. This gives pointwise convergence a.e. for subsequences.

Weak solution: Since $C_0^\infty(\Omega)^m$ is dense in X we use test functions $\varphi \in C_0^\infty(\Omega)^m$ and a pure time function $\chi \in C_0^\infty(S)$. Set $\varphi = (\varphi_{\nu}(x_K))_{K \in \mathcal{V}, \nu=1, \dots, m}$ and $\varphi_{\nu, h}(x) = \varphi_{\nu}(x_K)$ for $x \in K$, $K \in \mathcal{V}$. By $\varphi_{\nu, l}$ we denote the piecewise affine interpolation of φ . Note that $\varphi_{\nu, l} \in L^\infty(\Omega)$. Let be $\bar{\chi}(t_n) = (\chi(t_n) + \chi(t_{n-1}))/2$ for $n = 1, \dots, N$. We note that

$$\begin{aligned} t_\delta^{(n)} \bar{\chi}(t_n) &= \frac{(t_\delta^{(n)})^2}{2t_\delta^{(n)}} (\chi(t_n) + \chi(t_{n-1})) \\ &= \frac{1}{t_\delta^{(n)}} \int_{t_{n-1}}^{t_n} \chi(t_n)(t - t_{n-1}) + \chi(t_{n-1})(t_n - t) dt \\ &= \int_{t_{n-1}}^{t_n} (K_{\mathcal{D}} \chi)(t) dt, \quad n = 1, \dots, N \end{aligned} \quad (3.12.78)$$

and $(K_{\mathcal{D}}\chi) \rightarrow \chi$ in $L^2(S)$ (use the mean-value form of the Taylor series remainder). We use test functions $\varphi \bar{\chi}(t_n)$, $t \in (t_{n-1}, t_n]$, $n = 1, \dots, N$ for (3.3.4) and sum over $n = 1, \dots, N$ to obtain $S_1 + S_2 = S_3$ with

$$\begin{aligned} S_1 &:= \sum_{n=1}^N t_{\delta}^{(n)} \bar{\chi}(t_n) \left\langle \frac{\mathbf{u}(t_n) - \mathbf{u}(t_{n-1})}{t_{\delta}^{(n)}}, \boldsymbol{\varphi} \right\rangle_{\mathbb{R}^{Mm}}, \\ S_2 &:= \sum_{n=1}^N t_{\delta}^{(n)} \bar{\chi}(t_n) \sum_{\nu=1}^m \sum_{\sigma=T|L \in \mathcal{E}_{int}} T_{\sigma} Y_{\nu}^{\sigma} (a_{\nu L} - a_{\nu K}) (\varphi_{\nu L} - \varphi_{\nu K}), \\ S_3 &:= \sum_{n=1}^N t_{\delta}^{(n)} \bar{\chi}(t_n) \sum_{\nu=1}^m \sum_{K \in \mathcal{V}} |K| R_{\nu K}(\mathbf{a}_K) \varphi_{\nu K} = \sum_{\nu=1}^m \int_Q (K_{\mathcal{D}}\chi)(t) R_{\nu, h}(\mathbf{a}_h) \varphi_{\nu, h} dx dt. \end{aligned}$$

We remark that the interpretation of the discrete sums as an integral over $Q = S \times \Omega$ like in the last two lines is crucial for the convergence proof. In the following we will do this several times. We define

$$\begin{aligned} \widehat{S}_1 &:= \int_S \langle \widehat{u}', \boldsymbol{\varphi} \rangle_X \chi dt, \\ \widehat{S}_2 &:= \sum_{\nu=1}^m \int_S \chi \int_{\Omega} D_{\nu}(\widehat{a}) \bar{u}_{\nu} \nabla \widehat{a}_{\nu} \cdot \nabla \varphi_{\nu} dt dx, \\ \widehat{S}_3 &:= \sum_{\nu=1}^m \int_S \chi \int_{\Omega} R_{\nu}(\widehat{a}) \varphi_{\nu} dx dt. \end{aligned}$$

Time derivative: Using (3.12.78) the term S_1 can be interpreted as

$$S_1 = \sum_{\nu=1}^m \int_Q \varphi_{\nu, h} ((K_{\mathcal{D}}u_{\nu, h})(t))' (K_{\mathcal{D}}\chi)(t) dx dt.$$

By partial time integration (and $K_{\mathcal{D}}\chi(0) = \chi(0) = \chi(t_N) = K_{\mathcal{D}}\chi(t_N) = 0$) we obtain

$$S_1 = - \sum_{\nu=1}^m \int_{\Omega} \int_S \varphi_{\nu, h} (K_{\mathcal{D}}u_{\nu, h})(t) ((K_{\mathcal{D}}\chi)(t))' dt dx.$$

Now we use $\varphi_h \rightarrow \varphi$ in Y , $(K_{\mathcal{D}}\chi)' \rightarrow \chi'$ in $L^2(S)$, $K_{\mathcal{D}}u_h \rightarrow \widehat{u}$ in $L^2(S, Y)$ and $\widehat{u} \in H^1(S, X^*)$ to find

$$\lim_{\text{size } \mathcal{D} \rightarrow 0} S_1 = - \int_S \langle \widehat{u}, \boldsymbol{\varphi} \rangle_Y \chi' dt = \widehat{S}_1.$$

Reactions: Using (3.12.78) the reaction term S_3 can be written in the form $S_3 = S_{31} + S_{32}$,

$$S_{31} := \sum_{\nu=1}^m \int_S (K_{\mathcal{D}}\chi)(t) \int_{\Omega} (R_{\nu,h}(a_h) - R_{\nu}(\hat{a})) \varphi_{\nu,h} dx dt,$$

$$S_{32} := \sum_{\nu=1}^m \int_S (K_{\mathcal{D}}\chi)(t) \int_{\Omega} R_{\nu}(\hat{a}) \varphi_{\nu,h} dx dt.$$

Since $\varphi_h \rightarrow \varphi$ in Y we find $S_{32} \rightarrow \hat{S}_3$ and by the Cauchy-Schwarz inequality we get

$$S_{31} \leq \sum_{\nu=1}^m \|(K_{\mathcal{D}}\chi)\varphi_{\nu,h}\|_{L^2(Q)} \|R_{\nu,h}(a_h) - R_{\nu}(\hat{a})\|_{L^2(Q)}.$$

For the first term we find $\lim_{\text{size } \mathcal{D} \rightarrow 0} \|(K_{\mathcal{D}}\chi)\varphi_{\nu,h}\|_{L^2(Q)} = \|\chi\varphi_{\nu}\|_{L^2(Q)}$ and due to (3.12.77) the second term tends to zero, and we find $\lim_{\text{size } \mathcal{D} \rightarrow 0} S_{31} = 0$, hence $\lim_{\text{size } \mathcal{D} \rightarrow 0} S_3 = \hat{S}_3$.

Diffusion: The average Y_{ν}^{σ} , over a diamond $D_{\sigma} = T_{\sigma K} \cup T_{\sigma L}$ related to the edge σ , can be expressed by

$$|D_{\sigma}|Y_{\nu}^{\sigma} = |T_{\sigma K}|\bar{u}_{\nu K}D_{\nu K}(\mathbf{a}_K) + |T_{\sigma L}|\bar{u}_{\nu L}D_{\nu L}(\mathbf{a}_L) = \int_{D_{\sigma}} \bar{u}_{\nu,h}D_{\nu,h}(a_h) dx$$

since $|T_{\sigma K}| = |T_{\sigma L}| = \frac{1}{2}|D_{\sigma}|$ and $\bar{u}_{\nu K}D_{\nu K}(\mathbf{a}_K)$ constant on K . The weak gradient formulation and (3.12.78) yield

$$\begin{aligned} S_2 &= \sum_{n=1}^N t_{\delta}^{(n)} \bar{\chi}(t_n) \sum_{\nu=1}^m \sum_{\sigma=T|L \in \mathcal{E}_{int}} |D_{\sigma}|Y_{\nu}^{\sigma} 2 \frac{(a_{\nu L} - a_{\nu K})}{d_{\sigma}} \frac{(\varphi(x_L) - \varphi(x_K))}{d_{\sigma}} \\ &= \sum_{\nu=1}^m \int_S (K_{\mathcal{D}}\chi)(t) \int_{\Omega} \bar{u}_{\nu,h}D_{\nu,h}(a_{\nu,h}) \nabla_w a_{\nu,l} \cdot \nabla \varphi_{\nu,l} dx \\ &= S_{21} + S_{22} \end{aligned}$$

since the tangential component of the strong gradient is orthogonal to the normal component of the weak gradient. Here

$$S_{21} := \sum_{\nu=1}^m \int_S (K_{\mathcal{D}}\chi)(t) \int_{\Omega} (D_{\nu,h}(a_h)\bar{u}_{\nu,h} - D_{\nu}(\hat{a})\bar{u}_{\nu}) \nabla_w a_{\nu,l} \cdot \nabla \varphi_{\nu,l} dx dt,$$

$$S_{22} := \sum_{\nu=1}^m \int_S (K_{\mathcal{D}}\chi)(t) \int_{\Omega} D_{\nu}(\hat{a})\bar{u}_{\nu} \nabla_w a_{\nu,l} \cdot \nabla \varphi_{\nu,l} dx dt.$$

Applying the Cauchy-Schwarz inequality and the boundedness of $K_{\mathcal{D}}\chi\nabla\varphi_{\nu,l}$ and of the reference densities we find with a constant $c > 0$

$$S_{21} \leq \sum_{\nu=1}^m c \|\nabla_w a_{\nu,l}\|_{L^2(S, L^2(\Omega)^2)} \|\bar{u}_{\nu,h} D_{\nu,h}(a_h) - \bar{u}_{\nu} D_{\nu}(\hat{a})\|_{L^2(S, L^2)}.$$

Using (3.12.77), the estimate $\|\nabla_w a_{\nu,l}\|_{L^2(S, L^2(\Omega)^2)} \leq 2\|\nabla a_{\nu,l}\|_{L^2(S, L^2(\Omega)^2)}$, and Lemma 3.12.2 we obtain $S_{21} \rightarrow 0$ as $\text{size}(\mathcal{D}) \rightarrow 0$. For S_{22} we use weak-strong convergence with $a_l \rightarrow \hat{a}$ in $L^2(S, X)$ and $(K_{\mathcal{D}}\chi)\varphi_l \rightarrow \chi\varphi$ in $L^2(S, X)$ to find $\lim_{\text{size}(\mathcal{D}) \rightarrow 0} S_{22} = \hat{S}_2$ hence $\lim_{\text{size}(\mathcal{D}) \rightarrow 0} S_2 = \hat{S}_2$. Since $S_1 + S_2 = S_3$ and $\lim_{\text{size}(\mathcal{D}) \rightarrow 0} S_i = \hat{S}_i$ for $i = 1, 2, 3$ we have shown that $(\hat{u}, \hat{v}) = (\hat{u}, \ln \hat{a})$ is a weak solution of the continuous problem (2.2.P) in the sense of (2.2.3). \square

Corollary 3.12.4. *We assume A1, A2, A3, and A4. Furthermore we assume that Ω , the reaction rate coefficients $k_{(\alpha,\beta)}(\cdot, \mathbf{y})$, the diffusion coefficients $D_{\nu}(\cdot, \mathbf{y})$, and the reference densities \bar{u}_{ν} satisfy Assumption A7 for all $(\alpha, \beta) \in \mathcal{R}$, $\mathbf{y} \in \mathbb{R}^m$ and $\nu = 1, \dots, m$. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n=1}^N)$ be a sequence of discretizations fulfilling Assumptions A5 and A6. Moreover, let (\hat{u}, \hat{v}) be a weak solution of the continuous problem (2.2.P). Then there exists a subsequence of solutions (u_h, v_h) to (3.3.P_D) with the convergence properties stated in Theorem 3.12.3 such that*

$$\|F_h(u_h) - F(\hat{u})\|_{L^2(S)} \rightarrow 0 \text{ as } \text{size}(\mathcal{D}) \rightarrow 0,$$

where $F_h(u_h)$ denotes the corresponding piecewise constant reconstruction of the discrete free energy $\hat{F}(\mathbf{u})$, i.e.,

$$F_h(u_h) := \sum_{\nu=1}^m \int_{\Omega} \left(u_{\nu,h} \ln \frac{u_{\nu,h}}{\bar{u}_{\nu,h}} - u_{\nu,h} + \bar{u}_{\nu,h} \right) dx = \hat{F}(\mathbf{u}).$$

Moreover, for the relative free energy $\Psi_h(u_h) := F_h(u_h) - F(u_h^*)$ holds

$$\|\Psi_h(u_h) - \Psi(\hat{u})\|_{L^2(S)} \rightarrow 0 \text{ as } \text{size}(\mathcal{D}) \rightarrow 0. \quad (3.12.79)$$

Proof. Let c be a positive constant with varying meaning. Thanks to Theorem 3.8.3 we get

$$\begin{aligned} \|\ln(u_h/\bar{u}_h) - \ln(\hat{u}/\bar{u})\|_{L^2(S,Y)} &= \|\ln a_h - \ln \hat{a}\|_{L^2(S,Y)} \\ &\leq \frac{\|a_h - \hat{a}\|_{L^2(S,Y)}}{\min_{\nu=1,\dots,m} \text{ess inf}_{x \in \Omega} \hat{a}_{\nu}(x)} \rightarrow 0 \end{aligned}$$

and together with Lemma C.5 and $u_h \rightarrow \hat{u}$ in $L^2(S, Y)$ we obtain the following estimate

$$\begin{aligned} S_1 &= \|F_h(u_h) - F(\hat{u})\|_{L^2(S)}^2 \\ &\leq c \|\ln(u_h/\bar{u}_h) - \ln(\hat{u}/\bar{u})\|_{L^2(S,Y)}^2 + c \|u_h - \hat{u}\|_{L^2(S,Y)}^2 + c T \|\bar{u}_h - \bar{u}\|_Y^2 \rightarrow 0 \end{aligned}$$

as $\text{size}(\mathcal{D}) \rightarrow 0$. Since $u_h^* = \frac{\bar{u}_h}{\bar{u}} u^*$ and $a_h^* = a^*$ holds (see [56, Corollary 3.1]) and by Lemma C.5, we find

$$\|F(u_h^*) - F(u^*)\|_{L^2(S)}^2 \leq cT \sum_{\nu=1}^m \|a_\nu^*(\ln a_\nu^* - 1) + 1\|_{L^\infty(\Omega)} \|\bar{u}_{\nu,h} - \bar{u}_\nu\|_{L^2(\Omega)}^2 \rightarrow 0,$$

from which the second assertion follows. \square

Remark 3.12.5. *In the case where the diffusion coefficients only depend on the space variable and not on the state variable we have uniqueness of the weak solution (\hat{u}, \hat{v}) of the continuous problem (2.2.P) as proven in [63, Theorem 5.6]. Assume by contradiction that there exist two converging subsequences, then one can repeat the proof of Theorem 3.11.1 and Section 3.12 and use [63, Theorem 5.6] to conclude that there exists only one limit point. Therefore not only a subsequence converges, but so does the whole sequence.*

3.13 Comments on 3D

This section discusses the generalization of the previous obtained results to three space dimensions. All results of the Sections 3.4- 3.11 can be extended to three space dimension. However, the reaction order in Section 3.8, 3.9, and 3.10 is more restrictive than in two space dimension.

In the proof of the uniform exponential decay of the discrete free energy (see Theorem 3.9.3) the restriction of the reaction order is required by some steps in the proof due to the use of a discrete Sobolev-Poincaré inequality. Using the new results concerning the discrete Sobolev-Poincaré inequality in [10], see (B.22), instead of [60, Theorem 2.2] the estimate of (3.9.41), (3.9.42), and (3.9.52) holds true for $p \in [1, 6]$. Hence the estimate of (3.9.46) and (3.9.56) is valid up to a reaction order less or equal to 3. In two space dimension this result remains true for reaction of arbitrary finite order.

The uniform upper boundedness of the discrete solutions is obtained by Moser iteration, which needs an initial bound in some L^p , $p \geq 1$ space. In order to get an L^2 bound (see (3.8.28)), we estimate the growth of the reactions terms by a Gagliardo-Nirenberg inequality stated in Corollary B.6. The resulting H^1 -norm is compensated by the H^1 -norm of the diffusion term. From (B.20) follows that $p \leq 2 \frac{d+1}{d} = \frac{8}{3}$ in (B.23) for $d = 3$. Hence on the results of (B.23) our technique to get uniform upper works also in three space dimensions, provided that the order of the reaction source terms (see (2.1.1)) is less or equal to $\frac{5}{3}$. This is exactly the same restriction as in the continuous situation, see Chapter 2 and [46].

In the proof of convergence, we use the feature that in two space dimensions, the associated discrete H^1 -seminorm is equivalent to the H^1 -seminorm of the linear finite element reconstruction method. Numerical experiments suggest that they are also equivalent in three space dimensions

3 The discretized problem

under some conditions on the mesh. If this result would be proved rigorously, the proofs in Section 3.12 can be extended to three space dimensions.

4

Implementation of the scheme

Abstract In Chapter 3 a scheme is presented which is unconditionally stable and respects the physical properties of the continuous problem. Now we provide ideas how to implement the method in case of MMH kinetics, and to solve the discrete nonlinear system of equations by Newton's method. Our aim is to preserve the qualitative properties of the scheme during the calculations, especially for long time intervals. During the evolution of the system, the invariance property of the solution (3.4.1) is violated for finite precision arithmetic. This leads to a nonphysical behavior of the solutions, see Remark 3.7.5 and Section 5.2. Therefore we present a possible approach to preserve the invariance property of the solution (3.4.1) up to roundoff errors. The focus of Subsection 4.2 is the solution of the resulting linear systems, which must be solved in every Newton step. The implementation was done in Fortran and is parallelized by OpenMP [109]. Some core routines (BDF time stepping and vector routines) are borrowed from [49].

In this chapter, the reaction rates, diffusion coefficients, and reference densities are assumed to be constant per triangle and the interfaces of different materials are aligned to the triangle edges, see also Assumption A7. Let $|V_{KT}|$ be the Voronoi volume of the vertex $x_K \in T$ with respect to a simplex $T = (x_K, x_L, x_M) \in \mathcal{P}^3$ of the Delaunay mesh, see Remark 3.2.1. With $\bar{u}_{\nu T}$, $k_{(\alpha,\beta)T}$ we denote the constant value of the reference densities and reaction rates on a triangle $T \in \mathcal{T}$ of the Delaunay mesh $(\mathcal{P}, \mathcal{T})$.

4.1 Solution of the nonlinear equations

Due to the implicit time discretization of the discrete problem (3.3. P_D), it is necessary to solve a nonlinear system of equations in every time step. This means, in every time step we search for a tuple of vectors $(\tilde{\mathbf{u}}, \tilde{\mathbf{v}})$ which is a solution of the discrete Problem (3.3. P_D). Following the idea of [43], in the developed code the assembly of the discrete Problem (3.3. P_D) is done element wise over the triangles of the Delaunay triangulation (see Remark 3.2.1).

For that reason, we introduce the vector of reference densities multiplied by the volume K in the form

$$\bar{\mathbf{u}}_\nu = \left(\sum_{T \in \mathcal{N}_T(K)} |V_{KT}| \bar{u}_{\nu T} \right)_{K \in \mathcal{V}}, \quad \bar{\mathbf{u}} = (\bar{\mathbf{u}}_\nu)_{\nu=1}^m.$$

Since the reference densities are fixed over the evolution, the vector can be precomputed.

Considering the discrete Problem (3.3. P_D), we search for a vector of chemical potentials \mathbf{w} such that

$$\hat{\mathcal{B}}(\mathbf{w}) = \frac{\bar{\mathbf{u}} e^{\mathbf{w}} - \mathbf{u}(t_{n-1})}{t_\delta^{(n)}} + \hat{A}\mathbf{w} = 0, \quad \hat{E}\mathbf{w} \in \hat{\mathcal{U}} - \mathbf{U} \quad (4.1.1)$$

holds. Note that the operator $\hat{\mathcal{B}}$ is different to the operator $\hat{\mathcal{B}}$ introduced in Chapter 3. Newton's method uses a Taylor expansion of (4.1.1) around a predictor $\mathbf{v}^{(0)} = \ln \mathbf{a}^{(0)} = \ln \frac{\mathbf{u}^{(0)}}{\bar{\mathbf{u}}}$ for the solution at time t_n . The resulting linear system

$$J_{\hat{\mathcal{B}}}(\mathbf{v}^{(0)})\delta_v = -\hat{\mathcal{B}}(\mathbf{v}^{(0)}), \quad (4.1.2)$$

consists of the Jacobian matrix $J_{\hat{\mathcal{B}}}$, the function $\hat{\mathcal{B}}$, and the Newton update δ_v .

The solution δ_v of the linear system can be used to obtain a better approximation $\mathbf{v}^{(1)} = \mathbf{v}^{(0)} + \delta_v$ of the solution of (4.1.1). The iteration process is repeated, expected the convergence of the method, until a sufficiently accurate value is reached.

The appropriate choice of the initial guess $\mathbf{v}^{(0)}$ for the solution at time t_n has a big impact on convergence property of the method. If $\mathbf{v}^{(0)}$ is far away from the solution, the Newton method is out of the quadratic convergence region and needs a lot of steps to reach this region. As predictor $\mathbf{v}^{(0)}$, we use an exponential extrapolation (in time) of the last two solutions in activities. In other words, a linear extrapolation of the last two solutions in chemical potentials

$$\mathbf{v}^{(0)} = \mathbf{v}(t_{n-1}) + \frac{t_n - t_{n-2}}{t_{n-1} - t_{n-2}} (\mathbf{v}(t_{n-1}) - \mathbf{v}(t_{n-2})).$$

Since $\mathbf{a}^{(0)} = e^{\mathbf{v}^{(0)}}$ holds, it follows that the predictor preserves the positivity of the chemical activities. The vector $\mathbf{v}^{(0)}$ does not fulfill the invariance property of the solution, see Lemma 3.4.1. Linear extrapolation of the chemical activities in time preserves the mass invariance property as long as the two previously accepted solutions fulfill this property (exploit linearity of the extrapolation and of the affine space $\hat{\mathcal{U}} + \mathbf{U}$). But the linear extrapolation can violate the positivity of the chemical activities, hence the domain of the operator \hat{A} is violated and the behavior of the free energy is undefined. In the Example 5.4, we return to this topic again.

Remark 4.1.1. We have carried out some experiments with an explicit time integrator as predictor, i.e., an explicit Euler method

$$\left. \begin{aligned} \frac{\mathbf{u}(t_i) - \mathbf{u}(t_{i-1})}{t_\delta} + \widehat{A}\mathbf{v}(t_{i-1}) &= \mathbf{0}, & \mathbf{u}(t_i) &= \widehat{E}\mathbf{v}(t_i), & i \geq 1, \\ \mathbf{u}(0) &= \mathbf{u}(t_{n-1}). \end{aligned} \right\}$$

or a second order Adams–Bashforth method (see [24, 70])

$$\left. \begin{aligned} \frac{\mathbf{u}(t_i) - \mathbf{u}(t_{i-1})}{t_\delta} + \frac{3}{2}\widehat{A}\mathbf{v}(t_{i-1}) - \frac{1}{2}\widehat{A}\mathbf{v}(t_{i-2}) &= \mathbf{0}, & \mathbf{u}(t_i) &= \widehat{E}\mathbf{v}(t_i), & i \geq 2, \\ \mathbf{u}(0) &= \mathbf{u}(t_{n-1}). \end{aligned} \right\}$$

and a reduced time step t_δ , but still violating the CFL condition $c \frac{t_\delta}{(\text{diam } K)^2} \leq 1$ with some positive constant c depending on the data of the problem, see [116]. The instability results in an inaccurate forecast of the solution at time t_n . It does not solve the problem of positivity and preserving of invariants.

In the following we describe the different parts of the Jacobian $J_{\widehat{B}} = J_T + J_R + J_D$, consisting of the Jacobian of the time derivative J_T , the Jacobian of the flux term J_D and the Jacobian of the reaction term J_R .

Time derivative The function of the time derivative can be expressed as

$$\frac{\mathbf{u}^{(0)} - \mathbf{u}^{(K)}(t_{n-1})}{t_\delta^{(n)}}.$$

Differentiating the term with respect to the chemical potentials $\mathbf{v}^{(0)}$ we obtain

$$J_T := \frac{1}{t_\delta^{(n)}} \left[\mathbf{u}^{(0)} \right]. \quad (4.1.3)$$

Both the function and the Jacobian can be computed in a loop *over all vertices* in the mesh.

Reaction term Let

$$\mathbf{k}_{(\alpha, \beta)} = \left(\sum_{T \in \mathcal{N}(K)} |V_{KT}| k_{(\alpha, \beta)T} \right)_{K \in \mathcal{V}}$$

be the vector containing the reaction rates multiplied by the volume of the control volume. The reaction term of all species can be written in the form

$$\mathbf{R} = \left(\sum_{(\alpha, \beta) \in \mathcal{R}} (\alpha_\nu - \beta_\nu) \mathbf{k}_{(\alpha, \beta)} \left(e^{\alpha \cdot \mathbf{v}_K^{(0)}} - e^{\beta \cdot \mathbf{v}_K^{(0)}} \right)_{K \in \mathcal{V}} \right)_{\nu=1}^m. \quad (4.1.4)$$

Since the reaction rates are independent of the state, the vector can be precomputed at the beginning. Only the last vector $\left(e^{\alpha \cdot \mathbf{v}_K^{(0)}} - e^{\beta \cdot \mathbf{v}_K^{(0)}} \right)_{K \in \mathcal{V}}$ changes from Newton step to Newton step.

Linearizing R around $\mathbf{v}^{(0)}$, we obtain $J_R := J_{RS} J_{RK} J_{RV}$ with

$$J_{RS} = \mathcal{S} \otimes Id_M, \quad J_{RK} = \bigoplus_{(\alpha, \beta) \in \mathcal{R}} [\mathbf{k}_{(\alpha, \beta)}],$$

$$J_{RV} = \left(\left(\text{diag}_{K \in \mathcal{V}} \left(\alpha_\nu e^{\alpha \cdot \mathbf{v}_K^{(0)}} - \beta_\nu e^{\beta \cdot \mathbf{v}_K^{(0)}} \right) \right)_{\nu=1}^m \right)_{(\alpha, \beta) \in \mathcal{R}}.$$

Here $\mathcal{S} \otimes Id_M$ denotes the Kronecker product of the two matrices \mathcal{S} and Id_M and $\bigoplus_{(\alpha, \beta) \in \mathcal{R}} [\mathbf{k}_{(\alpha, \beta)}]$ the direct sum of all diagonal matrices $[\mathbf{k}_{(\alpha, \beta)}]$, $(\alpha, \beta) \in \mathcal{R}$.

The matrix J_{RK} can be precomputed and the matrix J_{RS} is a permutation matrix, only the matrix J_{RV} has to be recomputed in every Newton step, but this is only a loop *over all vertices* in the mesh. We remark that the vector of reactions \mathbf{R} can be written in a symmetric form, see [90], however the Jacobian J_R is non symmetric. Only at thermodynamic equilibrium the matrix J_R is symmetric, since then $e^{\alpha \cdot \mathbf{v}_K^{(0)}} = e^{\beta \cdot \mathbf{v}_K^{(0)}}$ holds in every control volume $K \in \mathcal{V}$ and for all $(\alpha, \beta) \in \mathcal{R}$, see Section 3.7.

Flux term In order to assemble the flux term over a triangle T we introduce the following matrix

$$G_2 = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}. \quad \text{Then } \mathbf{1}G_2 = \mathbf{0}, \quad (G_2^T G_2)_{ii} > 0, \quad (G_2^T G_2)_{i>j} < 0. \quad (4.1.5)$$

The diagonal matrix $[\gamma]$ contains the geometric weights per simplex (the transmissibility T_σ of the edge σ). The element matrix G_2 maps nodal values to edge differences on a triangle. The local (per triangle) analogon of the flux term (3.3.3) reads as follows

$$\sum_{T \in \mathcal{T}} G_2^T [Y_{\nu T}] [\gamma] G_2 \mathbf{a}_{\nu T}^{(0)} \quad (4.1.6)$$

where $\mathbf{a}_{\nu T}^{(0)} = e^{\mathbf{v}_{\nu T}^{(0)}}$ is the vector containing the activities of species ν in the nodes of triangle T . Here the diagonal matrix $[Y_{\nu T}]$ is created from the vector $Y_{\nu T}$ containing the values of the mean value Y_ν^σ of the involved edges of the triangle T . If the diffusion coefficient and the reference density are assumed to be piecewise constant on the triangle T , then the matrix $[Y_{\nu T}]$ is a scaled identity matrix.

The Jacobi matrix is obtained by differentiating the term (4.1.6) with respect to $\mathbf{v}^{(0)}$, which results in

$$\sum_{T \in \mathcal{T}} G_2^T [Y_{\nu T}] [\gamma] G_2 [\mathbf{a}_{\nu T}^{(0)}]. \quad (4.1.7)$$

This form has the advantage, that all computations can be done on a triangle and then the local quantities can be distributed into the global Jacobi matrix J_D of the system and the global function vector. Every species creates a diagonal block of the size $M \times M$ in the matrix J_D . The sparsity pattern of all these diagonal blocks is fixed for a given mesh. Therefore, the displacement vector mapping the local element matrices in the global matrix J_D can be computed once, which speeds up the assembly.

The boundary conforming Delaunay property yields $T_\sigma \geq 0$ and therefore $[\gamma] \geq 0$, see [58, Remark, p. 2504]. If the diffusion coefficients are independent of the state, the resulting matrix is an irreducible M-Matrix, since $J_D[\mathbf{a}]^{-1}$ is symmetric positive semidefinite.

Remark 4.1.2. *If the diffusion coefficients dependent on the state variable one has to apply the product rule to (4.1.6) and differentiate $[Y_{\nu T}]$ with respect to the dependent chemical potentials. For the first term of the product we obtain (4.1.7). For the differentiation of the vector $Y_{\nu T}$ with respect to the dependent chemical potentials it is useful to apply the identity*

$$\sum_{T \in \mathcal{T}} G_2^T[Y_{\nu T}][\gamma] G_2 \mathbf{a}_{\nu T}^{(0)} = \sum_{T \in \mathcal{T}} G_2^T[\gamma][G_2 \mathbf{a}_{\nu T}^{(0)}] Y_{\nu T} \quad (4.1.8)$$

and then to differentiate with respect to the dependent chemical potentials. Every depending species in a diffusion coefficient creates an additional off-diagonal block of the size $M \times M$ in the matrix J_D . The structure of this block is equal to the structure of the diagonal block, since (4.1.7) and (4.1.8) have a identical structure per triangle.

Mass invariance After linearization we search for a vector δ_v such that

$$J_{\hat{\mathcal{B}}}(\mathbf{v}^{(0)}) \delta_v = -\hat{\mathcal{B}}(\mathbf{v}^{(0)}) \quad (4.1.9)$$

and $\hat{E}(\mathbf{a}^{(0)} \delta_v) \in \hat{\mathcal{U}} + \mathcal{U}$. This means, we are only interested in solutions fulfilling Lemma 3.4.1. The mass invariance property is essential for large time steps. Only together with this property the free energy takes its minimal value at thermodynamic equilibrium, see Section 3.4. Numerical experiments in Section 5.2 show a drift of the invariants after some time steps. Hence we take care of the constraint. This phenomenon is also known in the context of ODE's with integral constraints, see [24, 71, 92].

Let $m_{\mathcal{S}^\perp} = \dim \hat{\mathcal{U}}^\perp$ be the number of invariants in the system. The function $g : \mathbb{R}^{Mm} \rightarrow \mathbb{R}^{m_{\mathcal{S}^\perp}}$ is introduced by

$$g(\mathbf{w}) = \left(\left\langle \bar{\mathbf{u}} e^{\mathbf{w}} - \mathbf{U}, \mathbf{w}^\perp \right\rangle_{\mathbb{R}^{Mm}} \right)_{\mathbf{w}^\perp \in \hat{\mathcal{U}}^\perp}, \quad \mathbf{w} \in \mathbb{R}^{Mm},$$

see Section 3.4. Then the constraint $\widehat{E}\mathbf{v}^{(i)} \in \widehat{\mathcal{U}} + \mathbf{U}$ is equivalent to $g(\mathbf{v}^{(i)}) = \mathbf{0}$, $i \in \mathbb{N}$. Linearizing g around $\mathbf{v}^{(0)}$ leads to the linear system $J_N(\mathbf{v}^{(0)})\boldsymbol{\delta}_\lambda = -g(\mathbf{v}^{(0)})$ with the Jacobian matrix $J_N^T \in \mathbb{R}^{Mm \times m_{S^\perp}}$ and a vector $\boldsymbol{\delta}_\lambda \in \mathbb{R}^{m_{S^\perp}}$. The columns of the Jacobian consist of

$$J_N^T(\mathbf{v}^{(0)}) = \left(\left[\mathbf{w}^\perp \right] \mathbf{u}^{(0)} \right)_{\mathbf{w}^\perp \in \widehat{\mathcal{U}}^\perp}. \quad (4.1.10)$$

Following [52, §4] we introduce the following two bilinear continuous forms $a(\cdot, \cdot) : \mathbb{R}^{Mm} \times \mathbb{R}^{Mm} \rightarrow \mathbb{R}$ and $b(\cdot, \cdot) : \mathbb{R}^{Mm} \times \mathbb{R}^{m_{S^\perp}} \rightarrow \mathbb{R}$ by

$$\begin{aligned} a(\boldsymbol{\delta}_v, \mathbf{w}) &= \langle J_{\widehat{\mathcal{B}}}\boldsymbol{\delta}_v, \mathbf{w} \rangle \quad \forall \mathbf{w}, \boldsymbol{\delta}_v \in \mathbb{R}^{Mm}, \\ b(\boldsymbol{\delta}_v, \boldsymbol{\mu}) &= \langle J_N\boldsymbol{\delta}_v, \boldsymbol{\mu} \rangle \quad \forall \boldsymbol{\delta}_v \in \mathbb{R}^{Mm}, \quad \forall \boldsymbol{\mu} \in \mathbb{R}^{m_{S^\perp}}. \end{aligned}$$

We consider the following problem: Given $\widehat{\mathcal{B}}(\mathbf{v}^{(0)}) \in \mathbb{R}^{Mm}$ and $\mathbf{g}(\mathbf{v}^{(0)}) \in \mathbb{R}^{m_{S^\perp}}$, find $(\boldsymbol{\delta}_v, \boldsymbol{\delta}_\lambda) \in \mathbb{R}^{Mm} \times \mathbb{R}^{m_{S^\perp}}$ such that

$$\begin{aligned} a(\boldsymbol{\delta}_v, \mathbf{w}) + b(\mathbf{w}, \boldsymbol{\delta}_\lambda) &= \langle -\widehat{\mathcal{B}}(\mathbf{v}^{(0)}), \mathbf{w} \rangle \quad \forall \mathbf{w} \in \mathbb{R}^{Mm}, \\ b(\boldsymbol{\delta}_v, \boldsymbol{\mu}) &= \langle -\mathbf{g}, \boldsymbol{\mu} \rangle \quad \forall \boldsymbol{\mu} \in \mathbb{R}^{m_{S^\perp}}. \end{aligned}$$

Since $b(\boldsymbol{\delta}_v, \boldsymbol{\mu}) = \langle J_N\boldsymbol{\delta}_v, \boldsymbol{\mu} \rangle = \langle \boldsymbol{\delta}_v, J_N^T\boldsymbol{\mu} \rangle$ holds, we can write: Find $\boldsymbol{\delta} := (\boldsymbol{\delta}_v, \boldsymbol{\delta}_\lambda) \in \mathbb{R}^{Mm} \times \mathbb{R}^{m_{S^\perp}}$ such that

$$\mathcal{K}(\mathbf{v}^{(0)})\boldsymbol{\delta} = \begin{pmatrix} J_{\widehat{\mathcal{B}}}(\mathbf{v}^{(0)}) & J_N^T \\ J_N & 0_{m_{S^\perp}} \end{pmatrix} \begin{pmatrix} \boldsymbol{\delta}_v \\ \boldsymbol{\delta}_\lambda \end{pmatrix} = - \begin{pmatrix} \widehat{\mathcal{B}}(\mathbf{v}^{(0)}) \\ \mathbf{g}(\mathbf{v}^{(0)}) \end{pmatrix} = \mathbf{f}. \quad (4.1.11)$$

If $(\boldsymbol{\delta}, \boldsymbol{\lambda})$ is a solution to (4.1.11), then $\widehat{E}(\mathbf{v}^{(0)} + \boldsymbol{\delta}_v) \in \widehat{\mathcal{U}} + \mathbf{U}$ holds and $\boldsymbol{\delta}_v$ is a solution of the initial problem (4.1.9). The resulting linear system (4.1.11) has a saddle-point structure and the Jacobian \mathcal{K} is a sparse, non symmetric block matrix with dimension $(mM + m_{S^\perp}) \times (mM + m_{S^\perp})$ (in case of MMH $(4M + 2) \times (4M + 2)$).

Remark 4.1.3. *The part $J_R + J_D$ of the Jacobian $J_{\widehat{\mathcal{B}}}$ is underdetermined. If the time step is sufficiently small, the linear system (4.1.2) is regularized by J_T (see (4.1.3)). Since this constraint contradicts long time calculations, the linear system (4.1.2) is regularized by J_N . Further, in order to guarantee the solvability of the system in finite precision arithmetic near the thermodynamic equilibrium (for large time steps) m_{S^\perp} additional Dirichlet values are introduced. This means we delete m_{S^\perp} linear dependent rows and columns in $J_{\widehat{\mathcal{B}}}$ and solve a regularized system. The Newton update $\boldsymbol{\delta}_v$ is then obtained using the Schur complement with respect to the regularized linearized reaction-diffusion system. The Dirichlet values are introduced by adding a large value on the diagonal of the Jacobian at the position responsible for the deleted equations. The selection of the deleted equations is arbitrary. In the implementation,*

we delete the equations responsible for control volume K where the largest dissipation (rate) density occurs.

In case of the MMH mechanism $m_{\mathcal{S}^\perp} = \dim(\mathcal{S}^\perp) = 2$ holds, therefore we have two Lagrange multipliers responsible for the two mass invariants. We delete one equation in the block of \mathcal{X}_2 and one equation in the block of \mathcal{X}_3 . The resulting full Jacobian of the system has four additional columns and rows.

Algorithm The implemented Newton method is depicted in Algorithm 4.1.4.

Algorithm 4.1.4 (Newton($\mathbf{a}^{(0)}, \varepsilon, \hat{c}$)).

Input: Predictor $\mathbf{a}^{(0)} = \mathbf{e}^{\mathbf{v}^{(0)}}$, $\varepsilon_r > 0, \varepsilon_m > 0, \varepsilon_\Psi \geq 0, \hat{c} > 0, \bar{n}_{nonl}$

Output: Approximation $\mathbf{a}^{(n_{nonl})}$ of the solution at time t_n after $n_{nonl} = i$ Newton steps.

Set $i = 0$.

Do:

1. Solve $\mathcal{K}(\mathbf{v}^{(i)})\boldsymbol{\delta} = \mathbf{f}$
2. If $|\delta_{v,j}| < \hat{c}$ for at most one $j = 1, \dots, M$ set $\delta_{v,j} = \text{sign}(\delta_{v,j})\hat{c}$
3. Update $\mathbf{v}^{(i+1)} = \mathbf{v}^{(i)} \mathbf{e}^{\boldsymbol{\delta}_v}$ and $i = i + 1$

As long as:

- $i \leq \bar{n}_{nonl}$, or
- $\|\boldsymbol{\delta}_v\|_{L^\infty} > \varepsilon_r$, and
- $\|g(\mathbf{v}^{(i)})\|_{L^\infty} > \varepsilon_m$, and
- $\hat{\Psi}(\bar{\mathbf{u}}\mathbf{a}^{(i)}) - \hat{\Psi}(\mathbf{u}(t_{n-1})) > \varepsilon_\Psi \hat{\Psi}(\mathbf{U})$ holds.

Since the chemical potentials are the primary variables, the update is multiplicative rather than additive, which preserves the positivity of the activities during the iteration, see [48]. In step 2 of Algorithm 4.1.4 we cut large updates in the potentials. This is an 'implicit' damping of the Newton method due to the reliability of the first order expansion of the exponential function (Boltzmann statistics).

The Newton iteration stops if $\|\boldsymbol{\delta}_v\|_{L^\infty} \leq \varepsilon_r$, where $\boldsymbol{\delta}_v$ is the Newton update of chemical potentials and ε_r is a given tolerance [48]. The second criterion $\|g(\mathbf{v}^{(i)})\|_{L^\infty} \leq \varepsilon_m$ controls the error of the mass invariants. The third criterion controls the decay of the relative free energy. Due to Lemma 3.6.1 this solution is only true if the free energy decays in time. Due to roundoff errors near the thermodynamic equilibrium this criterion cannot be exactly fulfilled and must be relaxed by some positive ε_Ψ . In Chapter 5 we choose $\varepsilon_r = 10^{-4}$, $\varepsilon_m = 10^{-15}$, and $\varepsilon_\Psi = 10^{-10}$.

Typically the Newton method requires more steps to fulfill the mass invariance error, hence the first criterion is fulfilled with higher accuracy than required.

4.2 Solution of linear equations

After linearization it remains to solve a linear system per Newton step. In a first step we answer the question invertibility of the linear system 4.1.2. In a second step we solve the linear system by a direct solver.

Direct solution

The full Jacobian \mathcal{K} is inverted using the Schur complement of the matrix. Multiplying \mathcal{K} by $-J_{\hat{\mathcal{B}}}^{-1}$ and adding the first block row and the second block row we obtain

$$\begin{pmatrix} J_{\hat{\mathcal{B}}} & J_N \\ 0 & -J_N^T J_{\hat{\mathcal{B}}}^{-1} J_N \end{pmatrix} \begin{pmatrix} \delta_v \\ \lambda \end{pmatrix} = \begin{pmatrix} \hat{\mathcal{B}} \\ \mathbf{g} - J_N^T J_{\hat{\mathcal{B}}}^{-1} \hat{\mathcal{B}} \end{pmatrix}. \quad (4.2.12)$$

In a first step the systems

$$J_{\hat{\mathcal{B}}} \tilde{\delta} = \hat{\mathcal{B}}, \quad J_{\hat{\mathcal{B}}} \tilde{\delta} = \hat{\mathcal{B}}$$

are solved using the sparse direct solver Pardiso [117]. Next the correction λ is obtained by solving $-J_N^T \tilde{\delta} = \mathbf{g} - J_N^T \tilde{\delta}$ using the LAPACK routine dgesv. Using λ we obtain the Newton update by $\delta = \tilde{\delta} - \delta_{J_N} \lambda$. This calculation is applicable if $\text{rk } J_N \ll \text{rk } J_{\hat{\mathcal{B}}}$, which holds for many applications. In the example of MMH $\text{rk } J_N$ is 4 (two invariants and two 'Dirichlet points'). In order to save some computational time, the LU-factorization of $J_{\hat{\mathcal{B}}}$ can be frozen over some Newton steps. The Newton update is computed using the frozen LU-factorization as preconditioner used in cgs iterative method. A new factorization is calculated if the iteration time is larger than the factorization time, see [112].

Remark 4.2.1. *We have done experiments with fully iterative methods for the solution of the linear system in case of MMH-Kinetic. Due to the large differences of the diffusion and reaction timescales, the strong coupling of the equations, and the constraints of the mass invariants, the development of a fully iterative method is challenging. The different regimes of the system need appropriate preconditioners which cover the physics of the system. In addition, the method must ensure the invertibility of the block row containing the equations of species 2 and 3 independently of the time step. This is a topic of future research.*

4.3 Adaptive time stepping

An important topic with respect to long time calculations is the adaptation of the time step. The goal is to make large time steps under the consideration that the discretization error is acceptable and the qualitative properties are preserved. In this spirit, the presented method adapts the time step to

the number of Newton steps and to the variation of suitable functionals. In contrast to the standard adaptive time step schemes [24, 71, 87], the main difference in this approach is the use of suitable functionals providing informations about the time scales of the system. This approach works well in the simulation of semiconductors, see [42, 57].

Let \bar{n}_{nonl} be the maximal number of nonlinear iterations and n_{ord} the order of the time discretization scheme, see [24, 70]. In case of the implicit Euler method $n_{ord} = 1$ holds. Moreover let g_{inc} , g_{dec} be given factors for the increment, decrement of the time step, and g_{sav} is a safety factor.

Let n_F be the number of all controlled functionals and the predicted value $\mathbf{u}^{(0)}$ for the solution $\mathbf{u}(t_n)$ at time t_n , $n > 0$. The method calculates the maximum relative error between the predicted value $\widehat{\Theta}_i^{(1)} = \widehat{\Theta}_i(\mathbf{u}^{(0)})$ and the computed value $\widehat{\Theta}_i^{(2)} = \widehat{\Theta}_i(\mathbf{u}(t_n))$ of all controlled functionals $i = 1, \dots, n_F$, i.e.,

$$\omega = \max_{i=1, \dots, n_F} \frac{|\widehat{\Theta}_i^{(1)} - \widehat{\Theta}_i^{(2)}|}{\left(|\widehat{\Theta}_i^{(2)}| + \varepsilon_{abs}\right) \varepsilon_{rel}}$$

with the relative and absolute error tolerance ε_{rel} and ε_{abs} . The decision whether a time step is rejected or accepted is done by the following two rules:

- The step is rejected if the difference between predictor and corrector is large or and the new time step $t_\delta^{(n+1)}$ is calculated by

$$t_\delta^{(n+1)} = \begin{cases} t_\delta^{(n)} g_{dec}, & \text{if } \omega > 4 \text{ or } n_{nonl} > \bar{n}_{nonl}, \\ g_{sav} \frac{t_\delta^{(n)}}{g_{inc} \omega^{1/(1+n_{ord})}}, & \text{if } \omega > 1. \end{cases}$$

- If the difference between predictor and corrector is acceptable $\omega \leq 1$ a new time step is obtained by

$$t_\delta^{(n+1)} = \min \left(g_{sav} t_\delta^{(n)} \omega^{-1/(1+n_{ord})}, t_\delta^{(n)} g_{inc} \right).$$

and if $3n_{nonl} > \bar{n}_{nonl}$ or $t_\delta^{(n+1)} \leq t_\delta^{(n)}$ the time step is adjusted by the number of Newton iterations, i.e.,

$$t_\delta^{(n+1)} = \begin{cases} (t_\delta^{(n+1)} + t_\delta^{(n)})/2 & \text{if } 2.3n_{nonl} < \bar{n}_{nonl}, \\ t_\delta^{(n)}/2 & \text{if } 1.5n_{nonl} < \bar{n}_{nonl}, \\ t_\delta^{(n)} g_{dec} & \text{if } n_{nonl} > \bar{n}_{nonl}. \end{cases}$$

We choose the relative free energy $\widehat{\Psi}$ and the dissipation \widehat{D} of the system as suitable functionals. As already mentioned in Section 3.6, the free energy is a Lyapunov function of the system, which provides information about the timescales of the system and measures the distance to thermodynamic equilibrium, see also Chapter 5. In Chapter 5 we choose $\bar{n}_{nonl} = 10$, $\varepsilon_{rel} = 0.3$, $\varepsilon_{abs} = 1$, $g_{inc} = 1.6$, $g_{dec} = 0.6$, and $g_{sav} = 0.8$.

Remark 4.3.1. *The method is implemented as backward differentiation formula (BDF) (see [24, 71]), so that it is possible to use a higher integration order. Since a BDF method of order $s > 1$ requires $s - 1$ solutions from the previous time steps, a second integration method must be applied to obtain the necessary initial data. This can be done by using a $s - 1$ step BDF method. We mention that a BDF method is only A-stable for $s \leq 2$ (second Dahlquist barrier), see [71]. Numerical experiments show that higher orders cause more time step rejections on coarse meshes due to non resolved features of the solution, see Section 5.5. Up to now, there are no theoretical investigations about monotonous decay of the free energy in case of higher order time discretizations ($s > 1$). Possible starting points are [18, 29, 44].*

Remark 4.3.2. *We also experimented with other functionals like the distance (2.3.7). In [42] the author notes that this measure is effective for the van Roosebrock system. In our situation we saw no significant differences between (2.3.7) and the used relative free energy and the dissipation rate of the system as controlled functionals.*

5

Numerical Example

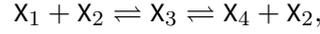
Abstract The aim of this chapter is to illustrate the behavior of the analyzed method. Since the system is nonlinear and an analytical solution is not known, our focus is the preservation of physical properties, a demonstration of the different timescales of the system, and finally reaching the thermodynamic equilibrium. The examples are selected in such a way that they consider various properties of the scheme. In order to cover the wide range of applications in biology and chemistry, the used parameters have sometimes a large variation. The ranges are inspired by [30, 41, 106]. At the beginning, we rephrase the Michaelis-Menten-Henri reaction mechanism. Section 5.2 illustrates the preservation of the invariance property of the discrete solution. In Section 5.3, we consider the first simple example including a heterogeneous material. The solutions of the system may contain layers on the boundary of two materials. In the forthcoming section, the method is applied to examples of different complexity. The example of Section 5.4 includes anisotropic meshes and is inspired by a process coming from the manufacturing of semiconductors, see Section 1.3. Then we are interested in the convergence behavior of the relative free energy and dissipation rate on varying meshes. Additionally we replace the implicit Euler method by a second order backward differentiation formula for time integration. In Section 5.6, we consider an example, where three materials meet each other in one point. The challenges of Section 5.7, a daisy with 12 petals, are a non convex domain, vanishing reaction rate coefficients in some regions of the domain. The chapter is closed by an example in Section 5.8 about the behaving of the scheme if the reactions becomes fast and a summary in Section 5.9.

5.1 Introduction

In the following we present different examples involving the MMH reaction mechanism [17], see Example 1.4.1. The reaction step is present in the modeling of a pattern doubling process by a catalytic cross-linking of a spacer, triggered by residual acid diffusion from a previously developed primary

structure into the spacer, see Subsection 1.3.2. Double patterning is a possible option used in optical lithography to create lines with a structure size less than 48nm, see Section 1.3.

As already stated in Example 1.4.1, the reaction step can be symbolically expressed by



where X_1 is a precursor, X_2 is an acid, X_3 is an intermediate, and X_4 is a cross-linked polymer. Correspondingly, the set \mathcal{R} consists of two pairs of vectors, namely $\alpha_1 = (1, 1, 0, 0)$ and $\beta_1 = (0, 0, 1, 0)$ for the first reaction as well as $\alpha_2 = (0, 0, 1, 0)$ and $\beta_2 = (0, 1, 0, 1)$ for the second reaction. Due to the definition of the reaction term (1.4.6), the net production rates of the species are given by

$$\begin{aligned} R_1(a) &= -k_{(\alpha_1, \beta_1)}(a_1 a_2 - a_3), \\ R_2(a) &= -k_{(\alpha_1, \beta_1)}(a_1 a_2 - a_3) && + k_{(\alpha_2, \beta_2)}(a_3 - a_2 a_4), \\ R_3(a) &= +k_{(\alpha_1, \beta_1)}(a_1 a_2 - a_3) && - k_{(\alpha_2, \beta_2)}(a_3 - a_2 a_4), \\ R_4(a) &= && + k_{(\alpha_2, \beta_2)}(a_3 - a_2 a_4). \end{aligned}$$

The stoichiometric subspace \mathcal{S} and its orthogonal complement \mathcal{S}^\perp are spanned by

$$\mathcal{S} = \text{span}\{(1, 1, -1, 0), (0, -1, 1, -1)\}, \quad \mathcal{S}^\perp = \text{span}\{(1, 0, 1, 1), (0, 1, 1, 0)\}.$$

Every element of \mathcal{S}^\perp represents one invariant of the system, i.e., $R_1 + R_3 + R_4 \equiv 0$ and $R_2 + R_3 \equiv 0$ hold, and hence

$$\int_{\Omega} (u_1 + u_3 + u_4)(t) dx = \int_{\Omega} (U_1 + U_3 + U_4) dx, \quad (5.1.1)$$

$$\int_{\Omega} (u_2 + u_3)(t) dx = \int_{\Omega} (U_2 + U_3) dx \quad \forall t \geq 0 \quad (5.1.2)$$

are conserved during the time evolution. At thermodynamic equilibrium the chemical activities are constant, therefore the solution can be obtained by solving

$$0 = R_1(a^*) = a_1^* a_2^* - a_3^*, \quad 0 = R_4(a^*) = a_2^* a_4^* - a_3^*$$

together with (5.1.1), (5.1.2). In the model that we have in mind, see Subsection 1.3.2, the diffusion coefficient of X_2 is given by

$$D_2(a_1, a_4) = D_{20} \exp \left(-\varphi_3 \frac{\varphi_1 \bar{u}_1 a_1 + (1 - \varphi_1) \bar{u}_4 a_4}{\varphi_2 \bar{u}_1 a_1 + (1 - \varphi_2) \bar{u}_4 a_4} \right),$$

where $\varphi_1, \varphi_2, \varphi_3$ are constants, see [41]. All other diffusion coefficients are assumed to be piecewise constant and independent of the concentration.

5.2 Homogeneous material

In a first example we illustrate the effect of the four Lagrangian multipliers (two for the invariants and the two for the Dirichlet values). The calculation is done for the domain $\Omega = (-1, 1)^2$ on a grid with 3260 nodes created by Triangle [120]. In order to show the long time stability of the method, we choose the time interval $S = [0, 10^{25}]$.

We choose diffusion coefficients $(D_1, D_{20}, D_3, D_4) = (0.01, 10^{-4}, 0.1, 1)$, $\varphi_1 = 1, \varphi_2 = 1, \varphi_3 = 0$ and reaction coefficients $k_{+1} = 10^9, k_{-1} = 3 \cdot 10^2, k_{+2} = 3 \cdot 10^{-5}, k_{-2} = 40$. Using the results of Section A.1, we can choose the reference densities and reaction rate coefficients such that the chemical potentials are constant in thermodynamic equilibrium.

The species concentrations are plotted in Figure 5.1 at times $t = 10^{-10}, 0.1, 1.0, 10.0$ and 10^{10} . Every column in Figure 5.1 represents a gray straight line in Figure 5.2. First the green humps of X_1 and X_2 react. Then X_1 diffuses and eats X_2 away. If the hump of X_2 is eaten up, the system reacts to the steady state.

We consider three experiments: Each of the three figures in Figure 5.2 show the time dependence of the relative free energy and the dissipation rate divided into the reaction and diffusion part, as well as the relative error of the invariants (5.1.1) and (5.1.2).

1. In the first experiment, the only stopping criterion for Newton's method is $\|\delta_v\|_{L^\infty} \leq \varepsilon_r$, where δ_v is the Newton update of chemical potentials and ε_r is a given tolerance, see Algorithm 4.1.4. In Figure 5.2a, the free energy decays monotonously and the different parts of the dissipation decay, too. Near to the thermodynamic equilibrium, the time step is not sufficiently small to guarantee the invertibility of the Jacobian in finite arithmetic. Therefore the system violates the invariants to the extent that the linear system is solvable since the stopping criterion ε_r (see Algorithm 4.1.4) is too slack. The invariant violation results in a larger deviation of the relative free energy from its minimum value and from the thermodynamic equilibrium solution. Moreover the method is not stable at thermodynamic equilibrium. After some time, the relative free energy begins to grow strongly.
2. The growth of the invariant error can be controlled by a sharper ε_r . Another option is to include the relative mass invariant error as an additional stopping criterion. This was done in the second calculation. The results are depicted in Figure 5.2b. The invariants are well preserved during the calculation and the free energy decays as expected. Long before the thermodynamic equilibrium of the system, the simulation runs into trouble due to the singularity of the Jacobian in finite precision arithmetic. Hence the time step must be reduced in order to

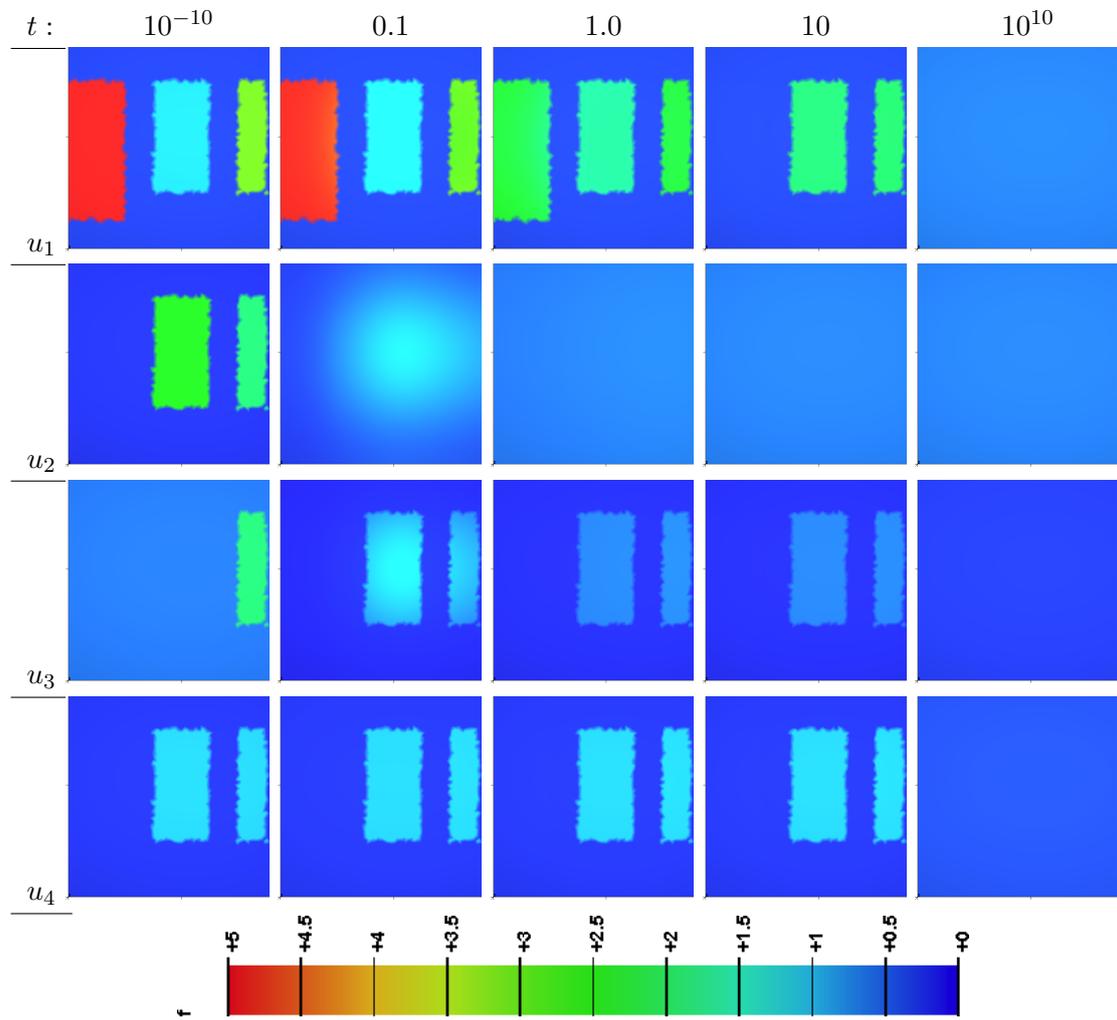


Figure 5.1: Concentrations of all species at different times. Every column represents a gray straight line in Figure 5.2

regularize the system. The simulation was cancelled after 835 time steps ($t = 7.381 \cdot 10^{-10}$) due to the large number of rejected time steps.

3. In the last experiment, we use the proposed method with the four Lagrangian multipliers in order to obtain the solution. The relative error of the two invariants (5.1.1) and (5.1.2) is of the order of machine epsilon $2.22 \cdot 10^{-16}$. The plot of Figure 5.2c shows the expected monotonous (and exponential) decay of the free energy and the expected non-negativity of the dissipation rate up to roundoff errors.

The number of time steps and Newton steps are summarized in Table 5.1. In contrast to the first naive approach, the last approach results in a slightly higher number of Newton steps. But as already

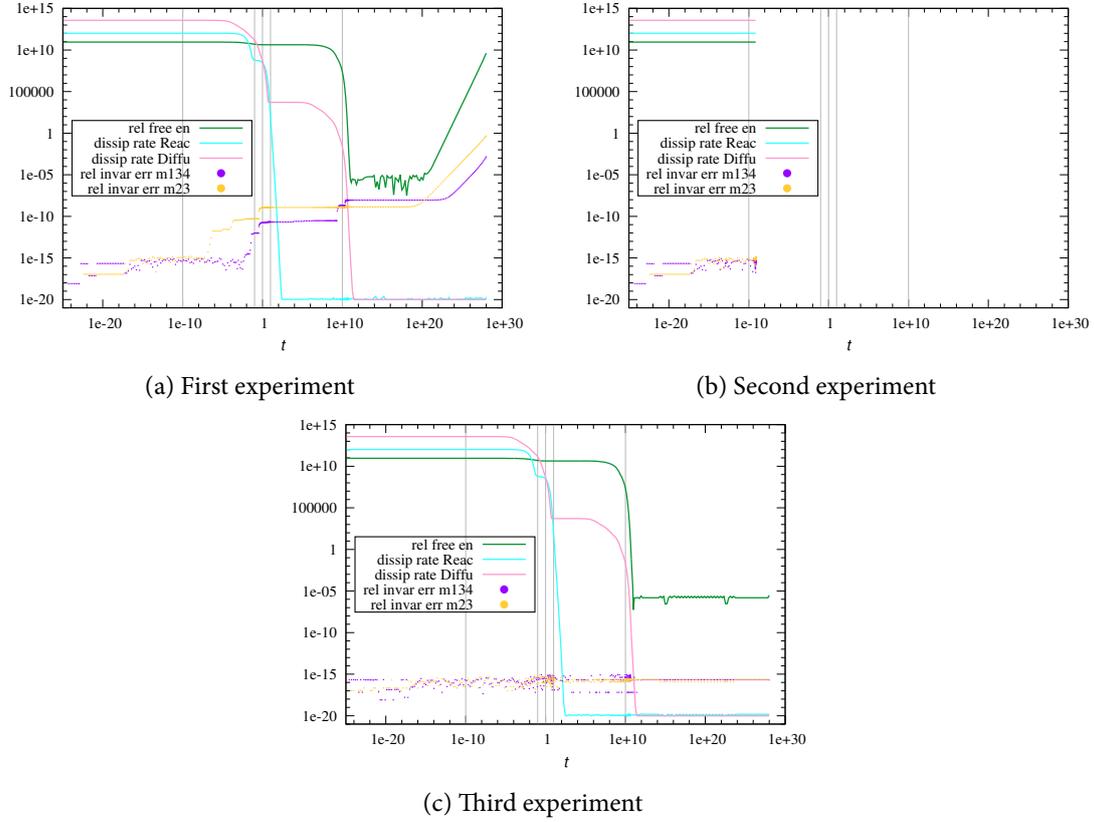


Figure 5.2: Time dependence of the relative free energy (red) and dissipation rate (diffusion part - green, reaction part - blue). Invariants conservation (rel. error, light blue, pink). Every gray straight line represents one column in Figure 5.1

mentioned, the first approach violates the invariance property of the solution which leads to an improper approximation of the thermodynamic equilibrium solution. Due to the time step constraint, we register a huge number of rejected time steps.

Alg. variant	1	2	3
Acc. Newton	654	<i>NA</i>	871
$\#t_\delta$ tot.(rej.)	413 (0)	835 (286)	402 (0)
Ratio	1.58	<i>NA</i>	2.17

Table 5.1: Summary of the simulation: accumulated Newton steps, total number of time steps including rejections and mean value of Newton steps per time steps. The second experiment was cancelled at $t = 7.381 \cdot 10^{-10}$

5.3 Layer behavior

We consider the MMH-kinetic on a domain $\Omega = \Omega_1 \cup \Omega_2$ consisting of two different materials $\Omega_1 = (0, 1] \times (0, 1)$ and $\Omega_2 = [1, 2) \times (0, 1)$. The mesh is depicted in Figure 5.3 and consists of 609 nodes and 808 triangles. The subdomains are aligned to the triangle edges. The mesh is strongly graded near the boundary of the two different materials, since we expect strong gradients in the species concentration in this region.

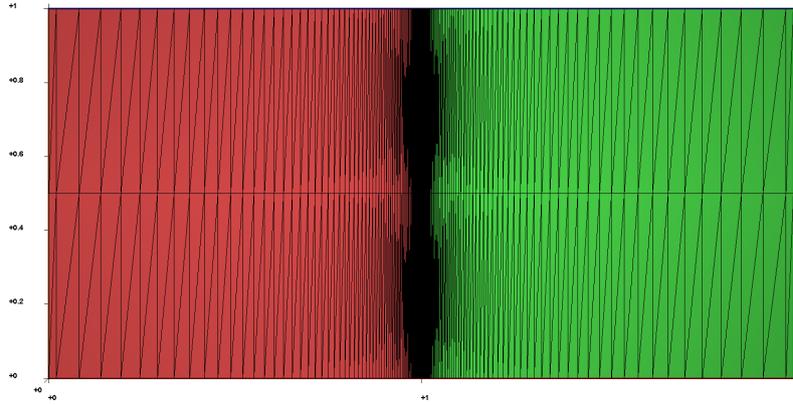


Figure 5.3: The used acute mesh which is graded in the neighborhood of the layers

The material data are collected in Table 5.2. The parameters model the situation where X_1 is

	Ω_1 and Ω_2
$(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$	$(10^{-18}, 10^{-8}, 1, 10^{20})$
(U_1, U_2, U_3, U_4)	$(1, 0.5, 10^{-20}, 10^{-20})$
(D_1, D_{20}, D_3, D_4)	$(10^5, 10^5, 10^5, 10^5)$
(K_1, K_2)	$(10^{-9}, 10^6)$

Table 5.2: Material parameters for both regions. The diffusion is assumed to be independent of the state variable

catalytically converted into X_4 . The lifetime of the intermediate X_3 is short and the catalyst X_2 is released after the reaction is finished.

Variation of the reference density of X_1 In the first experiment we vary the reference density \bar{u}_1 in Ω_2 in the range of $\bar{u}_{1,h}(x) = (1, 0.4, 0.133) \cdot 10^{-18} \forall x \in \Omega_2$. For $\bar{u}_{1,h}(x) = 10^{-18} \forall x \in \Omega_2$ both materials are equal and the system acts like an ODE system due to the spatially homogeneous initial

value. For the three other values of the reference density the reaction runs faster in Ω_2 than in Ω_1 (Note that $k_{(\alpha_1, \beta_1)} = k_{+1} \bar{u}_1 \bar{u}_2$, see also Section A.1).

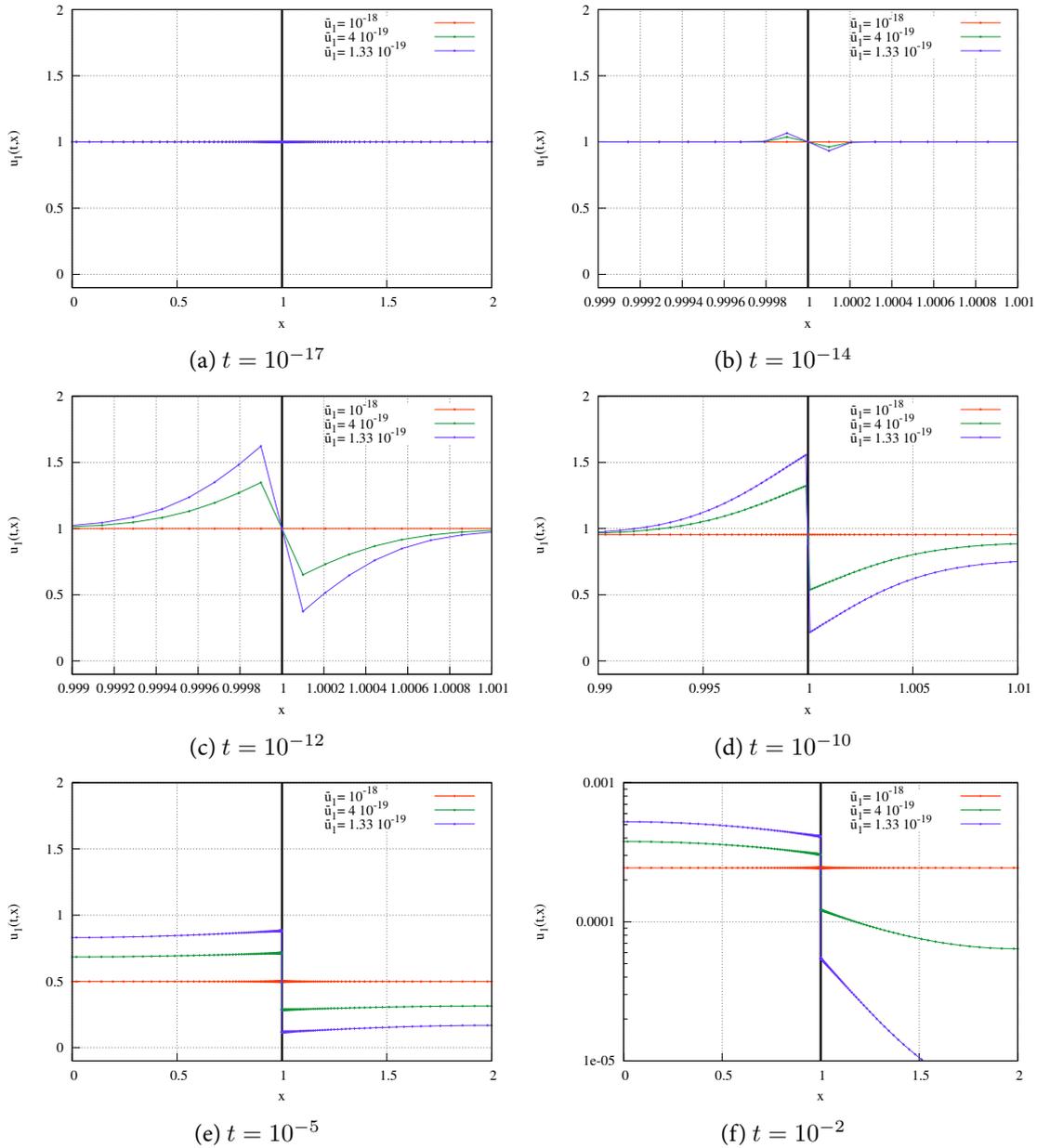


Figure 5.4: Cross section along $y = 0$ of the concentration $u_{1,h}$ for different reference densities $\bar{u}_{1,h}(x) = (1, 0.4, 0.133) \cdot 10^{-18} \forall x \in \Omega_2$ at different time steps. For a better visibility of the layer, the x-axis and y-axis have a different scaling. Note also the log scale of the y-axis in Figure 5.4f

In Figure 5.4 the evolution of the concentration of species X_1 along $y = 0$ is depicted at different times. Until $t = 10^{-17}$ the concentration of X_1 is constant. Then the system recognizes the different

materials and the resulting gradient in the activities. The diffusion processes start to equilibrate the gradient in the activities which results in a layer in the neighborhood of the boundary of Ω_1 and Ω_2 in the concentration of X_1 . The concentration jumps from the last mesh point contained in Ω_1 and not lying on the boundary of the regions to the mesh point lying on the boundary and the first mesh point contained in Ω_2 . The height of the layer seems proportional to the difference of the reference densities. The maximum is located in the region, where the reference density is higher, i.e., the layer increases left to the boundary and decreases right to the boundary, see Figure 5.4b and 5.4c. Swapping the parameter in Ω_1 and Ω_2 changes the location of the extreme values of the layer. Due to the diffusion process the layer is propagated from the boundary into the domain.

In Figure 5.4e and 5.4f the reaction of X_1 and X_2 to X_3 starts and the concentration of X_1 decays in space. Since the speed of the forward and backward reaction depends on the reference densities, the decay is differently fast in Ω_2 , i.e., the species with smallest reference density decays faster in Ω_2 than all others.

Variation of diffusion coefficients In this experiment, we vary uniformly the diffusion coefficient $D_\nu = \bar{D}$, $\nu = 1, \dots, 4$. We set $\bar{D} = 10^{-2}, 10^5, 10^6$. All other parameters are summarized in Table 5.3. The evolution of the concentration of X_1 is depicted in Figure 5.5 at different times. We see that

	Ω_1	Ω_2
$(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$	$(10^{-18}, 1, 1, 10^8)$	$(4 \cdot 10^{-18}, 1, 1, 10^8)$
(U_1, U_2, U_3, U_4)	$(1, 0.5, 10^{-20}, 10^{-20})$	$(1, 0.5, 10^{-20}, 10^{-20})$
\bar{D}	10^5	$10^{-2}, 10^5, 10^6$
(K_1, K_2)	$(10^{-9}, 10^6)$	$(10^{-9}, 10^6)$

Table 5.3: Material parameters for the red and green region. The diffusion is assumed to be independent of the state variable

larger diffusion coefficients lead to a larger width of the layer in Figure 5.5b, since the smoothing effect of the diffusion is greater. Due to the faster diffusion the red curve is strongly propagated into the domain, see Figure 5.5c and 5.5d. Since the diffusion coefficient corresponding to the blue curve is very small, the system acts like an ODE-system and therefore the concentration profile stays constant until the reaction starts in Figure 5.5c. The species X_1 decays and creates X_3 therefore the concentration of X_1 decreases. Since the reference density of X_1 is larger in Ω_2 than in Ω_1 the concentration profiles decay faster in Ω_1 than in Ω_2 , see Figure 5.5d. In Figure 5.5e the diffusion has almost completely equalized the gradient in the activities and therefore the green and the red curve are close together. In the calculation for the blue curve the diffusion process starts which creates a layer near the material boundary. From Figure 5.5e to 5.5f the conversion of X_1 to X_4 proceeds and

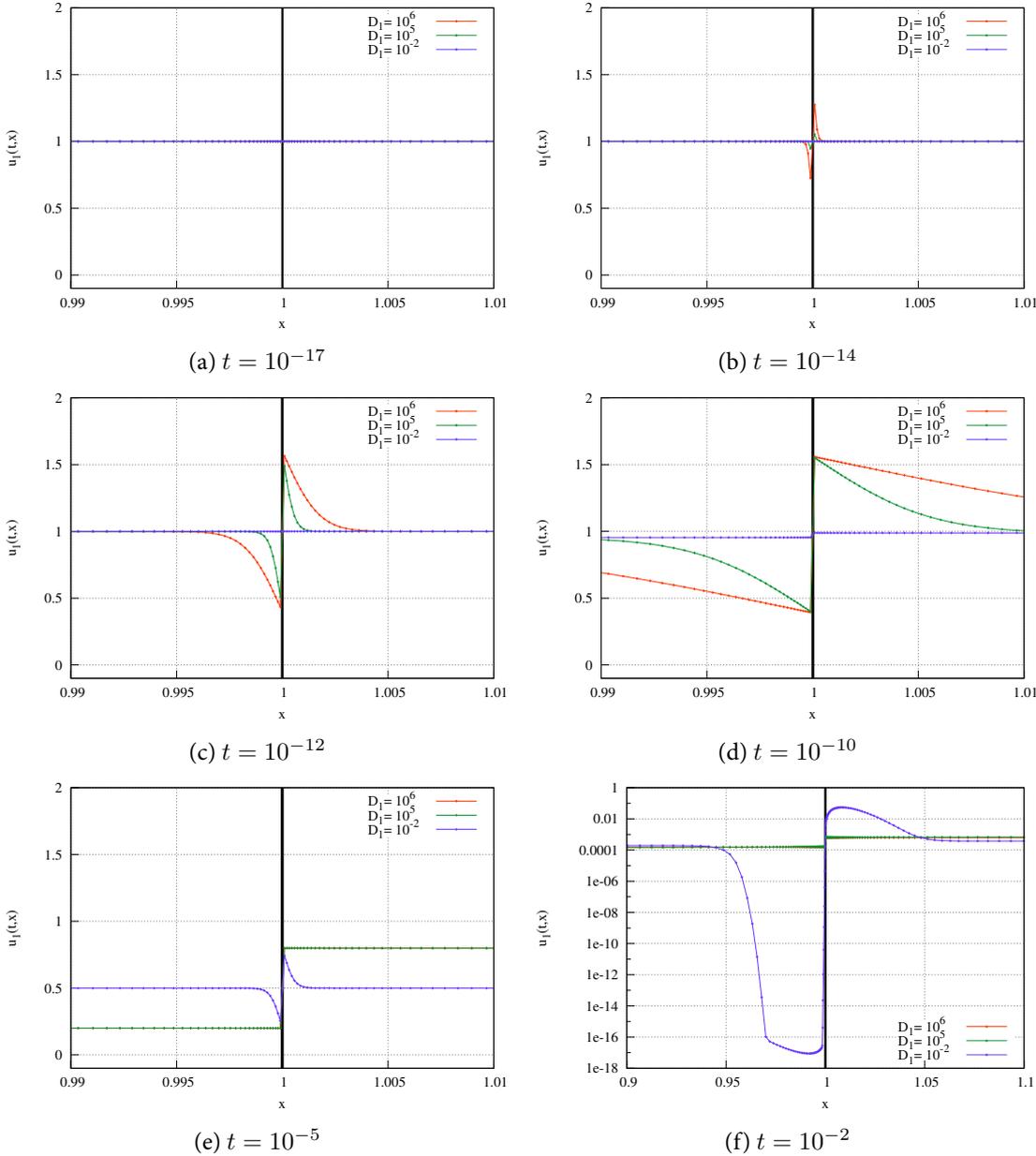


Figure 5.5: Cross section along $y = 0$ of the concentration $u_{1,h}$ for different diffusion coefficients at different times. The red line is covered by the green one. Note the different scaling of the x-axis and in Figure 5.5f the log scale of the y-axis

the concentration of X_1 decays. Since the diffusion coefficients of the blue curve are very small the layer is still there.

Summary We saw that heterogeneous materials lead to jumps on the material boundary in the species concentration. Due to different constellations of the simulation parameters and the size of the chemical activities, the behavior near the layer is variable. For example it would be possible that in Ω_1 the reaction creates X_1 and in Ω_2 the reaction destroys X_1 , i.e., the reaction acts like a dipole near the internal boundary. The maximum and minimum of the layer is located on the closest mesh point inside each region. The spatial expansion of the layer depends on the size of the diffusion coefficients. Fast reactions and slow diffusion lead to very sharp layers, which are hardly resolved by the mesh.

5.4 Moving front

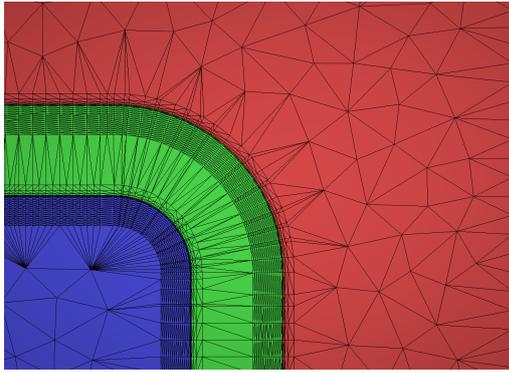
We consider a heterostructure consisting of two different materials. The example illustrates the process introduced in Section 1.3.2, where a pattern is duplicated by catalytic cross-linking of a spacer, triggered by residual acid diffusion from a previously developed primary structure into the spacer. A description of the process is given in Subsection 1.3.1. The interface between different materials is aligned to the Delaunay triangulation that is dual to the Voronoi grid. One half of the used mesh is depicted in Figure 5.6. Instead of showing the Voronoi boxes we show the dual Delaunay triangulation. The materials are represented by different colors in the mesh and the used parameters are collected in Table 5.4.

	red and green region	blue region
$(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$	$(10^{-16}, 10, 10, 10^{12})$	$(10^{-20}, 1, 10^{-11}, 1)$
(U_1, U_2, U_3, U_4)	$(1, 10^{-15}, 10^{-15}, 10^{-15})$	$(10^{-15}, 10^{-15}, 1, 10^{-15})$
(D_1, D_{20}, D_3, D_4)	$(10^{-10}, 200, 10^{-10}, 10^{-12})$	$(2 \cdot 10^{-10}, 200, 2 \cdot 10^{-10}, 2 \cdot 10^{-12})$
$(\varphi_3, \varphi_1, \varphi_2)$	$(3.37, 0.65, 0.95)$	$(0, 0.65, 0.95)$
(K_1, K_2)	$(10^{-9}, 10^3)$	$(10^{-21}, 0.1)$

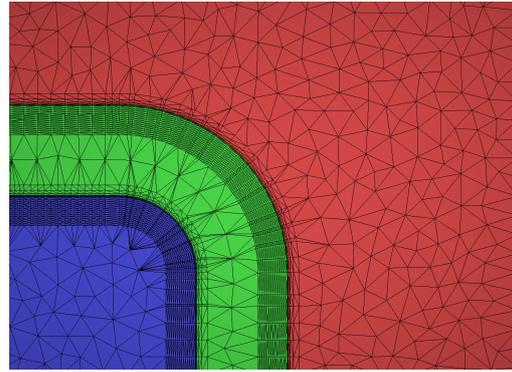
Table 5.4: Material parameters for the three different regions. The parameters for the red and green region are equal. Note the large differences in the reaction coefficients

The simulation parameters in the blue region are chosen in such a way that the catalyst X_2 is created by the intermediate X_3 . In the red and green region the forward direction of the MMH mechanism dominates. We expect that the catalyst X_2 from the blue region diffuses and creates together with X_1 the intermediate X_3 which quickly degrades into X_4 by releasing X_2 . This catalytic cycle leads to a moving reaction front from the primary region into the over-coated region, i.e. the thickness of the reaction front of X_4 grows in the red region, see Figure 5.11. On the slow timescales the created structures dissolve by diffusion and the thermodynamic equilibrium solution is attained. Technologically, the thickness of the reaction front of X_4 can be controlled by an additional fast

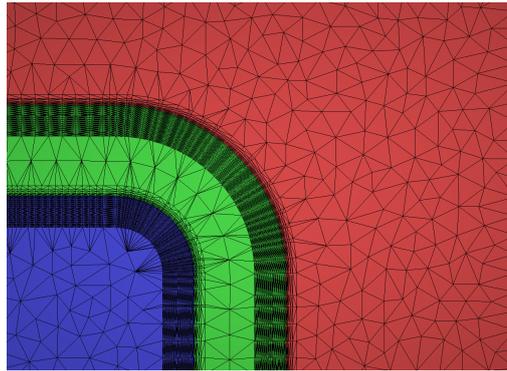
reaction which neutralizes the catalyst X_2 , therefore, the MMH reaction stops, see Subsection 1.3.1. But this part of the technological process is not included in the example. Based on the behavior of the system we align the mesh in the green area. The large gap is introduced in order to show that the used method is also stable on coarse mesh regions, see Figure 5.6.



(a) Mesh \mathcal{M}_1 with 7105 nodes and 13873 triangles



(b) Mesh \mathcal{M}_2 with 7766 nodes and 15195 triangles



(c) Mesh \mathcal{M}_3 with 13944 nodes and 27505 triangles

Figure 5.6: Different meshes used in the simulation. The material are represented by a different color

In order to illustrate that the method converges we consider three different meshes with varying numbers of mesh points, see Table 5.5. The meshes \mathcal{M}_1 and \mathcal{M}_2 are evolved from \mathcal{M}_3 by coarsening the green layer and removing some spider in the mesh \mathcal{M}_1 . The total number of time steps needed by the algorithm in order to solve the problem on the time interval S and the maximal relative error of the two invariants (5.1.1), (5.1.2) are also stated in the table. The time steps are distributed by the time stepping mechanism presented in Section 4.3.

On \mathcal{M}_3 , the thermodynamic equilibrium is given by

$$a_{h_3}^* = (1.195 \cdot 10^{-12}, 1.915 \cdot 10^{-2}, 2.288 \cdot 10^{-14}, 1.195 \cdot 10^{-12}),$$

	# nodes	# tri.	tot.(rej.)	max rel. err (5.1.1)	max rel. err (5.1.2)
\mathcal{D}_1	7105	13873	7369 (25)	$9 \cdot 10^{-15}$	$9.4 \cdot 10^{-15}$
\mathcal{D}_2	7766	15195	10073 (32)	$9.3 \cdot 10^{-15}$	$9.4 \cdot 10^{-15}$
\mathcal{D}_3	13944	27505	11462 (33)	$7.6 \cdot 10^{-15}$	$9.8 \cdot 10^{-15}$

Table 5.5: Characterization of the different discretizations. Total number of time steps $\#t_\delta$ including (rejections) needed by the algorithm in order to solve the problem on $S = [0, 10^{20}]$

and the relative differences compared to the other meshes are smaller than $2 \cdot 10^{-8}$. Figure 5.7 shows the time evolution of the free energy, the dissipation rate divided into the reaction and diffusion part. Every time scale in the system represents one plateau in the curve of the dissipation rate.

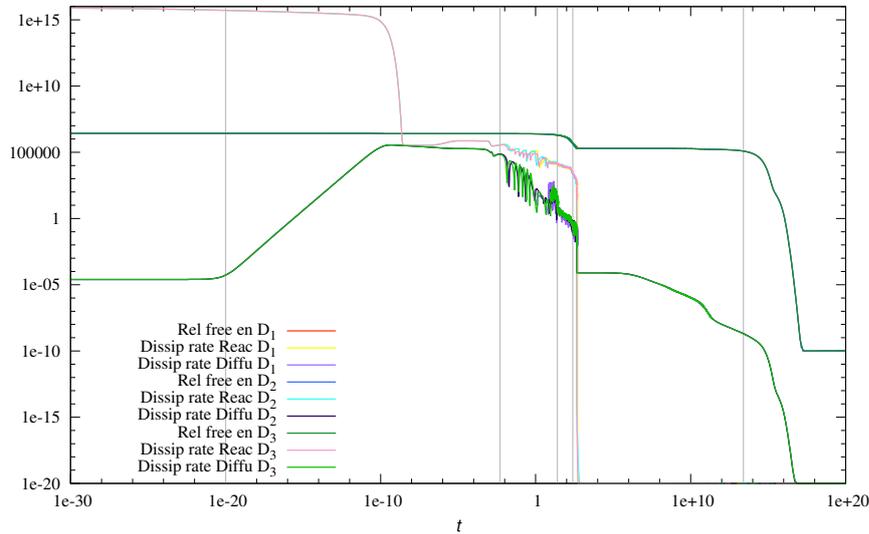


Figure 5.7: Evolution of the the relative free energy functional $\Psi_h(u_h)$ and the reactive and diffusive part of the dissipation rate $D_{R,h}$, $D_{D,h}$ for different discretizations \mathcal{D}_i , $i = 1, 2, 3$. Note that the dissipation rate and free energy is truncated at 10^{-20} and 10^{-10} , respectively. Every gray straight line represents one column in Figure 5.11

The reaction part and the diffusion part provide information whether the system is reaction dominated or diffusion dominated, i.e., in the first interval from 10^{-30} to 10^{-8} the system is reaction dominated. Then energy of the system disappears by the reaction and diffusion. After the whole region is filled with X_4 , the system is diffusion dominated. The curves of the relative free energy and the dissipation rate for different discretizations lie close by. Only during the front movement there are larger deviations, see Figures 5.7 and 5.8. The oscillating behavior of the dissipation rate curves correlates with the number of gridlines in the refined region. In this region, the front has to pass over the same distance, in order to jump to the next grid points. Therefore the period of the

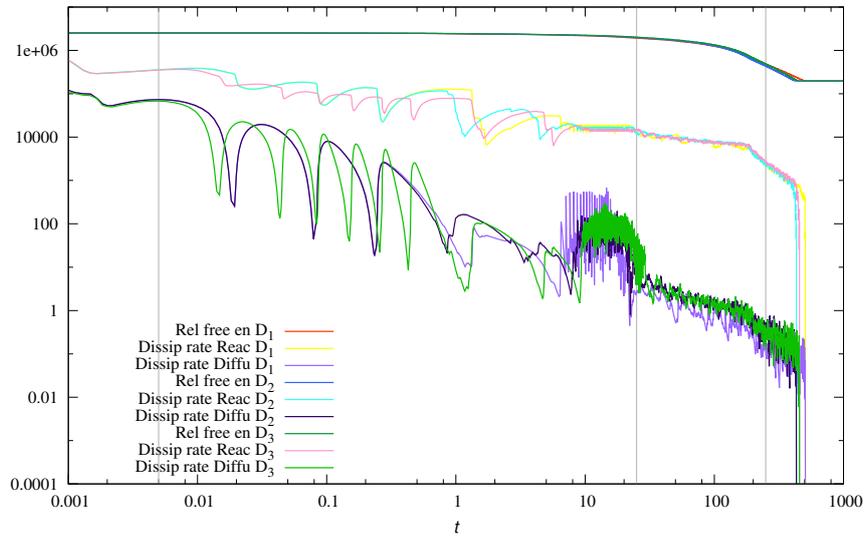


Figure 5.8: Zoom of Figure 5.7 into the region where the front is moving

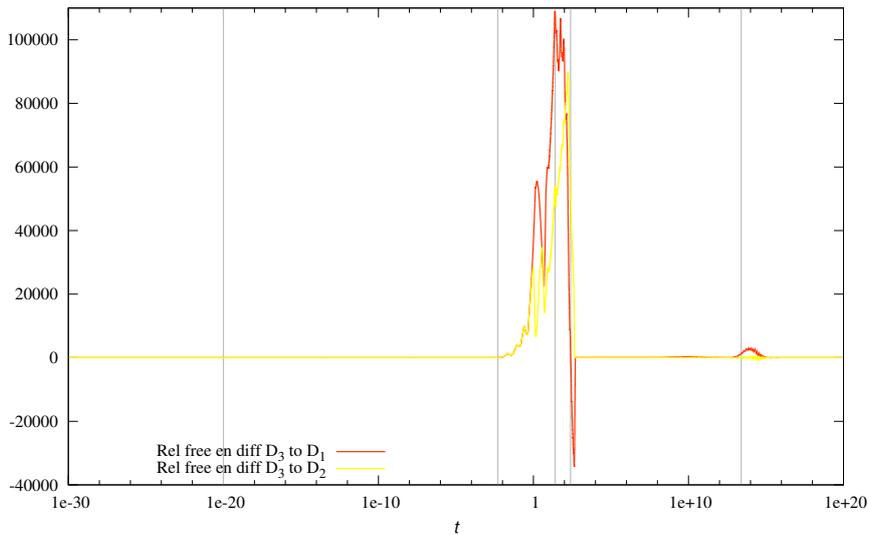


Figure 5.9: Evolution of the difference of the relative free energy functional $\Psi_h(u_h)$ of the discretizations \mathcal{D}_i $i = 1, 2$ relative to the relative free energy curve of discretization \mathcal{D}_3 . Every gray straight line represents one column in Figure 5.11

oscillations is larger than in the non-refined region. The dissipation curves are slightly shifted due to the different refinement level of the layer region, see Figure 5.8. Moreover in Figure 5.8 we see that a time delay in the dissipation curve results in a slower decay of the relative free energy.

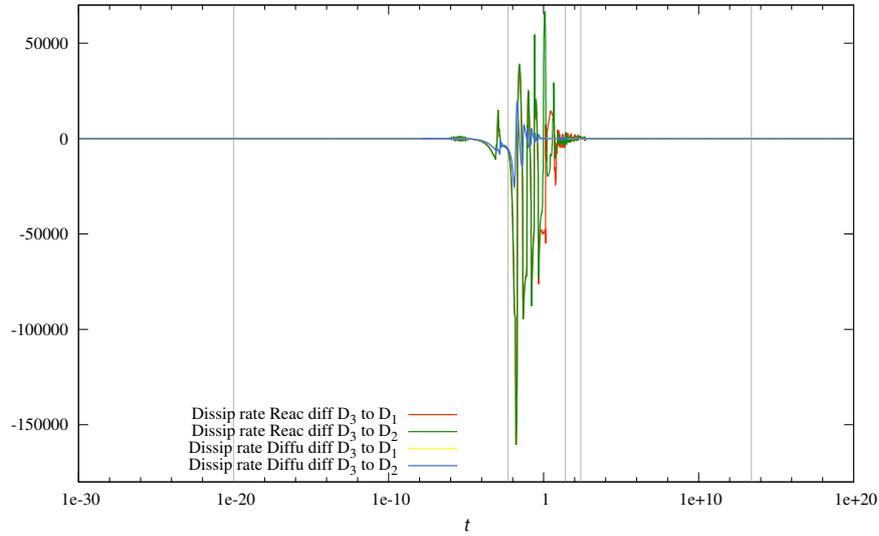


Figure 5.10: Evolution of the difference of the two parts of the dissipation rate $D_{D,h}$ and $D_{R,h}$ of the discretizations \mathcal{D}_i , $i = 1, 2$, relative to the relative free energy curve of discretization \mathcal{D}_3 . Every gray straight line represents one column in Figure 5.11

The deviation of the curves of discretizations \mathcal{D}_1 and \mathcal{D}_2 relating to the curves of \mathcal{D}_3 are plotted in Figure 5.9 and Figure 5.10. The largest differences in the free energy curves are during the front movement (see third to fourth column in Figure 5.11). After that, the curves approach each other. The largest deviations are obtained on the coarsest mesh \mathcal{D}_1 . The differences of the dissipation curves of discretization \mathcal{D}_1 and \mathcal{D}_2 relating to \mathcal{D}_3 are very small. The blue curve is covered by the purple one and the green one covers the red one.

In Table 5.6 the relative $L^2(S)$ -errors of the relative free energy and the dissipation rate are presented. Since the continuous solution of the reaction-diffusion system is not known, a reference solution for the finest discretization \mathcal{D}_3 is computed, and errors are measured against this discrete solution. We are only interested in the convergence behavior w.r.t. the spatial error. Therefore we repeat the simulation on \mathcal{M}_1 and \mathcal{M}_2 with the time steps used for the simulation on mesh \mathcal{M}_3 . The error norms are given by

$$\begin{aligned} \left\| \tilde{\Psi}_{h_{ij}} \right\|_{L^2(S)} &= \left\| \Psi_{h_i}(u_{h_i}) - \Psi_{h_j}(u_{h_j}) \right\|_{L^2(S)} \left\| \Psi_{h_j}(u_{h_j}) \right\|_{L^2(S)}^{-1}, \\ \left\| \tilde{D}_{h_{ij}} \right\|_{L^2(S)} &= \left\| D_{h_i}(u_{h_i}) - D_{h_j}(u_{h_j}) \right\|_{L^2(S)} \left\| D_{h_j}(u_{h_j}) \right\|_{L^2(S)}^{-1}. \end{aligned}$$

From Table 5.6, we conclude that small changes in the mesh quality lead to larger changes in the

	# \mathcal{P}	# t_δ tot.	$\ \tilde{\Psi}_{h_{i3}}\ _{L^2(S)}$	$\ \tilde{D}_{h_{i3}}\ _{L^2(S)}$
\mathcal{D}_1	7105	12020	$2.8157 \cdot 10^{-2}$	$4.6457 \cdot 10^{-6}$
\mathcal{D}_2	7766	11801	$2.2989 \cdot 10^{-4}$	$4.4736 \cdot 10^{-6}$

Table 5.6: $L^2(S)$ -norm of the relative free energy Ψ_{h_i} and the dissipation rate D_{h_i} on mesh \mathcal{M}_i , $i = 1, 2$, related to mesh \mathcal{M}_3 and scaled by the $L^2(S)$ -norm of the relative free energy and dissipation rate of \mathcal{M}_3 . The time steps for the calculation on $\mathcal{M}_1, \mathcal{M}_2$ are included in the time step sequence of \mathcal{M}_3

L^2 -error of the relative free energy than in the dissipation rate. In contrast to \mathcal{M}_1 we only remove some spider in the mesh \mathcal{M}_2 .

In Figure 5.11 the time evolution of the four species concentrations is shown at different time steps. Every column in Figure 5.11 is represented by a straight gray line in the Figures 5.7, 5.9 and 5.10. Up to $T = 5 \cdot 10^{-3}$, the species X_3 decomposes into X_2 and X_4 . Now the reaction of X_2 and X_1 starts and the profile of X_4 grows ($T = 25$ and $T = 250$). Since the lifetime of X_3 is very short, the concentration of X_3 is very small. Once X_1 is fully converted into X_4 , the dissipation rate of the reactions drops down and the slow diffusion and the backward reactions start (around at $t = 2.5 \cdot 10^{13}$). Finally the system converges to the thermodynamic equilibrium solution (last column).

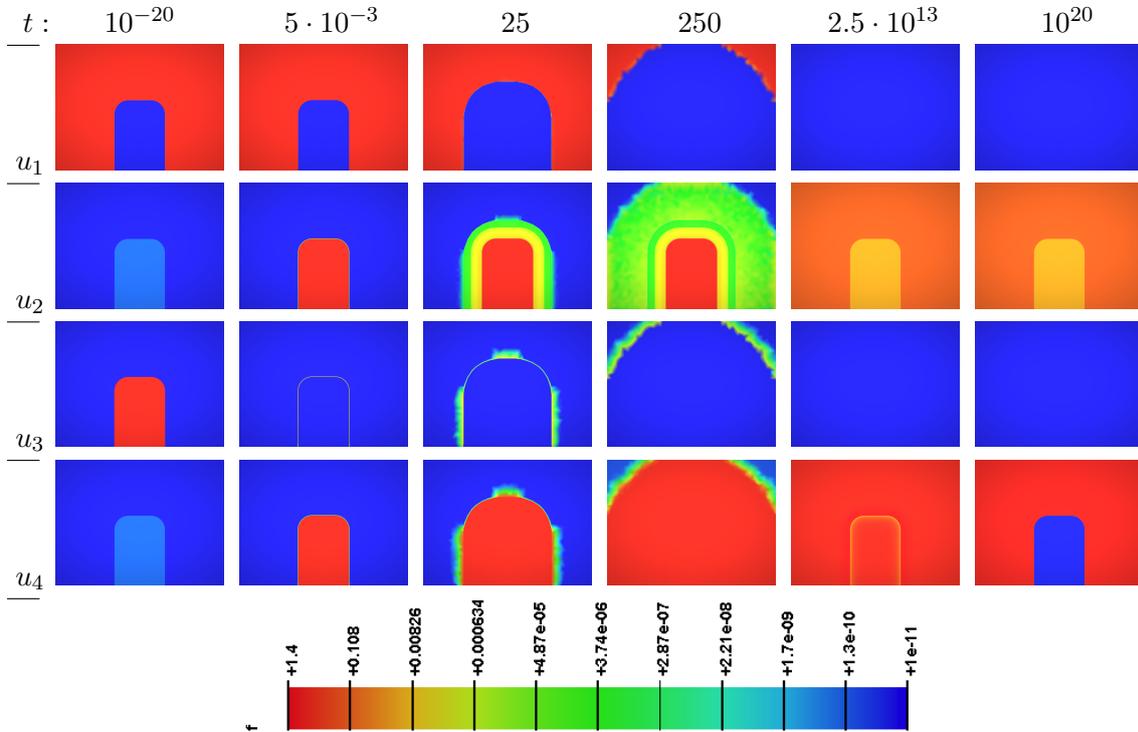


Figure 5.11: Concentrations of the four species at different times on discretization \mathcal{D}_3

The relative free energy $\Psi_h(u_h)$ decays to 10^{-10} and the chemical activities at the final simulation time $t = 10^{20}$ are close to the calculated thermodynamic equilibrium (5.4). In Figure 5.12 the relative difference in space is depicted and in Table 5.7 the maximal value of the differences relative to (5.4) is tabulated. The relative errors are close to the double precision machine epsilon ($2.22 \cdot 10^{-16}$).

Summarizing, we emphasize that our method can reach – up to machine epsilon – the thermodynamic equilibrium (approx. at $T = 10^{20}$) and shows the expected monotonous (and exponential) decay of the free energy and the expected non-negativity of the dissipation rate, see Lemma 3.6.1. Moreover the relative error of the two invariants (5.1.1) and (5.1.2) is of the order of machine epsilon over all orders of magnitude in time, see Table 5.5.

The Newton method shows almost quadratic convergence rates and passes the convergence criteria in 3 or 4 steps. During the marching of the reaction front the Newton iteration needs 5 or 6 steps, in order to pass the convergence criteria, see Table 5.5. Since the reaction front cannot be accurately predicted by the linear extrapolation predictor, the initial guess of the Newton method is not in the region of quadratic convergence resulting into a larger number of necessary Newton steps. To overcome this problem, one would need a more accurate prediction of the reaction front, which would be possible to implement, but was not in the focus of this contribution.

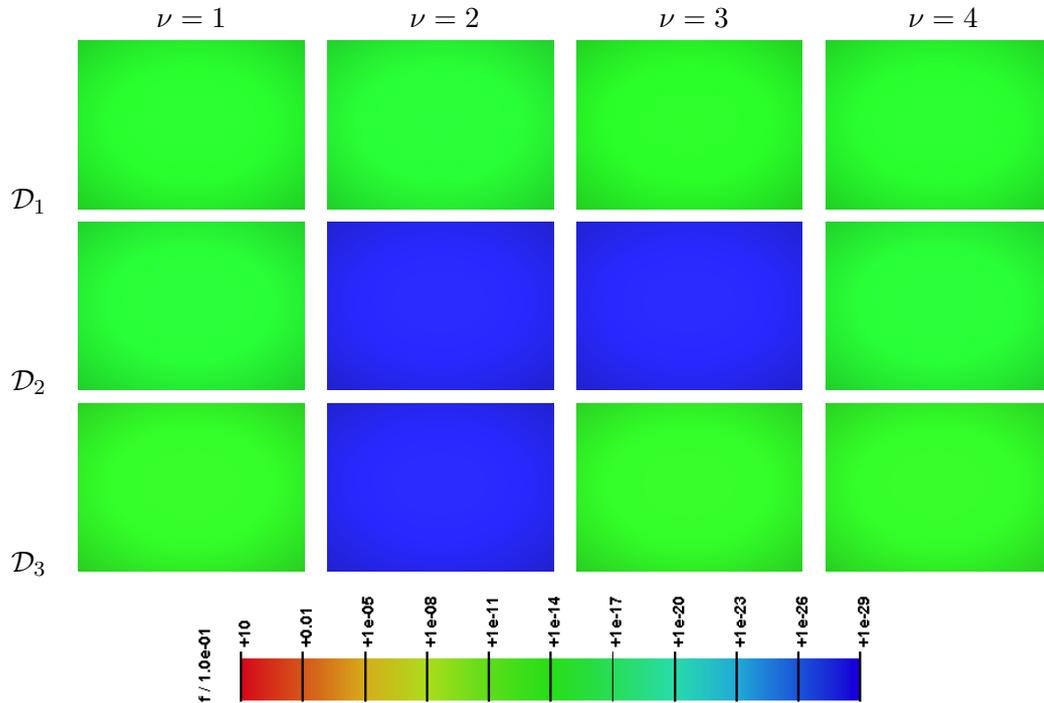


Figure 5.12: Relative distance of the activities $\text{err}_\nu = (a_{\nu,h}(T) - a_{\nu,h}^*)/a_{\nu,h}^*$, $\nu = 1, \dots, 4$, for different discretizations at time $T = 10^{20}$. Note the logarithmic scale and see also Table 5.7 for the maximum value

$\ \text{err}_\nu\ _{L^\infty(\Omega)}$	\mathcal{D}_1	\mathcal{D}_2	\mathcal{D}_3
$\nu = 1$	$6.761 \cdot 10^{-16}$	$3.381 \cdot 10^{-16}$	$2.366 \cdot 10^{-15}$
$\nu = 2$	$3.624 \cdot 10^{-16}$	$1.812 \cdot 10^{-16}$	0
$\nu = 3$	$1.104 \cdot 10^{-15}$	0	$2.345 \cdot 10^{-15}$
$\nu = 4$	$6.761 \cdot 10^{-16}$	$3.381 \cdot 10^{-16}$	$2.366 \cdot 10^{-15}$

Table 5.7: The relative error $\|\text{err}_\nu\|_{L^\infty(\Omega)} = \left\| (a_{\nu,h}(T) - a_{\nu,h}^*) / a_{\nu,h}^* \right\|_{L^\infty(\Omega)}$, $\nu = 1, \dots, 4$, for different discretizations at time $T = 10^{20}$

5.5 Nested Circles – Order of convergence

In the former example we saw that the method is applicable to real world problems coming from chemical engineering. Now we are more interested in the convergence behavior of the method. For this purpose we assume a circular domain consisting of three different materials. The values are comparable to Table 5.4. The parameters of the red material are equal to the parameters of the red material in Table 5.4. In the blue region the diffusivities are smaller than in the blue material in Table 5.4. Only the diffusivity of X_3 is larger. In the green region the forward and backward reaction from X_1 and X_2 to X_3 is assumed to be fast and the second one is slow compared to the first one. At the initial state the reactions are in equilibrium. Due to diffusion on the material interface, this equilibrium state will be violated and the reactions start. In Table 5.8, we collect the parameters of the blue and green material. The parameters in the red region are equal to those of Table 5.4.

	green region	blue region
$(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$	(1, 1, 1, 1)	$(10^{-20}, 1, 10^{-11}, 1)$
(U_1, U_2, U_3, U_4)	(1, 1, 1, 1)	$(10^{-20}, 10^{-16}, 1, 10^{-27})$
(D_1, D_{20}, D_3, D_4)	$(10^{-10}, 0.02, 10^{-5}, 10^{-12})$	$(10^{-10}, 250, 10^{-10}, 10^{-12})$
$(\varphi_3, \varphi_1, \varphi_2)$	(10, 0.01, 0.95)	(0, 0.65, 0.95)
(K_1, K_2)	$(10^{10}, 10)$	$(10^{-21}, 0.1)$

Table 5.8: Material parameters for the three different regions. Note the large differences in the diffusion and reaction coefficients

5.5.1 Uniformly refined mesh

In a first step we consider a uniformly refined mesh, see Figure 5.13. The meshes are created with Triangle [120] using a constraint on the triangle area. The reference mesh \mathcal{M}_5 consists of 43973 nodes. The different numbers of degrees of freedom are listed in Table 5.9. In addition to the implicit

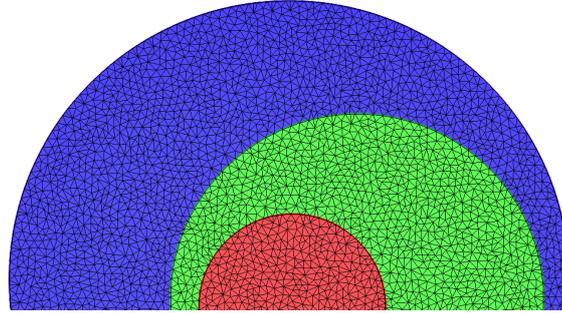


Figure 5.13: One half of the Delaunay mesh of discretization \mathcal{D}_2 . The different materials are represented by different colors

Mesh.	# nodes	# tri.	<i>ratio</i>	$\ \Psi_{h_i}\ _{L^2(S)}$	$\ D_{R,h_i}\ _{L^2(S)}$	$\ D_{D,h_i}\ _{L^2(S)}$
\mathcal{M}_1	465	861	0.01	$1.835 \cdot 10^8$	$3.082 \cdot 10^7$	$2.474 \cdot 10^2$
\mathcal{M}_2	2264	4360	0.05	$1.886 \cdot 10^8$	$3.294 \cdot 10^7$	$2.238 \cdot 10^2$
\mathcal{M}_3	4454	8687	0.1	$1.890 \cdot 10^8$	$3.339 \cdot 10^7$	$2.068 \cdot 10^2$
\mathcal{M}_4	22085	43666	0.5	$1.893 \cdot 10^8$	$3.397 \cdot 10^7$	$1.772 \cdot 10^2$
\mathcal{M}_5	43973	87233	1	$1.894 \cdot 10^8$	$3.414 \cdot 10^7$	$1.593 \cdot 10^2$

Table 5.9: Characterization of the different meshes and L^2 -norm of the initial free energy and dissipation. The quantity *ratio* is defined by $ratio = \#\mathcal{M}_i/\#\mathcal{M}_5$

Euler method (1-step), we use a 2-step backward differentiation formula (BDF-2) for time integration, see [71]. In case of time step rejections we return to the implicit Euler method (BDF-1) for the next time step and then we try to increase the order once again. Note that the BDF-2 method preserves the invariance property of the solutions (see Lemma 3.4.1), as long as the first step fulfills this property.

The total numbers of time steps that are needed by the algorithm in order to obtain the solution on the interval $S = [0, 10^{20}]$ are summarized in Table 5.10. The number of first and second order steps are collected in Table 5.10, too.

Evolution of the time step In Figure 5.14, we plot the growth of the time step over the time. The time step grows linearly in the log-log scale, with the exception where a new timescale is reached. The first deviation from linear growth in log-log-scale is due to the fast reaction in the blue region. During the movement of the front of X_4 , the time step is limited by the number of Newton steps.

The second deviation is less pronounced on the coarse mesh, due to a faster discrete diffusion on coarse meshes. The next deviation is due to the growth of the concentration on the boundary between the green and red region. The finer mesh resolves more effects of the solution, therefore the necessary time steps for \mathcal{D}_5 are smaller compared to the time steps of \mathcal{D}_1 .

Discret.	Mesh	$\#t_\delta$ tot.(rej.)	T1	$ratio\ sq$	$\#t_\delta$ tot.(rej.)	$ratio\ sq$	2/1-order
\mathcal{D}_1	\mathcal{M}_1	834 (4)		0.027	1528 (146)	0.19	1357;153
\mathcal{D}_2	\mathcal{M}_2	1104 (5)		0.047	3701 (536)	1.11	3183; 518
\mathcal{D}_3	\mathcal{M}_3	1693 (4)		0.11	4926 (768)	1.97	3927;999
\mathcal{D}_4	\mathcal{M}_4	3513 (3)		0.48	4950 (729)	1.99	3858;1092
\mathcal{D}_5	\mathcal{M}_5	5084 (3)		1	3507 (265)	1	2429;1078

Table 5.10: Characterization of the time discretizations. Total number of time steps t_δ (without rejections) that are needed in order to solve the problem on $S = [0, 10^{20}]$ and the squared ratio for implicit Euler (T1) and BDF-2 (T2). The quantity $ratio\ sq$ is defined by $(\#t_{\delta, \mathcal{M}_i} / \#t_{\delta, \mathcal{M}_5})^2$. Moreover the number of BDF-2 steps and BDF-1 steps during time integration with the 2-stage BDF method

During the front movement, the BDF-2 method spends much more time steps and the method switches back to the implicit Euler method. When the solutions evolve smoothly, the BDF-2 method works well. Especially during the growth of the concentration on the boundary between the green and red region, the deviation from linear growth in log-log-scale of the BDF-2 method is less pronounced compared to the implicit Euler method.

Due to the large number of rejections and time steps of the BDF-2 method, the computational cost is higher than for the implicit Euler method. On \mathcal{M}_5 , the BDF-2 method needs less time steps than the implicit Euler, since the spatial discretization error is sufficiently small. During the front movement and near the thermodynamic equilibrium the BDF-2 method switches back to the implicit Euler.

It is interesting to note that the ratio of the mesh nodes and the squared ratio of the executed time steps by the implicit Euler method are close together. Due to the large number of rejection on coarse meshes the ratio for the BDF-2 method is greater than one and not meaningful in comparison to the ratio of the mesh nodes.

Time integration using implicit Euler First we consider the implicit Euler method for time integration. In Figure 5.17 we depict the evolution of concentration of the four species on the mesh \mathcal{M}_1 and \mathcal{M}_5 at various times. Every column in Figure 5.17 corresponds to a vertical straight line in the relative free energy plots in Figure 5.15.

In a first step the system forms X_2 in the blue region. Due to the strongly spatially varying diffusivity of X_4 the concentration creates a layer on the boundary of two materials. Then the movement of X_4 into the red region starts. On the coarsest mesh the diffusion is larger, therefore the front moves faster and the dissipation rate disappears earlier in time. After the species X_4 fully fills up the red region, the layer on the boundary of the red and green region starts to grow, see the profile of u_{4,h_1} and u_{4,h_5} at time $t = 5.05 \cdot 10^6$. The growth is driven by reaction and diffusion and will become

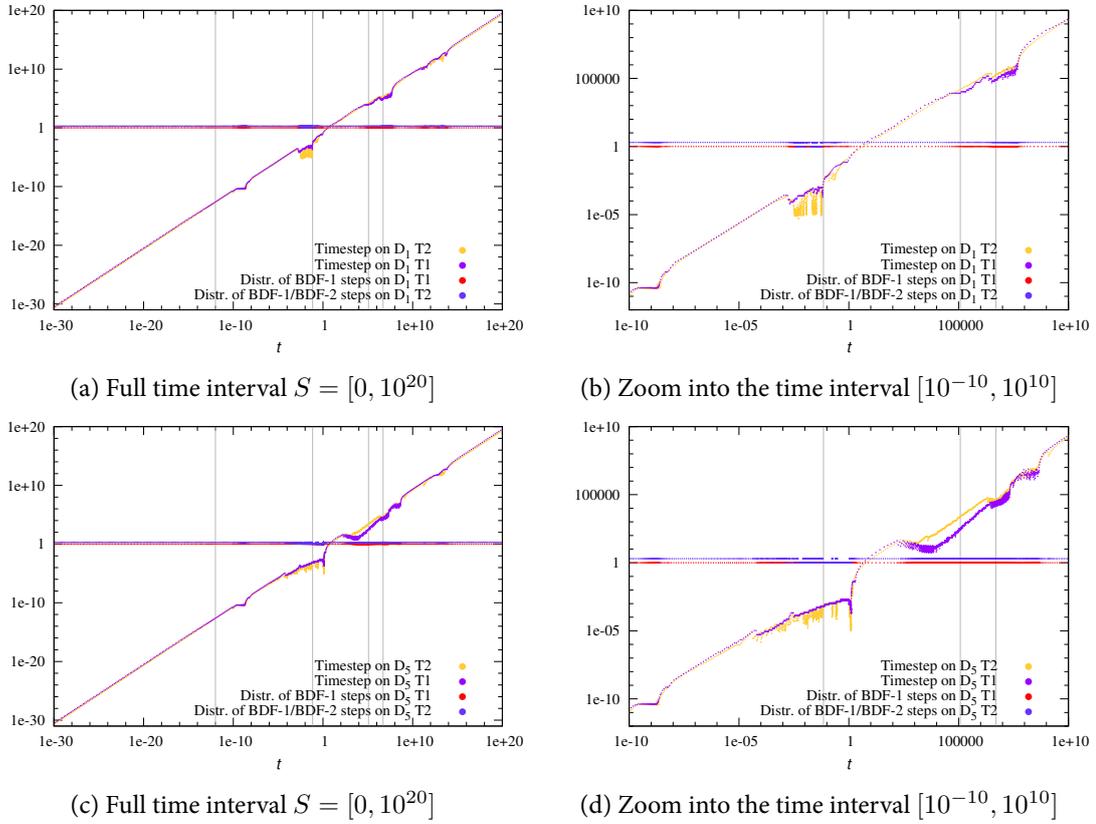


Figure 5.14: Evolution of the time step and the distribution of first and second order steps on a mesh \mathcal{M}_1 with BDF-1 and BDF-2 method. Every gray straight line represents one column in Figure 5.17

flatter on the area where the distance of the red and blue region is minimal. The species X_4 cannot penetrate into the red region due to its very small diffusivity. After the system is evolving on the slow time scale (around $t = 10^{10}$), the barrier dissipates, the concentration of species X_4 raises in the red region, and the system approximates the thermodynamic equilibrium (approximately at $t = 10^{20}$).

The evolution of the relative free energy functional Ψ_h is plotted in Figure 5.15a. The evolution of the relative free energy and the dissipation reflects the different timescales of the system.

Since the initial value is differently resolved by the meshes, the relative free energy varies on different meshes at the initial time. After the first fast reaction the curve of the relative free energy steps down. The reaction timescale is faster than all other diffusion timescales and therefore the curves nearly coincide. Coarse grids add numeric diffusion, hence the front movement is slower on fine grids and the energy curves decay accordingly.

Then further diffusion is blocked by the already described layer and this causes the synchronization of all free energy curves. The fill up of the red region by X_4 is diffusion dominated and depends on the refinement of the mesh. Therefore we see a time shift in the free energy curves. Finally the

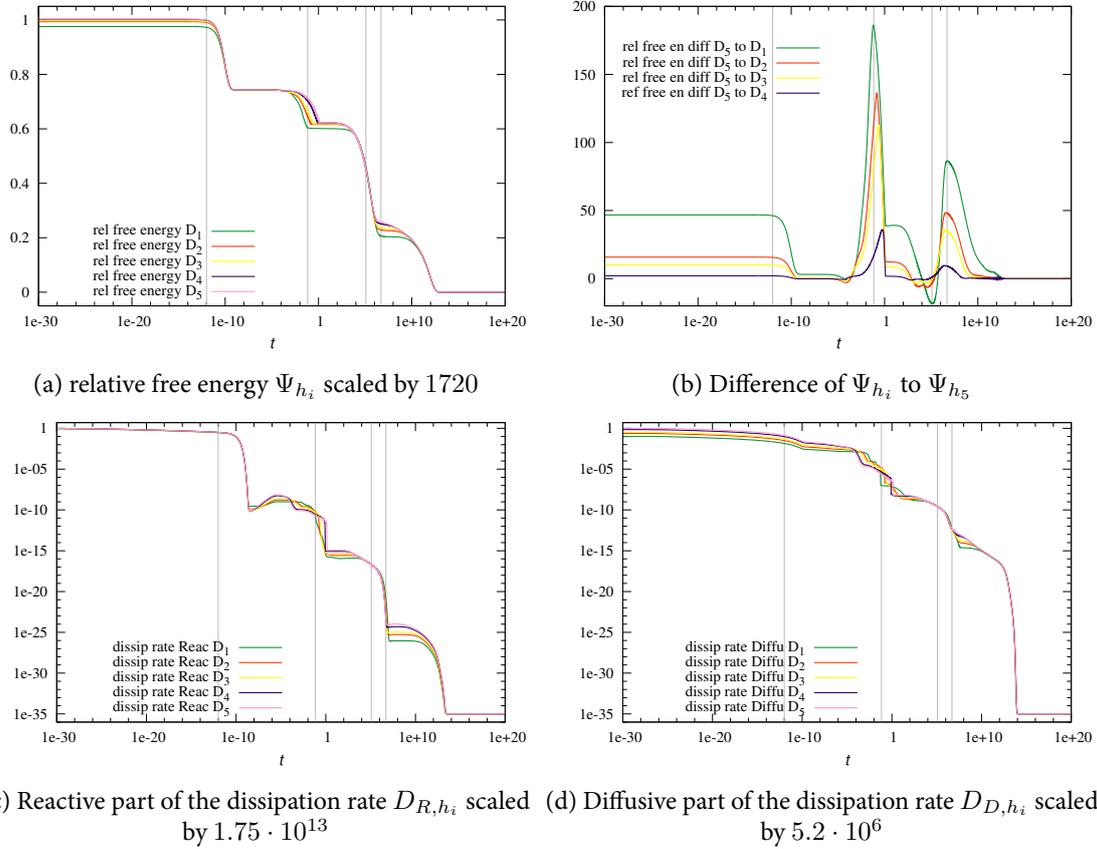


Figure 5.15: Evolution of the relative free energy functional, both parts of the dissipation rate for different discretization levels \mathcal{D}_i , $i = 1, \dots, 5$. Moreover the time evolution of $\Psi_{h_j} - \Psi_{h_5}$, $j = 1, \dots, 4$ is depicted. The values in the Figures 5.15a, 5.15d and 5.15c are scaled by 1720.0, $1.75 \cdot 10^{13}$ and $5.2 \cdot 10^6$, respectively. In Figure 5.15d and 5.15c we set the numerical zero to 10^{-35} . Every gray straight line represents one column in Figure 5.17

system approximates the thermodynamic equilibrium, which depends on the initial species masses, hence weakly on the spatial discretization. This is reflected in nearly coinciding energy curves.

In Figure 5.15b we plot the evolution of the absolute difference of the relative free energy curve associated to \mathcal{D}_i , $i = 1, \dots, 4$, and the relative free energy curve belonging to \mathcal{D}_5 .

The largest deviations occur at the end of the front movement. Moreover the maxima are slightly shifted, due to the faster evolution on the coarser meshes. The second peak represents the growing of X_4 on the boundary of the red and green region. The height of the peak depends on the resolution of the layer by the meshes. The growth of the layer is reaction dominated and the maxima are only slightly shifted. The fill up of the red region with X_4 is diffusion dominated. Therefore the curves of the free energy differences descent at different speeds into the thermodynamic equilibrium.

Convergence order In order to investigate the convergence behavior of the method w.r.t. the spatial error, we fix the temporal discretization for every mesh. This means, for every mesh \mathcal{M}_i , $i = 1, \dots, 4$, we run the simulations again with the time steps used in the simulation of \mathcal{D}_5 . Due to the shifts in the time scales on different meshes we allow the method to refine the given time steps in case of rejections. Therefore the temporal discretization is assumed to be comparable for every considered mesh.

In Table 5.11 the relative $L^2(S)$ -error of the relative free energy and the dissipation rate is presented. Since the continuous solution of the reaction-diffusion system is not known, a reference solution for the finest discretization \mathcal{D}_5 is computed, and errors are measured against this discrete solution. From these errors convergence rates are estimated by fitting. The error norms are given by

$$\begin{aligned} \left\| \tilde{\Psi}_{h_{ij}} \right\|_{L^2(S)} &= \left\| \Psi_{h_i}(u_{h_i}) - \Psi_{h_j}(u_{h_j}) \right\|_{L^2(S)} \left\| \Psi_{h_j}(u_{h_j}) \right\|_{L^2(S)}^{-1}, \\ \left\| \tilde{D}_{h_{ij}} \right\|_{L^2(S)} &= \left\| D_{h_i}(u_{h_i}) - D_{h_j}(u_{h_j}) \right\|_{L^2(S)} \left\| D_{h_j}(u_{h_j}) \right\|_{L^2(S)}^{-1}. \end{aligned} \quad (5.5.3)$$

Defining $h = \frac{1}{\sqrt{\#\mathcal{P}}}$, from Table 5.11 can be recognized that the $L^2(S)$ -errors of the relative free energy and the dissipation rate decay as h^2 and $h^{1.5}$. Therefore, we observe that our scheme allows by quadrupling the number of mesh points to divide the error of the discrete relative free energy by a factor of four. In order to get a better understanding of the influence of the time integration method

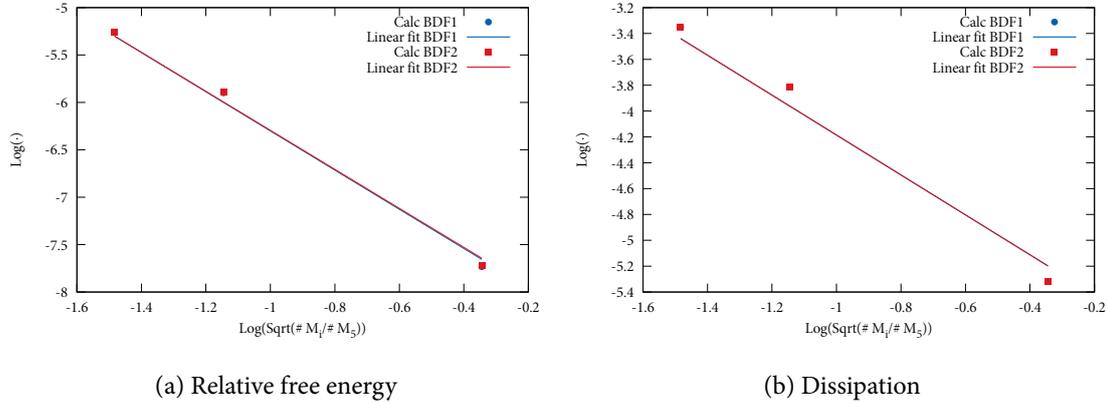


Figure 5.16: Linear fit of the values of Table 5.11 and 5.12 On both x-axis the value $\ln \sqrt{\#\mathcal{M}_i/\#\mathcal{M}_5}$ is plotted. On the y-axis the value $\ln \left\| \tilde{\Psi}_{h_{i5}} \right\|_{L^2(S)}$ (left) and $\ln \left\| \tilde{D}_{h_{i5}} \right\|_{L^2(S)}$ (right) is plotted. The blue points and lines are covered by the red ones. The slope of the linear function is -2.06 in Figure 5.16a and -1.54 in Figure 5.16b

on (5.5.3), we compute the relative $L^2(S)$ -error of the relative free energy and the dissipation rate. As before, we repeat all calculations with the time steps used on \mathcal{M}_5 and allow refinements in case

	$\#\mathcal{P}$	$\#t_\delta$ tot.	$\ \tilde{\Psi}_{h_{i5}}\ _{L^2(S)}$	$\ \tilde{D}_{h_{i5}}\ _{L^2(S)}$
\mathcal{D}_1	465	5136	$2.3871 \cdot 10^{-2}$	$9.7167 \cdot 10^{-2}$
\mathcal{D}_2	2264	5092	$5.2070 \cdot 10^{-3}$	$3.5054 \cdot 10^{-2}$
\mathcal{D}_3	4454	5095	$2.7435 \cdot 10^{-3}$	$2.2046 \cdot 10^{-2}$
\mathcal{D}_4	22085	5098	$4.3777 \cdot 10^{-4}$	$4.8919 \cdot 10^{-3}$
h^p			$p = 2.06$	$p = 1.5$

Table 5.11: $L^2(S)$ -norm of the relative free energy Ψ_{h_i} and the dissipation rate D_{h_i} of discretization \mathcal{D}_i , $i = 1, \dots, 4$, related to discretization \mathcal{D}_5 and scaled by the $L^2(S)$ -norm of the relative free energy and dissipation rate of \mathcal{D}_5

of rejections. In Figure 5.16 we see that the measured values of the BDF-1 and BDF-2 method nearly coincide, see also Tables 5.11 and 5.12. This means that we see the spatial convergence behavior of the method.

	$\#\mathcal{P}$	$\#t_\delta$ tot.	$\ \tilde{\Psi}_{h_{i5}}\ _{L^2(S)}$	$\ \tilde{D}_{h_{i5}}\ _{L^2(S)}$
\mathcal{D}_1	465	5187	$2.3704 \cdot 10^{-2}$	$9.7145 \cdot 10^{-2}$
\mathcal{D}_2	2264	5756	$5.2174 \cdot 10^{-3}$	$3.5047 \cdot 10^{-2}$
\mathcal{D}_3	4454	6070	$2.7629 \cdot 10^{-3}$	$2.2041 \cdot 10^{-2}$
\mathcal{D}_4	22085	6144	$4.4232 \cdot 10^{-4}$	$4.8908 \cdot 10^{-3}$
h^p			$p = 2.06$	$p = 1.5$

Table 5.12: $L^2(S)$ -norm of the relative free energy Ψ_{h_i} and the dissipation rate D_{h_i} on the meshes \mathcal{M}_i , $i = 1, \dots, 4$, related to mesh \mathcal{M}_5 and scaled by the $L^2(S)$ -norm of the relative free energy and dissipation rate of \mathcal{M}_5 . Time integration was done using the BDF-2 method

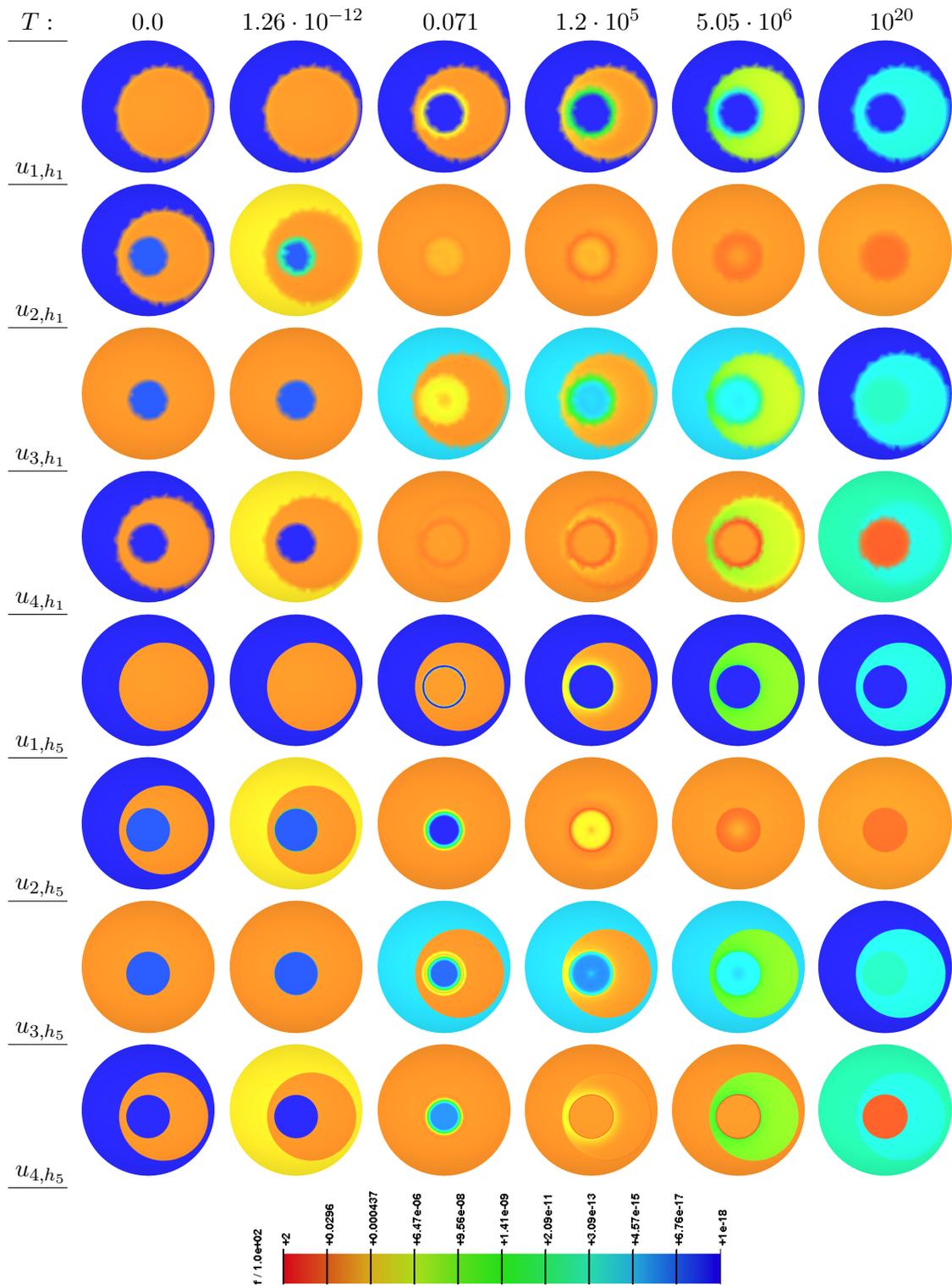


Figure 5.17: Concentrations of the four species at different times for discretizations \mathcal{D}_1 and \mathcal{D}_5 . Note the similarities of the coarse and fine solutions, e.g., at $t = 1.26 \cdot 10^{-12}$ and $t = 5.05 \cdot 10^6$. Be aware that differences are mainly caused by faster diffusion on coarse grids

5.5.2 Aligned mesh region

In this example we replace the unstructured triangulation in the red region by an aligned one, in order to get a better resolved moving front. The resulting discretizations are tabulated in Table 5.14 and one half of the mesh is shown in Figure 5.18. From the degrees of freedom, the discretizations

Mesh.	# nodes	# tri.	ratio	$\ \Psi_{h_i}\ _{L^2(S)}$	$\ D_{R,h_i}\ _{L^2(S)}$	$\ D_{D,h_i}\ _{L^2(S)}$
\mathcal{M}'_1	463	854	0.02	$1.857 \cdot 10^8$	$3.092 \cdot 10^7$	$2.729 \cdot 10^2$
\mathcal{M}'_2	1065	2034	0.48	$1.874 \cdot 10^8$	$3.197 \cdot 10^7$	$2.979 \cdot 10^2$
\mathcal{M}'_3	2235	4314	0.10	$1.886 \cdot 10^8$	$3.297 \cdot 10^7$	$2.378 \cdot 10^2$
\mathcal{M}'_4	2924	5691	0.13	$1.887 \cdot 10^8$	$3.299 \cdot 10^7$	$2.825 \cdot 10^2$
\mathcal{M}'_5	4753	9293	0.21	$1.889 \cdot 10^8$	$3.338 \cdot 10^7$	$2.333 \cdot 10^2$
\mathcal{M}'_6	7040	13862	0.31	$1.890 \cdot 10^8$	$3.340 \cdot 10^7$	$2.747 \cdot 10^2$
\mathcal{M}'_7	22423	44354	1.0	$1.881 \cdot 10^8$	$3.399 \cdot 10^7$	$1.873 \cdot 10^2$

Table 5.13: Characterization of the different meshes and L^2 -norm of the initial free energy and dissipation

Discretization	Mesh	$\#t_\delta$ tot.(rej.)	ratio sq
\mathcal{D}'_1	\mathcal{M}'_1	1046 (5)	0.025
\mathcal{D}'_2	\mathcal{M}'_2	932 (4)	0.020
\mathcal{D}'_3	\mathcal{M}'_3	996 (4)	0.023
\mathcal{D}'_4	\mathcal{M}'_4	1917 (4)	0.085
\mathcal{D}'_5	\mathcal{M}'_5	1742 (4)	0.070
\mathcal{D}'_6	\mathcal{M}'_6	3993 (3)	0.369
\mathcal{D}'_7	\mathcal{M}'_7	6575 (14)	1

Table 5.14: Characterization of the different discretizations. The numbers of time steps are needed by the scheme in order to solve the problem on $S = [0, 10^{20}]$ (Time integration done by implicit Euler)

\mathcal{D}'_1 and \mathcal{D}_1 , \mathcal{D}'_3 and \mathcal{D}_2 , \mathcal{D}'_5 and \mathcal{D}_3 , as well as \mathcal{D}'_7 and \mathcal{D}_4 are comparable. On some meshes the total number of time steps is slightly greater than on the unstructured mesh. In comparison to the ratio between space and time discretization, here the ratio of the number of time steps is smaller than the ratio the number of nodes, due to the irregular refinement of different regions. Since an analytical solution is not available, we compare the energy curves. In Figure 5.19a, 5.19d and 5.19c the time evolution of relative free energy and dissipation rate is plotted. The curves are comparable to the corresponding plots of the unstructured mesh. Only during the front movement, we see a more pronounced oscillating behavior in Figure 5.19d, due to the aligned structure of the mesh. The period

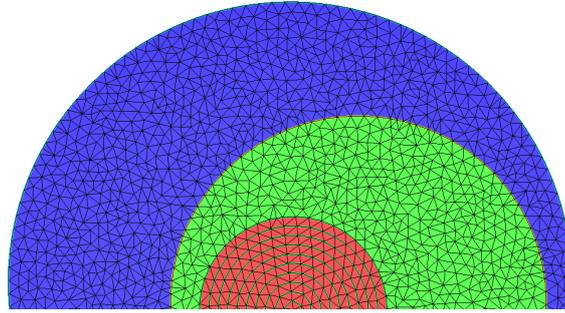


Figure 5.18: One half of the Delaunay mesh of discretization \mathcal{D}'_3 , which has approximately the same number of nodes as \mathcal{D}_2 . The different materials are represented by different colors

is correlated to the distance of two neighboring mesh lines. This behavior was already mentioned in Subsection 5.4.

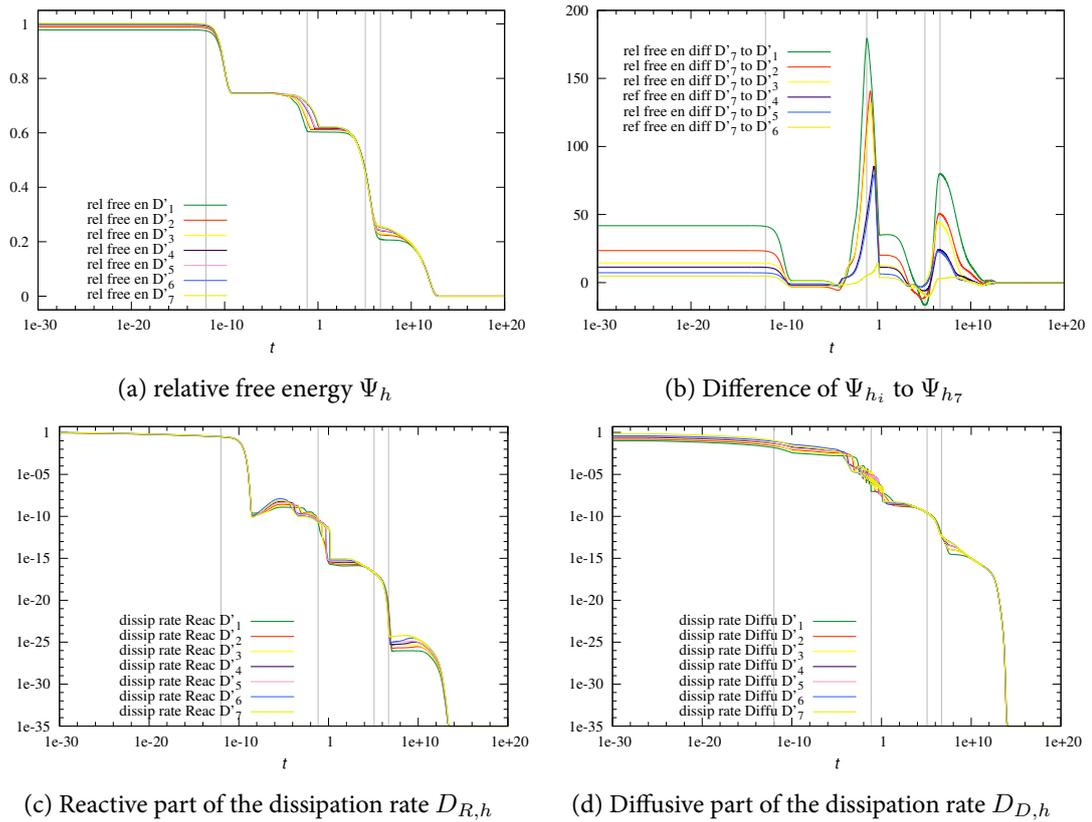


Figure 5.19: Evolution of the relative free energy functional, both parts of the dissipation rate for different discretization levels. Moreover the time evolution of $\Psi_{h_j} - \Psi_{h_7}$, $j = 1, \dots, 6$ is depicted. The values in the Figures 5.15a, 5.15d and 5.19c are scaled by 1720.0 , $1.75 \cdot 10^{13}$ and $5.2 \cdot 10^6$, respectively. In Figure 5.15d and 5.19c we set the numerical zero to 10^{-35} . Every gray straight line represents one column in Figure 5.17

At large times near the thermodynamic equilibrium, there is a strong deviation of the diffusive part of the dissipation rate, due to roundoff errors. The energy curves of the coarsest mesh are closer to the other curves, which could be a consequence of the aligned mesh. In Figure 5.19b the peaks are slightly smaller and less shifted than on the unstructured meshes. This means that the time scales on the different discretization are closer together. The $L^2(S)$ norm of the deviation is summarized in Table 5.15. The values are comparable to Table 5.11.

	$\#\mathcal{P}$	$\#t_\delta$ tot.	$\ \tilde{\Psi}_{h_{i7}}\ _{L^2(S)}$	$\ \tilde{D}_{h_{i7}}\ _{L^2(S)}$
\mathcal{D}'_1	463	6608	$2.1985 \cdot 10^{-2}$	$9.0280 \cdot 10^{-2}$
\mathcal{D}'_2	1065	6663	$1.1510 \cdot 10^{-2}$	$5.9365 \cdot 10^{-2}$
\mathcal{D}'_3	2235	6651	$2.7435 \cdot 10^{-3}$	$2.2046 \cdot 10^{-2}$
\mathcal{D}'_4	2924	6668	$3.9113 \cdot 10^{-3}$	$2.9382 \cdot 10^{-2}$
\mathcal{D}'_5	4753	6667	$2.1584 \cdot 10^{-3}$	$1.8019 \cdot 10^{-2}$
\mathcal{D}'_6	7040	6811	$1.8157 \cdot 10^{-3}$	$1.7470 \cdot 10^{-2}$
h^p			$p = 1.93$	$p = 1.30$

Table 5.15: $L^2(S)$ -norm of the relative free energy Ψ_{h_i} and the dissipation rate D_{h_i} of discretization \mathcal{D}'_i , $i = 1, \dots, 6$ related to discretization \mathcal{D}'_7 and scaled by the $L^2(S)$ -norm of the relative free energy and dissipation rate of \mathcal{D}'_7

Summary In this example the method can reach the thermodynamic equilibrium, after the decay of 14 magnitudes of the relative free energy. All timescales of the system can be resolved by the method, even if the front movement or the strong gradients in the direct neighborhood of interfaces are only roughly resolved — still on the very coarse mesh \mathcal{M}_1 the different energetic states of the system are present, of course with time shifts of their begin and end points (compare Figures 5.15, 5.17, 5.19d). In general, diffusion is faster on coarse grids due to the added numerical diffusion. Whenever a quasi steady state at a time scale is reached, or the system is 'waiting' before evolving on the next slower time scale, the free energy curves are synchronized. The BDF-2 method is cheaper for simulations where the spatial discretization error is sufficiently small. That means that the mesh must resolve more features of the solution and the evolution in time is smoother, compare Table 5.10. For both time integration methods BDF-1 and BDF-2 we see a convergence order of approximately 2 for the relative free energy and approximately 1.5 for the dissipation rate, which seems reasonable for a first order scheme, see [32].

5.6 Tiles

We consider a domain containing a point at which the boundaries of three materials meet. For simplicity, we choose the same parameters as in Subsection 5.5. Therefore we expect a similar behavior of the solution as in Subsection 5.5, except for singularities due to the triple point.

We adapt the Delaunay triangulation in the direction of the interface between different materials. On a domain $\Omega = (0, 3) \times (0, 3)$ we create a mesh consisting of 98262 triangles, 50668 nodes. The largest edge is $h_{max} = 0.12$ and the mesh is depicted in Figure 5.20.

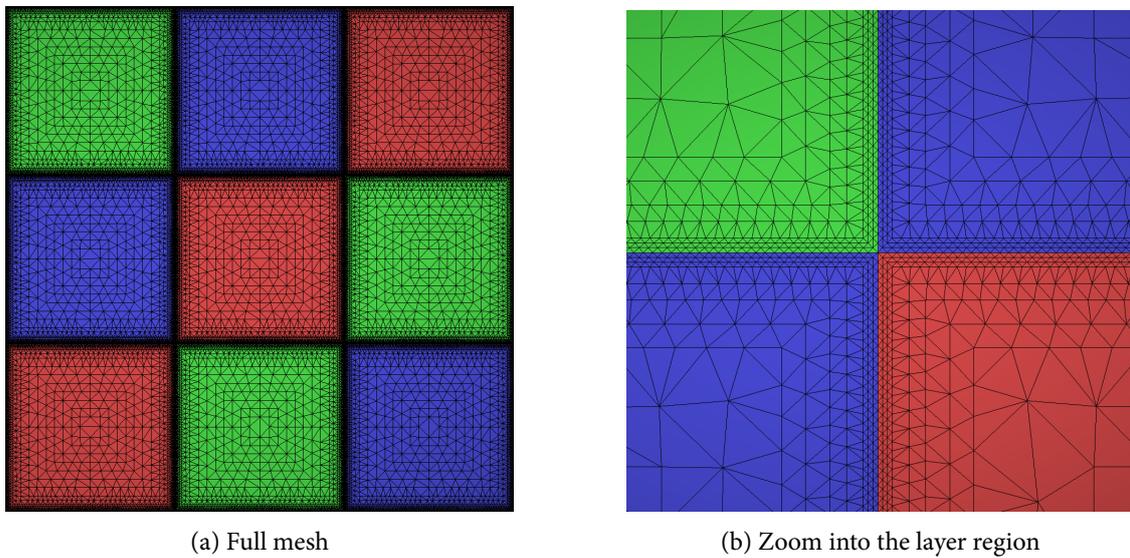


Figure 5.20: The used mesh with 50668 nodes and 98262 Delaunay triangles. Zoom into the adapted layer region

The evolution of the solution at different times is depicted in Figure 5.20. In a first step (in the interval $[0, 1 \cdot 10^{-9}]$ in Figure 5.23) the fast reaction acts in the blue region and creates X_4 and X_2 . Due to the space dependence of the reactions and the different diffusion coefficients, we see strong layers on the boundary of the blue and red material, see Figure 5.22a. On the triple points $(2, 2)$ (green, blue, red, blue) and $(3, 3)$ (red, green, blue, green) we see some small spikes. In the neighborhood of these points the reaction can act in three different ways. Then a bit earlier at $t = 1.214 \cdot 10^{-5}$ the front movement starts. In the red region X_1 will be fully converted into X_4 . Since the concentration of X_2 is high in all neighboring squares, the front stops in the tile $(2, 2)$ from every edge of the square and therefore the square in middle is filled up by X_4 at first, see columns three and four in Figure 5.21. During the front movement the layer in the concentration of X_4 on the boundary of the red and green region grows, see Figures 5.22b and 5.22c. The maximum of the layers is located only on the nodes lying on the boundary. Due to triple points the layers can be curved.

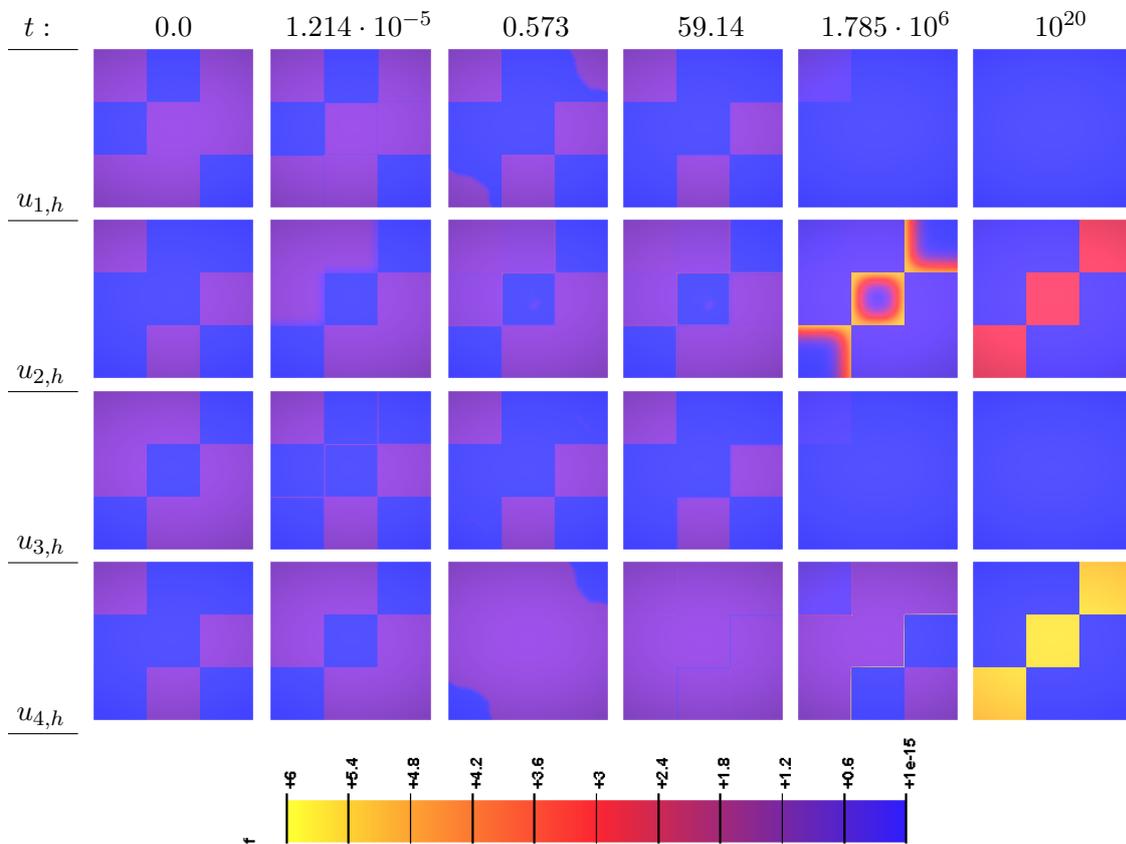


Figure 5.21: Concentrations of the four species at different times. Every column represents a straight gray line in Figure 5.23

After the red region is fully filled up by X_4 , X_1 and X_2 degrade in the blue and green regions. On the boundary of the red material, the concentration of X_4 grows. Especially at the triple point (2, 2) a peak of X_4 grows. At $t = 1.7153 \cdot 10^6$ the peak reaches its maximum of $13.6731 \cdot 10^4$. Almost all mass of X_4 is concentrated in this point, see Figure 5.22d. During this process it is essential that the mass is correctly divided. Otherwise it is not possible to reach the thermodynamic equilibrium.

Now the slow diffusion and reaction time scales come into play and the peak as well as the layers on the boundary will degrade and the concentration of the species X_4 raises in the red region. The system achieve the thermodynamic equilibrium, where almost all mass of the species X_1 and X_3 is transformed into X_2 and X_4 and is located in the red region.

Since the parameters are the same as in Subsection 5.5, the evolution of the relative free energy and the dissipation rate are comparable. Due to the better aligned mesh in the neighborhood of the layer the oscillating behavior of the dissipation rate is less pronounced. If the front reaches the

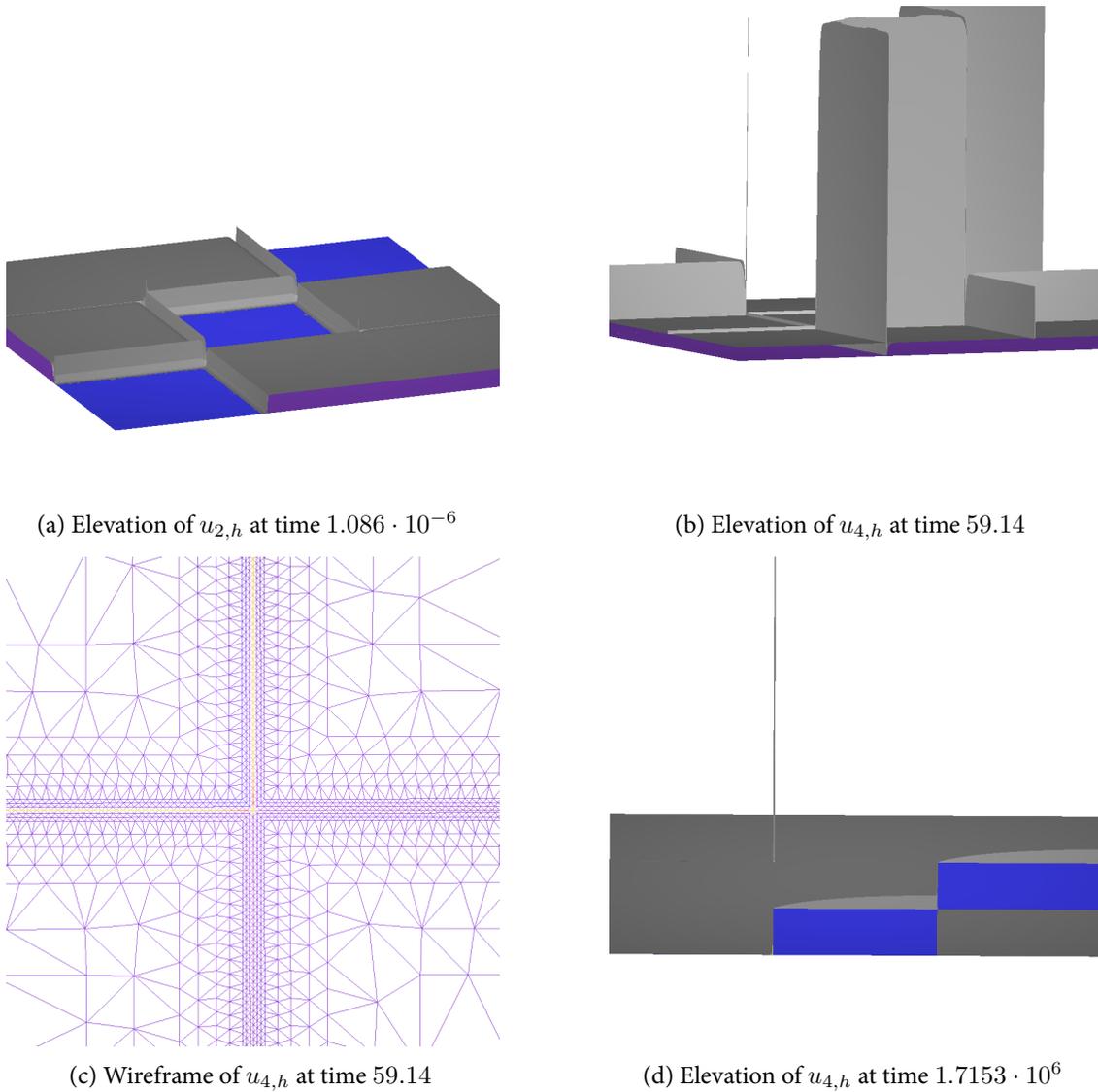


Figure 5.22: Details of the solution

coarse region of the squares, the oscillations are larger, since the front jumps from one grid line to the next, see Figure 5.23a.

The simulation needs 12778 time steps with 1688 rejections and 52215 Newton steps (in the mean four Newton steps per time step) in order to calculate the solution on the time interval $S = [0, 10^{20}]$ using the BDF-2 method. During the evolution the relative error of both invariants (5.1.1) and (5.1.2) is of the order of machine epsilon, see Figure 5.23c. Most of the time steps are executed during the front movement in the red material. In order to integrate from $[0, 2.18]$ we need 11902 steps. However integrating from $[2.18, 10^{20}]$ needs only 877 time steps. On the one hand, the reaction

front can not be accurately predicted by the linear extrapolation predictor, the initial guess of the Newton method is not in the region of quadratic convergence leading to a larger number of necessary Newton steps. On the other hand, the front movement needs a constant time step.

Similar to Subsection 5.5, most of the BDF-1 steps are done during the front movement, due to the spatially nonsmooth behavior of the solution. The BDF-2 method takes advantage in time intervals where the solution is spatially stationary or the solution evolves smoothly. These time intervals are the first fast reaction and the growth of the layers and peaks on the boundary.

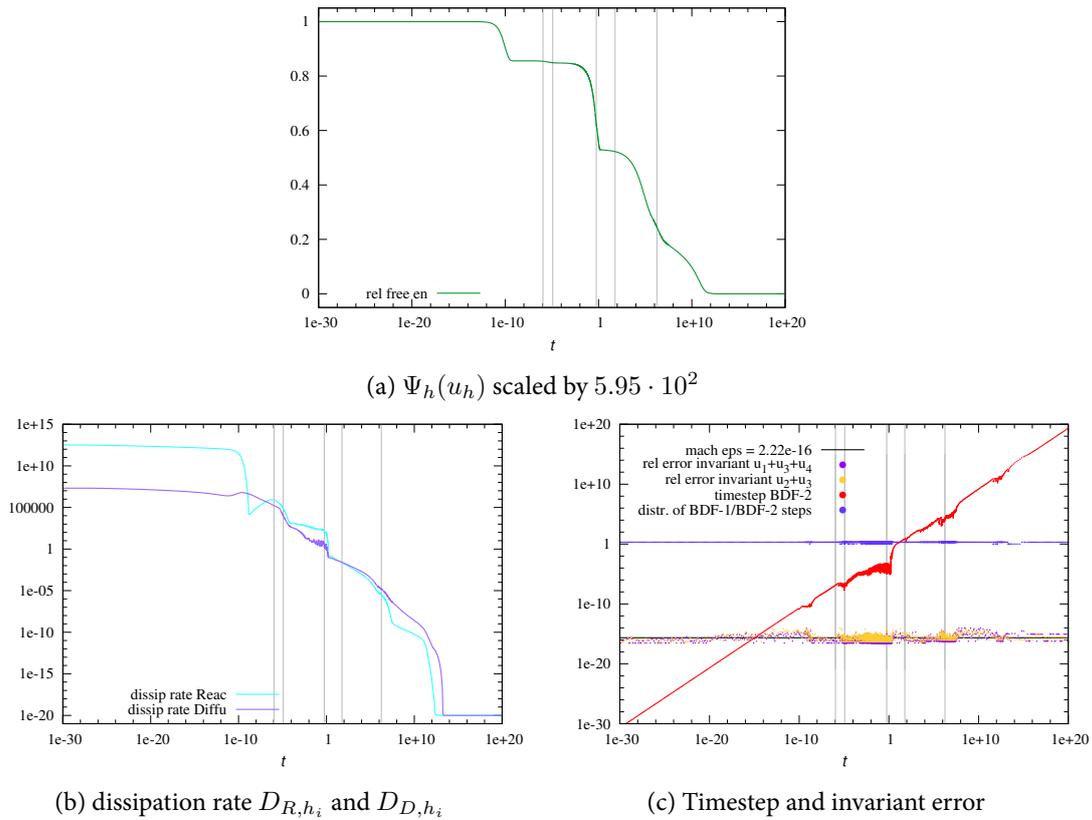


Figure 5.23: Figure 5.23a: Evolution of the relative free energy Ψ_{h_i} (Scaled by $5.95 \cdot 10^2$) and both parts of the dissipation rate D_{R,h_i} and D_{D,h_i} . (Numerical zero at 10^{-20}) Fig. 5.23c: Evolution of the time step over the time, distribution of BDF-1 and BDF-2 steps and the evolution of the relative invariant error. Every gray straight line represents one column in Figure 5.21

5.7 Daisy

In the example a non convex domain is used, which has the form of a daisy with a spanner inside. The large boundary and the interfaces are of special interest, because the initial data is chosen such

that the reactions are in equilibrium in the interior of the regions. In other words the equations are fulfilled up to boundary and interface layers which are very small in normal direction. Therefore the mesh is anisotropically refined along all interfaces, see Figure 5.24. The mesh consist of four different materials, 16812 triangles, and 8701 nodes. In every material we consider different reaction steps,

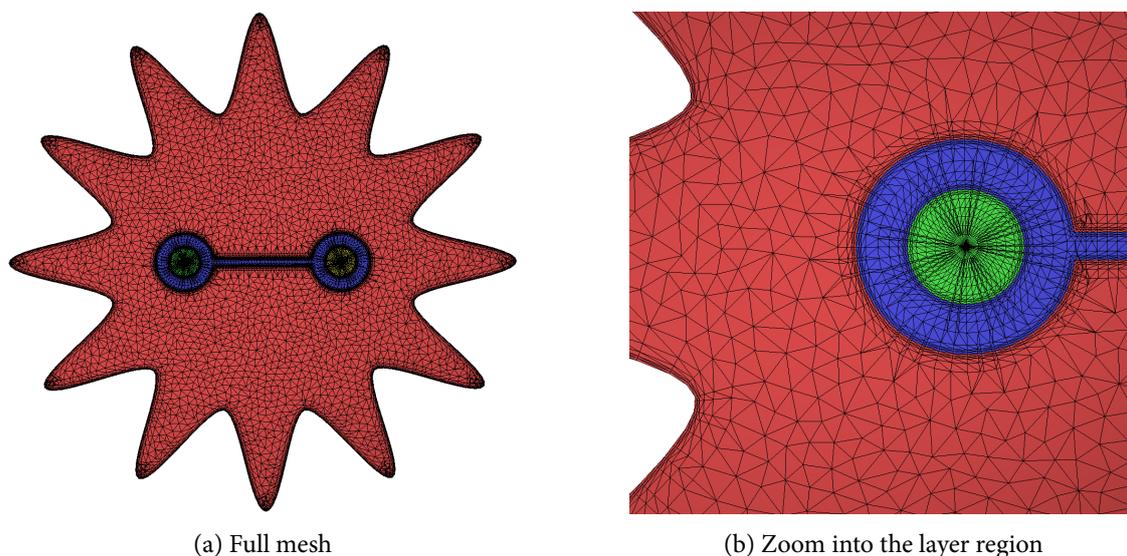
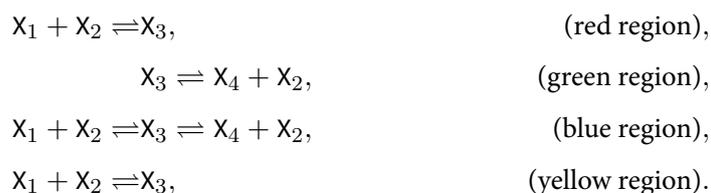


Figure 5.24: Daisy with 12 petals. The mesh has 8701 nodes and 16812 Delaunay triangles. The different materials are represented by different colors. Zoom into the spanner region

i.e. we assume



This means the first MMH reaction step is present in the red and yellow regions, the second MMH reaction step is active in the green region, and the full MMH mechanism only occurs in the blue region. The reaction step in the blue region is assumed to be fast, whereas all other reactions steps are act on a moderate or slow timescale. In the red and blue material, the diffusion coefficient of X_2 depends on the concentration of X_1 and X_4 . All other diffusion coefficients are independent of the state. Table 5.16 summarizes all parameters used in the calculation.

During time evolution, the concentration of X_2 decays strongly in the neighborhood of the domain boundary, see Figures 5.26a and 5.26b. In Figure 5.25 the evolution of the relative free energy, the dissipation rate and the conservation of the invariants is depicted. Every straight gray line represents

region	$(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$	(D_1, D_{20}, D_3, D_4)	$(\varphi_3, \varphi_1, \varphi_2)$	(K_1, K_2)
red	$(10^{-8}, 100, 10^{-11}, 10^{-8})$	$(10^{-2}, 1, 10^{-3}, 10^{-2})$	$(1, 0.65, 0.95)$	$(10^{-15}, 0)$
green	$(10^{-7}, 10^2, 10^{-3}, 0.1)$	$(10, 10^{-2}, 10^{-2}, 10^{-4})$	$(0, 0, 0)$	$(0, 10^{-5})$
blue	$(10^{-8}, 10^4, 10^{-3}, 10^{-15})$	$(10^3, 10^{-2}, 10^{-5}, 1)$	$(0.2, 0.65, 0.95)$	$(0.1, 10^{-4})$
yellow	$(10^{-4}, 1, 10^{-3}, 10^{-17})$	$(10^{-6}, 10^{-3}, 10, 10^{-5})$	$(0, 0, 0)$	$(10^{-9}, 0)$

Table 5.16: Material parameter for the three different regions. In the red and blue region the diffusion coefficient of x_2 depends on the state

one row in Figure 5.27. Due to the initial data, the reactions are locally in equilibrium and therefore

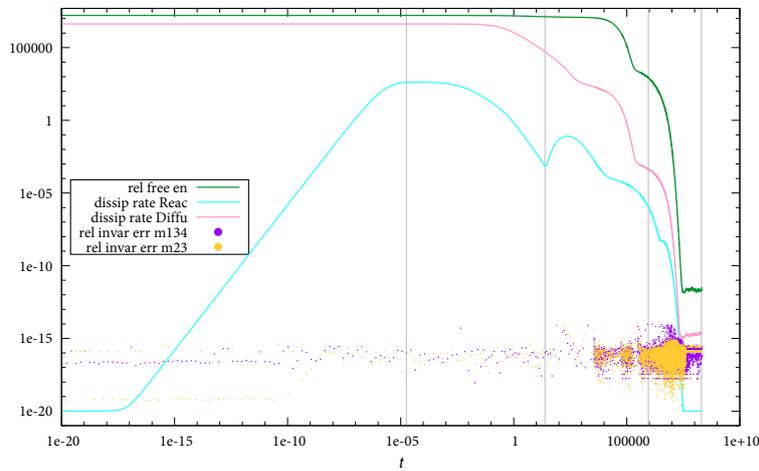


Figure 5.25: Evolution of the relative free energy $\Psi_h(u_h)$, the dissipation rate decomposed into the reaction part D_{R,h_i} and the diffusion part D_{D,h_i} , and the conservation of the invariants

the dissipation rate of the reactions is close to zero at the beginning, see Figure 5.25. The diffusion processes start immediately and trigger the reactions. Consequently the dissipation rate of the reactions must increase, compare Figure 5.25. In a first step, the system fills up the blue region with X_3 , see the second and the third column in Figure 5.27. Now, the dissipation rate of the reactions decays. Then the diffusion starts to smooth out the boundary layer of X_2 , see the fourth column in Figure 5.27. This causes the dissipation rate of the reactions to grow a second time. When this is finished the gradient of X_2 in the neighborhood of the domain boundary is smoothed out, see Figure 5.26c. Now the concentration of X_3 in the blue material decays. The species X_2 decays too and creates X_1 in the green and X_4 in the yellow region, see Figure 5.26d. Now the system proceeds fast to the thermodynamic equilibrium, see the fifth column in Figure 5.27 and the last gray straight line in Figure 5.25. The simulation ends at $t = 2 \cdot 10^8$. At this time the distance of the concentrations to the thermodynamic equilibrium solution is

$$\max_{\nu=1,\dots,4} \|u_{\nu,h} - u_{\nu,h}^*\|_{L^\infty} = 10^{-12}.$$

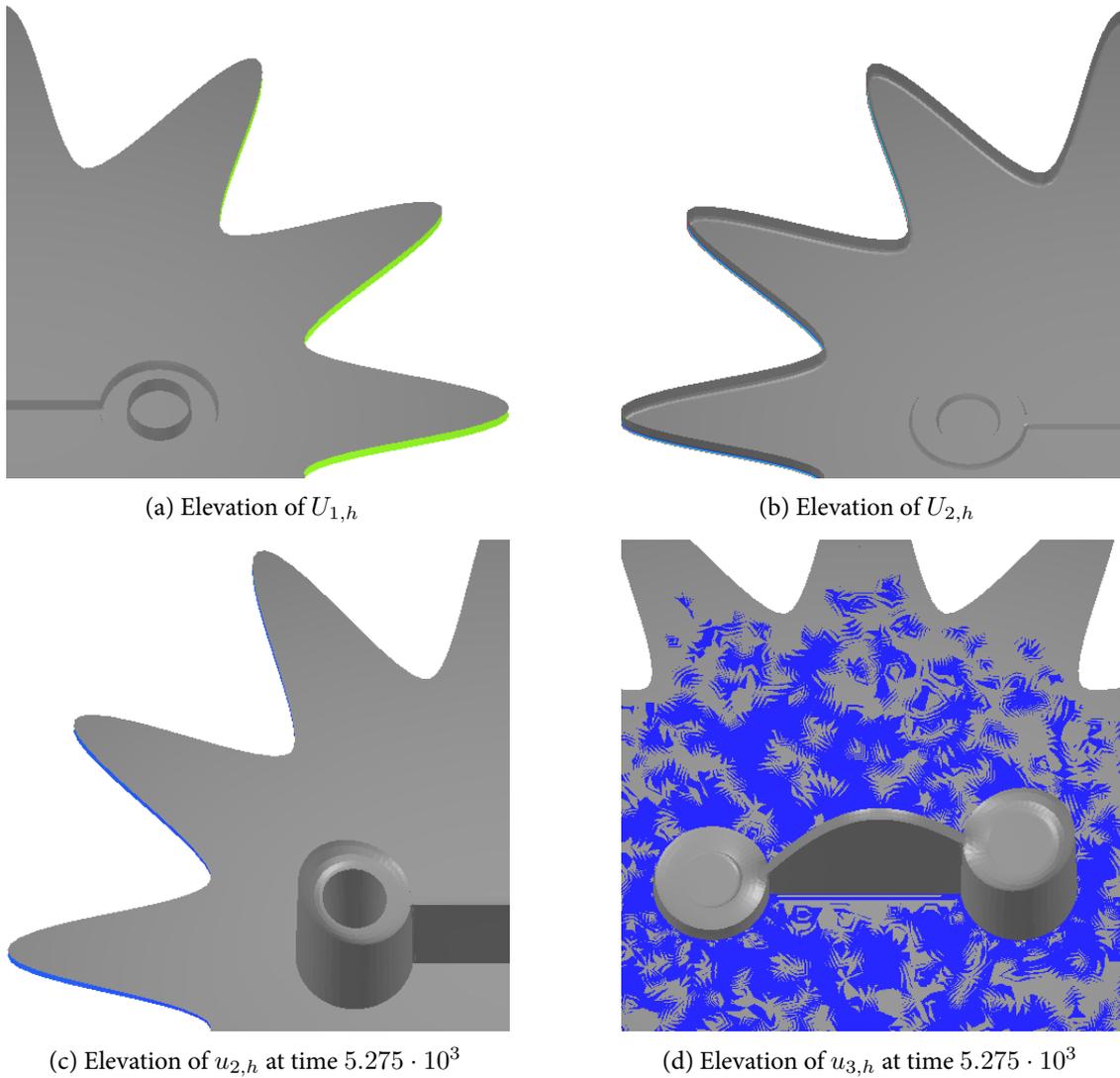


Figure 5.26: Details of the solution at different times. Note that the scale is different

Summary The method is able to handle the case with vanishing rate coefficients in parts of the domain. Subdomains without any reactions are possible, too. The equations in this subregion are regularized by the the global invariants. If in addition one diffusion coefficient is very small, the boundary of the subdomain acts, due to rounding, like an homogeneous Neumann boundary con-

dition for that equation. Hence the used regularization with one Dirichlet value is not sufficient (see Remark 4.1.3). With respect to the last Section 5.8, zero reaction rate constants in some subdomains can be viewed as a limit case of the full system.

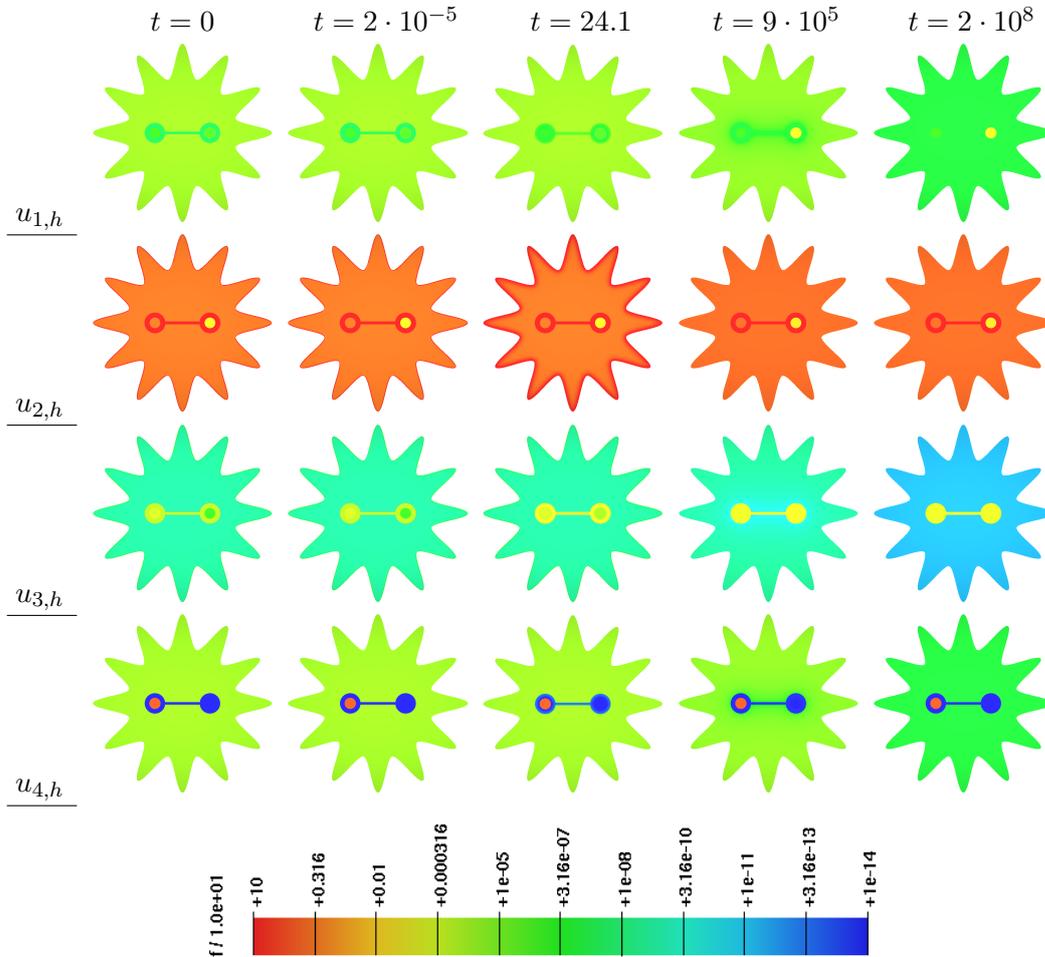


Figure 5.27: Concentrations of the four species at different times $t + \varepsilon$. Note the logarithmic scale of the pictures

5.8 Instantaneous reaction

In this section we are interested in the behavior of the code handling the instantaneous reaction limit. For this propose we chose the well studied example of [34, Section 6]. of a reversible dimerisation of *o*-phenylenedioxydimethylsilane (2,2-dimethyl-1,3,2-benzodioxasilole), which has been studied by hydrogen-1 Nuclear magnetic resonance spectroscopy, together with its limit system. The reaction

can be described quantitatively by a bimolecular 10-ring formulation reaction and a monomolecular back reaction, see [98]. Therefore the reaction step has the form



The set \mathcal{R} consist of the vectors $\alpha = (2, 0)$ and $\beta = (0, 1)$ and therefore

$$\mathcal{S} := \text{span} \{(2, -1)\}, \quad \mathcal{S}^\perp := \text{span} \{(1/2, 1)\}.$$

The fast reaction limit of special reaction-diffusion systems was considered in [11–14, 55, 113]. In order to prove convergence of the initial system to the limit system, the idea is to obtain a-priori bounds independent on the quantity which becomes large and than pass to the limit. Therefore the upper and lower bounds established in Section 3.8 and 3.10 are not suitable since they depend on the reaction rate constant of the reaction term. Whereas the L^1 -bounds obtained by the conservation of atoms (see Remark 3.6.2) stay valid.

We consider the one-dimensional domain $\Omega = (0, 0.1)$ and a finite interval $S = (0, T] \subset \mathbb{R}$, $T < \infty$. Assuming that the region is isolated (Neumann boundary conditions) and homogeneous and diffusion is modeled by Fick's law we obtain a system of two reaction diffusion equations

$$\left. \begin{aligned} u_1' - D_1 \Delta u_1 + 2\kappa (k_{1+} u_1^2 - k_{1-} u_2) &= 0 \\ u_2' - D_2 \Delta u_2 - \kappa (k_{1+} u_1^2 - k_{1-} u_2) &= 0 \end{aligned} \right\} \text{ in } S \times \Omega, \quad (5.8.5)$$

$$\mathbf{n} \cdot \nabla u_1 = \mathbf{n} \cdot \nabla u_2 = 0 \quad \text{on } S \times \Gamma,$$

$$u_1(0) = U_1, \quad u_2(0) = U_2 \quad \text{in } \Omega$$

with positive diffusion coefficients $D_1, D_2 \in \mathbb{R}$, a positive parameter $\kappa \in \mathbb{R}$, and positive reaction rate coefficients $k_{1+}, k_{1-} \in \mathbb{R}$. Since the material is assumed to be homogeneous the reference densities are spatially constant and therefore using Boltzmann statistics (1.4.3) and

$$k_{(\alpha, \beta)}(x) = k_{1+} \bar{u}_1(x)^2 = k_{1-} \bar{u}_2(x) \quad \forall x \in \Omega \quad (5.8.6)$$

the system (5.8.5) can be rewritten with $R_{(\alpha, \beta)} = k_{(\alpha, \beta)}(a_1^2 - a_2)$ in terms of chemical potentials v and concentrations u (see the continuous problem (2.2.P)), i.e.,

$$\left. \begin{aligned} u_1' - \nabla \cdot (D_1 u_1 \nabla v_1) + 2\kappa R_{(\alpha, \beta)} &= 0 \\ u_2' - \nabla \cdot (D_2 u_2 \nabla v_2) - \kappa R_{(\alpha, \beta)} &= 0 \end{aligned} \right\} \text{ in } S \times \Omega,$$

$$\mathbf{n} \cdot \nabla u_1 = \mathbf{n} \cdot \nabla u_2 = 0 \quad \text{on } S \times \Gamma,$$

$$u_1(0) = U_1, \quad u_2(0) = U_2 \quad \text{in } \Omega.$$

Well-posedness of problem (5.8.5) is considered in [11, 20] and the limit to an instantaneous reaction if κ becomes large is studied in [11]. In particular, using a suitable Lyapunov functional and compactness arguments, they proof that

$$u^\kappa \rightarrow u \text{ in } L^2(S, Y) \text{ as } \kappa \rightarrow \infty, \quad Y := L^2(\Omega)^2.$$

Here $u^\kappa = (u_1^\kappa, u_2^\kappa)$ is a solution of (5.8.5) for fixed parameter κ . Let $k_{eq} = k_{1+}k_{1-}^{-1}$ denotess the equilibrium constant of the reaction. The limit $\kappa \rightarrow \infty$ can be determined by

$$k_{eq}u_1^2 = u_2 \quad \text{and} \quad \frac{1}{2}u_1 + u_2 = \mathfrak{w} \quad (5.8.7)$$

where \mathfrak{w} is the unique weak solution of the nonlinear diffusion problem

$$\begin{aligned} \mathfrak{w}' - \Delta\varphi(\mathfrak{w}) &= 0 \text{ in } S \times \Omega, \\ \mathbf{n} \cdot \nabla\mathfrak{w} &= 0 \text{ on } S \times \Gamma, \\ \mathfrak{w}(0) &= \mathfrak{W} := \frac{1}{2}U_1 + U_2 \quad \text{in } \Omega \end{aligned} \quad (5.8.8)$$

with

$$\varphi = \left(\frac{D_1}{2}\text{id} + D_2\eta \right) \circ \left(\frac{1}{2}\text{id} + \eta \right)^{-1} \quad \text{and} \quad \eta(x) = k_{eq}x^2, \quad x \geq 0.$$

The first equation in (5.8.7) states the reaction is in equilibrium. The second one defines the conserved quantity by the reaction. An explicit unique representation of u_1 and u_2 can be obtained from $\frac{1}{2}\text{id} + \eta = \mathfrak{w}$, i.e.

$$\begin{aligned} u_1 &= \left(\frac{1}{2}\text{id} + \eta \right)^{-1}(\mathfrak{w}) = \frac{1}{4k_{eq}} \left(\sqrt{1 + 16k_{eq}\mathfrak{w}} - 1 \right), \\ u_2 &= \eta \circ \left(\frac{1}{2}\text{id} + \eta \right)^{-1}(\mathfrak{w}) = \frac{1}{16k_{eq}} \left(\sqrt{1 + 16k_{eq}\mathfrak{w}} - 1 \right)^2. \end{aligned} \quad (5.8.9)$$

Using (5.8.7) the limit system can be written in the form

$$\begin{aligned} (u_1 + 2u_2)' - \Delta(D_1u_1 + 2D_2u_2) &= 0, \quad k_{eq}u_1^2 = u_2 \text{ in } S \times \Omega, \\ \mathbf{n} \cdot \nabla(D_1u_1 + 2D_2u_2) &= 0 \text{ on } S \times \Gamma, \\ u_1(0) + 2u_2(0) &= U_1 + 2U_2 \quad \text{in } \Omega. \end{aligned} \quad (5.8.10)$$

The parabolic equation describes the evolution of the conserved quantity. Integrating this equation leads to the same global invariant as (5.8.5), see Remark 3.6.2, Example 3.8.1 and [111]. Problem 5.8.5 and 5.8.7 have the same thermodynamic equilibrium and dissipative structure.

The implicit Euler in time and finite volume in space discretization of 5.8.5 was considered in [34]. Using a discrete comparison principle, the existence and uniqueness of solutions for the discrete problem was proven. Moreover, they show that the discrete solutions $(u_{1,h}, u_{2,h})$ converge in $L^2(S, Y)$ to a solution of (5.8.8) if $\kappa \rightarrow \infty$ and $\text{size}(\mathcal{D}) \rightarrow 0$, see [34, Subsection 5.1]. Voronoi meshes represent a class of admissible finite volume meshes [32]. Hence for the special case of homogeneous materials, the finite volume method analyzed in [34] is equivalent to the method introduced in Section 3.2. Therefore we expect the same convergence behavior with respect to κ .

In the following we are interested in the experimental convergence order if κ becomes large, see [34, Section 6]. In order to compare the discrete solution $(u_{1,h}, u_{2,h})$, obtained with the scheme of Section 3.2, and the discrete solution $(\mathbf{u}_{1,h}, \mathbf{u}_{2,h})$ of the limit system the following quantities

$$E_a = \|\mathbf{u}_h/\bar{u}_h - u_h/\bar{u}_h\|_{L^2(S,Y)} \|\mathbf{u}_h/\bar{u}_h\|_{L^2(S,Y)}^{-1}, \quad E_r = \|(u_{1,h}/\bar{u}_{1,h})^2 - u_{2,h}/\bar{u}_{2,h}\|_{L^2(S,L^2)}$$

are introduced. The first quantity E_a measures the distance between the solution of the discretized problem and the solution of the discretized limit system and E_r measures if the reaction term is in equilibrium. The solution $(\mathbf{u}_{1,h}, \mathbf{u}_{2,h})$ is obtained by discretizing the conserved quantity (5.8.10). The discretized equation (5.8.10) preserves the same qualitative properties: mass invariant, thermodynamic equilibrium and dissipative structure.

The simulations were done on a domain $[0, 0.1] \times [0, 1]$ and the tensor product mesh consisting of 213 nodes (3 nodes in y direction and 71 nodes in x direction). The same parameter as given in [34, Section 6] are chosen, i.e. forward and backward reaction rate constants

$$k_{1+} \approx 1.072 \cdot 10^{-4} \text{L}^2 \text{mol}^{-2}, \quad k_{1-} \approx 2.363 \cdot 10^{-6} \text{L}^2 \text{mol}^{-2}$$

and diffusion constants

$$D_1 \approx 1.579 \cdot 10^{-9} \text{m}^2 \text{s}^{-1}, \quad D_2 \approx 1.042 \cdot 10^{-9} \text{m}^2 \text{s}^{-1}.$$

These parameters are experimentally estimated in the case of benzene as solvent at a temperature $T = 298\text{K}$. Due to (5.8.6), the reaction rate constants are equivalent to

$$\bar{u}_1 = 10^{-2}, \quad \bar{u}_2 = 0.4537, \quad k_{(\alpha,\beta)} = 1.072 \cdot 10^{-6}.$$

The value of κ is varied in the range of $[10^{-12}, 10^{12}]$, see Table 5.17. In order to fulfill the positivity of the initial data (see Assumption A1), we set

$$U_1(x) = \begin{cases} 10^{-30} & \text{for } x \in [0, 0.03], \\ \frac{1}{2} \sin\left(\frac{50\pi}{7}(x - 0.03)\right) & \text{for } x \in [0.03, 0.1], \end{cases}$$

$$U_2(x) = \begin{cases} \frac{1}{4} \cos\left(\frac{50\pi}{7}x\right) & \text{for } x \in [0, 0.07], \\ 10^{-30} & \text{for } x \in [0.07, 0.1]. \end{cases}$$

In Figure 5.28 the evolution in time of $(u_{1,h}, u_{2,h})$ along the line $y = 0$ for different values of κ is depicted. At first for smaller κ_i , $i = 1, 2$ (see Table 5.17), the concentrations are flattening in

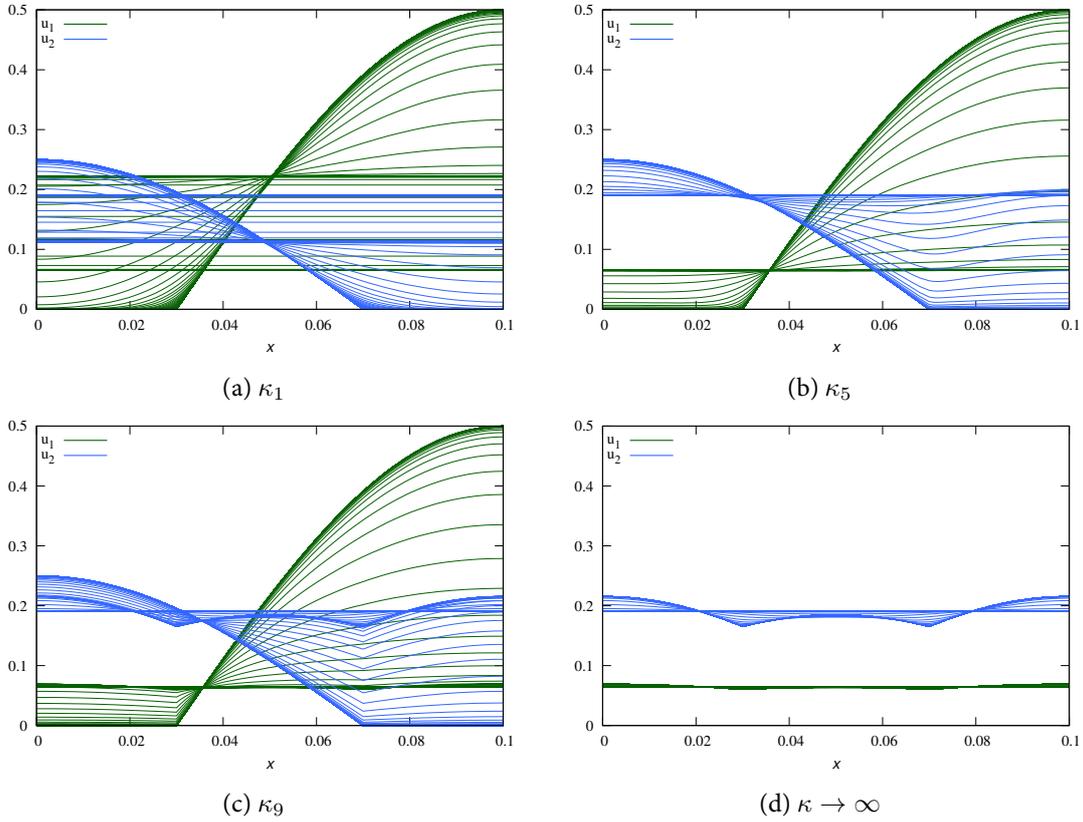


Figure 5.28: Evolution profile of the numerical solution $(u_{1,h}, u_{2,h})$ for various choices of κ . Every line represents an other time

space by diffusion and then the thermodynamic equilibrium is reached on the reaction timescale. For κ_i , $i = 3, 4, 5$, the diffusion and the reaction timescale is comparable, the concentration is uniformly transformed into the thermodynamic equilibrium. For larger κ_i the reaction is dominant.

The concentrations $(u_{1,h}, u_{2,h})$ become a w-profile which is smoothed out by diffusion and the thermodynamic equilibrium is reached. In the limit system the evolution of the initial value to the w-profile is replaced by an instantaneous reaction step (the solution of $a_{1,h}(x)^2 = a_{2,h}(x)$ and $u_{1,h}(x) + 2u_{2,h}(x) = U_{1,h}(x) + 2U_{2,h}(x)$ for every $x \in \Omega$), see (5.8.9).

i	κ_i	$\kappa_i k_{(\alpha,\beta)}$	E_a	E_r
1	10^{-12}	$1.072 \cdot 10^{-18}$	$1.772 \cdot 10^{-6}$	$9.205 \cdot 10^7$
2	10^{-9}	$1.072 \cdot 10^{-15}$	$5.266 \cdot 10^{-8}$	$2.724 \cdot 10^6$
3	10^{-4}	$1.072 \cdot 10^{-10}$	$1.600 \cdot 10^{-10}$	$8.513 \cdot 10^3$
4	10^{-2}	$1.072 \cdot 10^{-8}$	$2.033 \cdot 10^{-11}$	$1.337 \cdot 10^3$
5	1	$1.072 \cdot 10^{-6}$	$2.355 \cdot 10^{-12}$	$1.668 \cdot 10^2$
6	4^2	$6.861 \cdot 10^{-5}$	$2.127 \cdot 10^{-13}$	$1.507 \cdot 10^1$
7	4^4	$1.098 \cdot 10^{-3}$	$5.319 \cdot 10^{-14}$	3.766
8	10^6	1.072	$4.084 \cdot 10^{-15}$	$3.200 \cdot 10^{-1}$
9	10^8	$1.072 \cdot 10^2$	$1.359 \cdot 10^{-15}$	$1.090 \cdot 10^{-1}$
10	10^{12}	$1.072 \cdot 10^6$	$1.781 \cdot 10^{-17}$	$1.405 \cdot 10^{-3}$
$(\kappa_i/\kappa_{10})^p$			$p = 0.459$	$p = 0.450$

Table 5.17: The calculated values of E_a and E_r for different values of κ_i . A convergence rate of 0.459 for E_a and 0.450 for E_r with respect to κ_i is observed

Table 5.17 shows the values of the error E_a and E_r . For larger κ , the error becomes smaller which means that the system approximates the discrete solution $(u_{1,h}, u_{2,h})$ of the discretized limit system, see also Figure 5.30c and 5.30d. The experimentally observed order of convergence with respect to κ is close to $1/2$ (see Table 5.17) which is in good agreement with the theoretical results obtained in [34, Eq. (4.4) and Theorem 5.2]. Figure 5.30a shows the decay of the discrete relative free energy $\widehat{\Psi}$ for different values of κ . If the reaction is slower than the diffusion (compare the curves for κ_1 to κ_3 in Figure 5.30a), the concentration is equalized in space before the reaction timescale is reached. Therefore the free energy decays to a plateau. If the reaction timescale is reached the system approximates the thermodynamic equilibrium. If the reaction and the diffusion timescale are of the same order (compare the curves for κ_4 and κ_5 in Figure 5.30a), the initial value is uniformly transformed into the thermodynamic equilibrium. Hence the relative free energy decays without a plateau. If the reaction becomes faster than the diffusion (compare the curves for κ_6 to κ_{10} in Figure 5.30a), the relative free energy decays to a plateau. The value of $\widehat{\Psi}$ is approximately equal to the initial value of $\widehat{\Psi}$ for the limit system. At this value the concentration has the w-profile, see Figure 5.29b, 5.29a, and 5.30c. After that the w-profile is equalized by the diffusion process and the relative free energy decays to zero. Since the diffusion process is the same for finite κ and the limit

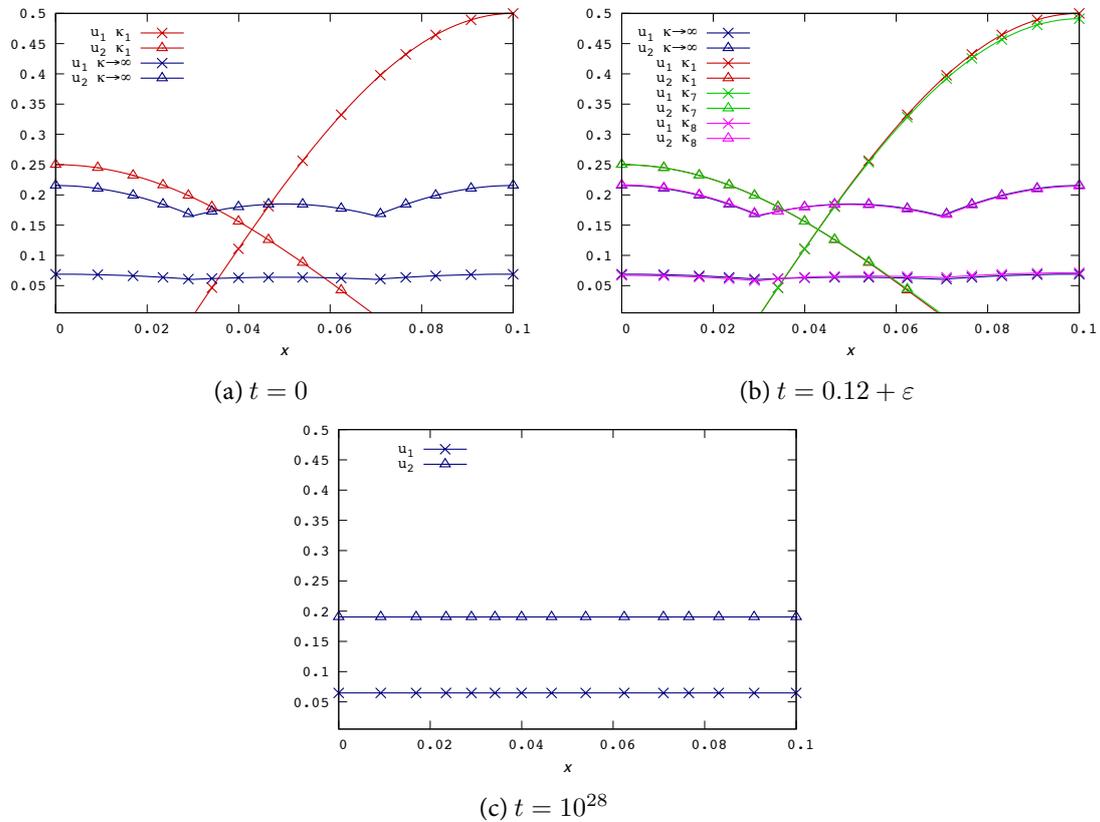


Figure 5.29: The concentration for different values of κ at time $t = 0, 0.12 + \varepsilon, 10^{28}$. The initial value of the limit system has a w-profile, which is approximated by the full system if κ becomes large.

case the curves of $\widehat{\Psi}$ stay close together. The dissipation of the discrete problem for different values of κ is depicted in Figure 5.30b. A larger value of κ leads to a higher dissipation since \widehat{D} is the time derivative of the free energy. The initial value of the dissipation in case of the instantaneous reaction limit is smaller, since in this case the first decay is infinitely fast.

Summary The example shows the connection between the different timescales of the system, defined by the reaction rate constants and the diffusion constants, and the decay of the relative free energy. The first downturn of the relative free energy is related to the fastest time scale and the last downturn with the slowest timescale, see Figure 5.30a. In the limit case the reaction timescale is infinity fast and therefore the initial value of the limit system is approximately the same as the value of the full system after the evolution of the fast reaction timescale is finished.

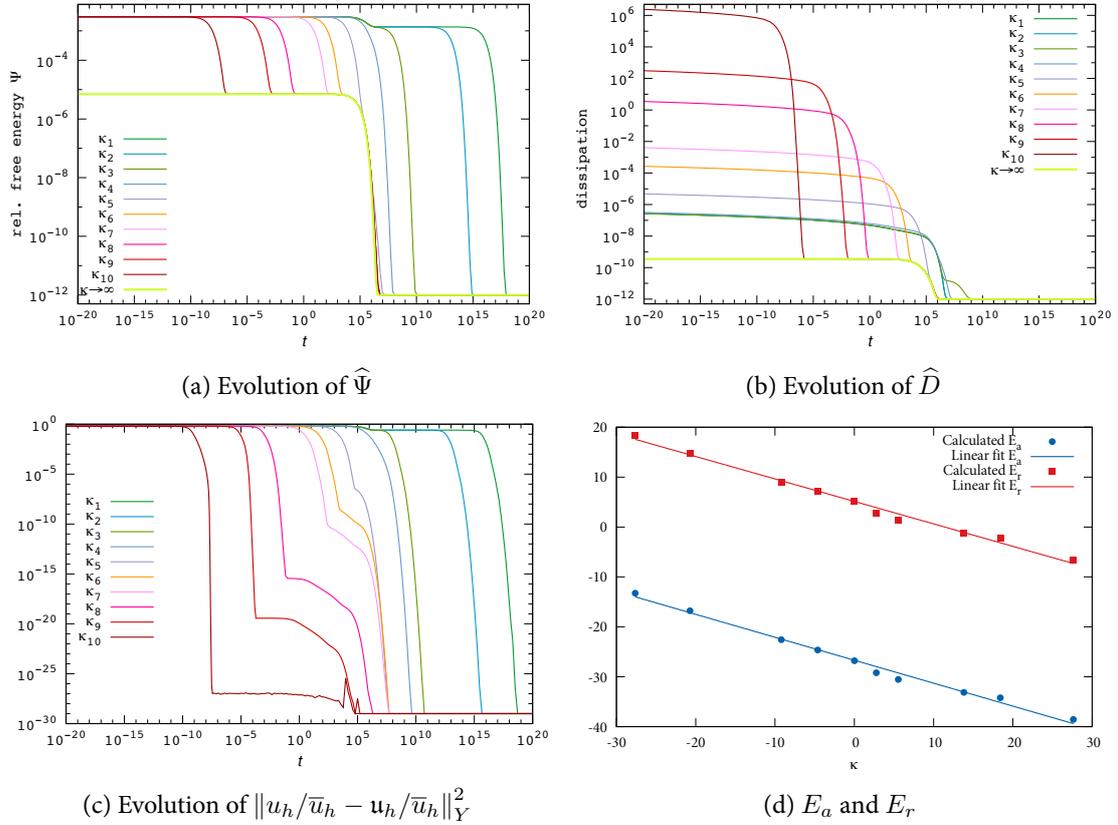


Figure 5.30: Evolution of the discrete relative free energy $\widehat{\Psi}$, the norm $\|u_h/\bar{u}_h - u_h/\bar{u}_h\|_Y^2$ and the errors E_a and E_r for different values of κ . In Figure 5.30a numerical zero is set to 10^{-12} and in Figure 5.30c to 10^{-29} . The slope of the linear fitted lines are approximately $-1/2$, see Table 5.17

5.9 Conclusion

The example of Section 5.2 shows the consequences if the invariants property (see Section 3.4) is violated by the method and illustrates the behavior of the introduced Lagrange multipliers. In the following example of Section 5.3 we saw that heterogeneous materials may lead to strong internal layers on the interfaces between two materials. Using strong varying diffusivities and fast reactions, it is possible to create 'boundary' reactions and the resulting layer is hardly to resolve by the mesh.

The examples of Section 5.4 and 5.5 show that the method can reach the thermodynamic equilibrium on different discretizations. The system travels on comparable energy curves into the thermodynamic equilibrium. The method attains comparable quasi stationary states but the arrival times vary due to the faster diffusion on coarser meshes. On quasi stationary points the energy curves come close together. Moreover we saw that a second order time integration method (BDF-2) is more effective on fine meshes, where more details of the solution are resolved and the evolution of the

solution is smoother. We establish comparable rates of convergences for the relative free energy (close to 2) and for the dissipation rate (close to 1.5) if the discretization size tends to zero.

The example of Section 5.6 takes points into account, where three materials meet each other in one point. In Section 5.7 we illustrate the behavior of the scheme in case of reaction rate constants vanishing on subdomains. Finally, in the example of Section 5.8 we illustrate the behavior of the method if the reaction timescale becomes fast. In all these examples of different complexity with many different fast and slow timescales the method can reach the thermodynamic equilibrium and all qualitative properties of the scheme are fulfilled up to roundoff errors.

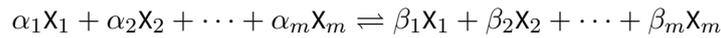
A

Appendix

A.1 Symmetric reaction coefficients

In the usual setting a reversible reaction is described by a reaction rate constant for the forward reaction and one for the reverse reaction. In this part we will show that both constants can be transformed to a single rate constant as defined in the model of Section 1.4. For this purpose we introduce the equilibrium constant of a reaction and the Wegscheider condition [126].

Assuming that, according to the mass action law, the reaction rate of a reversible reaction



for a pair $(\alpha, \beta) \in \mathcal{R}$ is given by

$$R_{(\alpha, \beta)}(\cdot, u) = (k_\alpha u^\alpha - k_\beta u^\beta), \quad (\text{A.1})$$

where $k_\alpha : \Omega \rightarrow \mathbb{R}_+$, $k_\beta : \Omega \rightarrow \mathbb{R}_+$ are positive functions of the spatial variable, and $u^\alpha := \prod_{\nu=1}^m u_\nu^{\alpha_\nu}$. Then the equilibrium constant of a reaction is introduced by $K_{(\alpha, \beta)}^{eq} = \frac{k_\alpha}{k_\beta}$ for $(\alpha, \beta) \in \mathcal{R}$. If during the evolution of time $K_{(\alpha, \beta)}^{eq} > u^{(\beta-\alpha)}$ holds, then the forward direction is dominant. Otherwise if $K_{(\alpha, \beta)}^{eq} < u^{(\beta-\alpha)}$ holds, then the backward reaction is dominant. The reaction is in equilibrium if $K_{(\alpha, \beta)}^{eq} = u^{(\beta-\alpha)}$ holds. For a single reaction the positive equilibrium solution is determined by the equilibrium constant. In complex reaction networks the Wegscheider condition imposes restrictions on the choice of the forward and backward rate constants, see [83, 126].

We introduce the stoichiometric matrix Γ and the column vector \mathbf{K} by

$$\Gamma = (\beta_1 - \alpha_1, \dots, \beta_m - \alpha_m)_{(\alpha, \beta) \in \mathcal{R}}, \quad \mathbf{K} = (K_{(\alpha, \beta)}^{eq})_{(\alpha, \beta) \in \mathcal{R}}.$$

To ensure the existence of a positive equilibrium solution $u^* = e^{\tilde{v}^*} > 0$ of

$$\Gamma \tilde{v}^* = \ln \mathbf{K} \quad (\text{A.2})$$

we assume that for every solution $\lambda = (\lambda_{(\alpha,\beta)})_{(\alpha,\beta) \in \mathcal{R}}$ of the linear system

$$\Gamma^T \lambda = \mathbf{0} \quad (\text{A.3})$$

the Wegscheider condition

$$\prod_{(\alpha,\beta) \in \mathcal{R}} k_{\alpha}^{\lambda_{(\alpha,\beta)}} = \prod_{(\alpha,\beta) \in \mathcal{R}} k_{\beta}^{\lambda_{(\alpha,\beta)}}$$

holds. This condition is necessary and sufficient for the existence of $u^* = e^{\tilde{v}^*} > 0$, see [66]. From the physical point of view the Wegscheider condition ensures the thermodynamical consistency of the reaction system. In mathematical sense, the condition means that the inhomogeneous linear system (A.2) has a solution iff the rank of the matrix is equal to the rank of the extended matrix. This means that every solution to (A.3) is orthogonal to the right-hand side of (A.2).

Assuming that the reactions satisfy the Wegscheider condition, there exist a vector \bar{u} such that

$$k_{(\alpha,\beta)} = k_{\alpha} \prod_{\nu=1}^m \bar{u}_{\nu}^{\alpha_{\nu}} = k_{\beta} \prod_{\nu=1}^m \bar{u}_{\nu}^{\beta_{\nu}}, \quad (\alpha, \beta) \in \mathcal{R},$$

holds and that the reaction rates (A.1) can be written in the form

$$R_{(\alpha,\beta)} = k_{(\alpha,\beta)} \left(a^{\alpha} - a^{\beta} \right),$$

which is exactly (1.4.5), see also [102]. The section is closed by two examples to illustrate the Wegscheider condition.

Example A.1 (MMH-Kinetic, see Example 1.4.1). *The set \mathcal{R} consists of the pairs $\alpha_1 = (1, 1, 0, 0)$, $\beta_1 = (0, 0, 1, 0)$, and $\alpha_2 = (0, 0, 1, 0)$, $\beta_2 = (0, 1, 0, 1)$. The linear system to determine a positive equilibrium solution is given by*

$$\begin{pmatrix} \ln K_{eq,1} \\ \ln K_{eq,2} \end{pmatrix} = \begin{pmatrix} -1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{pmatrix} \tilde{v}^*.$$

A solution of the system exists, since the Wegscheider condition $k_{\alpha_1}^{\lambda_1} k_{\alpha_2}^{\lambda_2} = k_{\beta_1}^{\lambda_1} k_{\beta_2}^{\lambda_2}$ is satisfied with $\lambda_1 = 0$ and $\lambda_2 = 0$.

Example A.2 (see [83, S. 109]). *We consider the reaction system*



This network creates X_5 over two different ways. On the one hand, the substance X_1 reacts together with X_2 to form the substance X_3 which then reacts together with X_4 to form X_5 . On the other hand, the species X_1 reacts together with X_4 to form the substance X_6 which further reacts with X_2 to form X_5 . The stoichiometric subspace \mathcal{S} consists of

$$\mathcal{S} = \text{span}\{(1, 1, -1, 0, 0, 0), (1, 0, 0, 1, 0, -1), (0, 0, 1, 1, -1, 0), (0, 1, 0, 0, -1, 1)\}$$

and the matrix Γ is given by

$$\Gamma = \begin{pmatrix} -1 & -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

The solution of (A.3) is spanned by $\lambda = (-1, 1, -1, 1)$ and the Wegscheider condition imposes the following dependencies on the reaction rates

$$\frac{k_2^+ k_4^+}{k_2^- k_2^-} = \frac{k_1^+ k_3^+}{k_1^- k_3^-}.$$

By the choice of appropriate reaction coefficients it is possible to create cyclic reactions. In this example such a behavior is excluded by the Wegscheider condition.

A.2 Discrete Gagliardo-Nirenberg inequality

In this section we provide a discrete Gagliardo-Nirenberg and a Sobolev-Poincaré inequality for boundary conforming Delaunay-Voronoi meshes. In the first version, the equivalence of the Voronoi finite volume L^p -norm, the Donald box finite volume L^p -norm (for the definition of a Donald box see [69]), and the linear finite element L^p -norm is exploited. The proof provides a simple way to translate continuous inequalities into the discrete world. A drawback of this approach is the limitation to two space dimensions.

During the work at this thesis, another proof of the discrete Gagliardo-Nirenberg and the Sobolev-Poincaré inequality was given by [10]. The version works in higher space dimensions and for general admissible finite volume meshes. Therefore it is possible to handle all (boundary conforming) Delaunay-Voronoi meshes, including arbitrarily anisotropic grids. We recapitulate the result and infer the special version of the Gagliardo-Nirenberg inequality (B.24) from (B.19).

A.2.1 1. Approach – restricted to 2d

In the following, we prove a discrete version of the Gagliardo-Nirenberg inequality [108] on Voronoi finite volume meshes. The approach is restricted to two space dimensions. First, we introduce

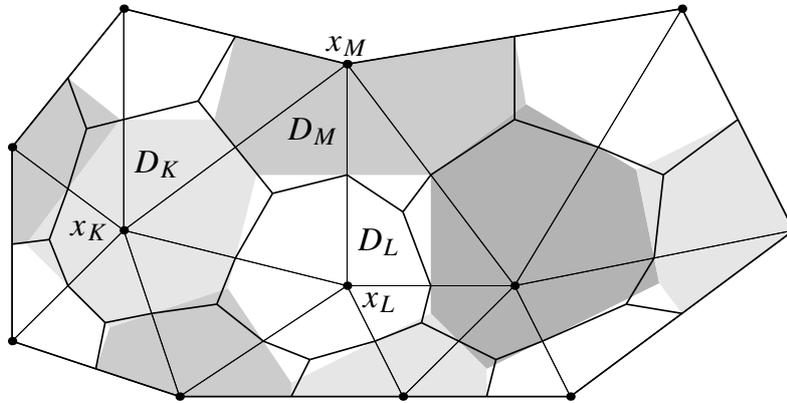


Figure A.1: Notation of the dual Voronoi mesh and the Donald box mesh. Voronoi boxes (gray areas), triangles (thin lines), Donald boxes (thick lines)

the following notation, see also Figure A.1 : The dual of a Voronoi mesh $\mathcal{M} = (\mathcal{P}, \mathcal{V}, \mathcal{E})$ is given by a tuple $(\mathcal{P}, \mathcal{T})$, where \mathcal{T} is a family of triangles T spanned by $(x_K, x_L, x_M) \in \mathcal{P}^3$ fulfilling $\overline{K} \cap \overline{L} \cap \overline{M} \neq \emptyset$, $K, L, M \in \mathcal{V}$. Furthermore, by $\mathcal{N}_T(K)$ we denote the set of all triangles sharing x_K as common vertex.

A *Donald box* $D_K = D(x_K)$ around a node $x_K \in \mathcal{P}$ is the polygonal control volume obtained by joining the barycenters and the midpoints of the edges of the triangles surrounding a node x_K ,

see [69, 116]. The area of a triangle $T \in \mathcal{T}$ is denoted by $|T|$ and the area of the part of a Donald box contained in the triangle T is given by $|D_K \cap T| = 1/3|T|$. By $X_D(\mathcal{M})$ we denote the set of all piecewise constant functions from Ω to \mathbb{R} which are constant on every Donald box D_K . The discrete Donald box finite volume L^p -norm is then introduced by

$$\|w_D\|_{L^p} = \left(\sum_{K \in \mathcal{V}} |D_K| |w_K|^p \right)^{1/p} \quad \forall w_D \in X_D(\mathcal{M}). \quad (\text{B.4})$$

Let $w_T(\mathbf{x})$ be the linear function on a triangle $T \in \mathcal{T}$ with nodes $(x_K, x_L, x_M) \in \mathcal{P}^3$ and values $w_T(x_K) = w_K, w_T(x_L) = w_L, w_T(x_M) = w_M$, where w_K , resp., is the value in the node $x_K \in \mathcal{P}$. The set of all these functions is denoted by $P_1(\mathcal{M})$. The finite element L^p -norm is then defined by

$$\|w_T\|_{L^p} = \left(\sum_{T \in \mathcal{T}} \int_T |w_T(\mathbf{x})|^p d\mathbf{x} \right)^{1/p} \quad \forall w_T \in P_1(\mathcal{M}). \quad (\text{B.5})$$

We make the following assumption on the mesh: There exist two constants $\underline{C}_D > 0, \overline{C}_D \geq 1$ such that for all $x_K \in \mathcal{P}$ the area of the Donald box around x_K can be estimated by the area of the Voronoi box K , i.e., it holds

$$\underline{C}_D |K| \leq |D_K| \leq \overline{C}_D |K| \quad \forall K \in \mathcal{V}. \quad (\text{B.6})$$

Remark B.1. Such an inequality is valid for all Delaunay-meshes where either:

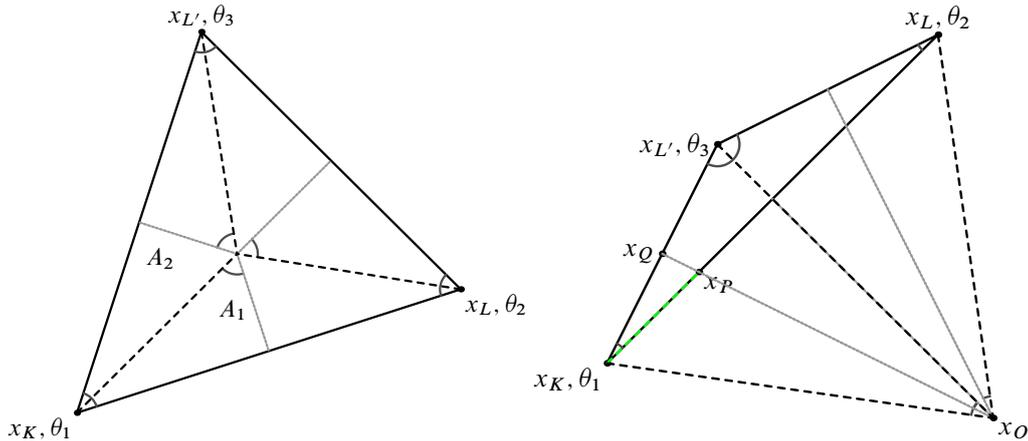


Figure A.2: Illustration of the two typical cases: acute triangles (left) and triangles with obtuse angles

1. All triangles with a common vertex $x_K \in \mathcal{P}$ are acute.

2. Or if one triangle hanging on a common vertex $x_K \in \mathcal{P}$ has an obtuse angle, then there exists a $0 < \theta_{min} \leq \pi/3$ such that the angles of all triangles hanging on a common vertex $x_K \in \mathcal{P}$ are greater or equal θ_{min} .

The area of a triangle is given by

$$|T| = \frac{d_{KL}^2 \sin \theta_1 \sin \theta_2}{2 \sin \theta_3} = \frac{d_{KL'}^2 \sin \theta_1 \sin \theta_3}{2 \sin \theta_2}, \quad (\text{B.7})$$

where d_{KL} and $d_{KL'}$ is the length of the edge $\overline{x_X x_L}$ and $\overline{x_X x_{L'}}$, resp.. The area $A = A_1 + A_2$ of the Voronoi snippet spanned by T can be calculated by

$$\begin{aligned} A &= A_1 + A_2 = \frac{d_{KL}^2}{8} \cot \theta_3 + \frac{d_{KL'}^2}{8} \cot \theta_2 \\ &= \frac{|T|}{4} \left(\frac{\cos \theta_3}{\sin(\theta_3 + \theta_2) \sin \theta_2} + \frac{\cos \theta_2}{\sin(\theta_3 + \theta_2) \sin \theta_3} \right) \\ &= \frac{|T|}{4} (1 + \cot \theta_3 \cot \theta_2). \end{aligned}$$

1. If all triangles with a common vertex $x_K \in \mathcal{P}$ are acute, then the Voronoi box K is convex and every triangle edge $\sigma = K|L \in \mathcal{E}_K$ is divided into two equal sized parts. Therefore dividing every triangle edge σ by the factor two, all triangles with the a common vertex $x_K \in \mathcal{P}$ are inside the Voronoi box K and it results $|K| \geq \frac{1}{4}|T|$. If the angle related to x_K is $\pi/2$ then $|K \cap T| = \frac{1}{2}|T|$. Hence $\frac{1}{4} \leq \frac{|K|}{|T|} \leq \frac{1}{2}$.
2. Let $0 < \theta_{minK}$. If $\pi/2 \leq \theta_1 \leq \pi - 2\theta_{minK}$ and $\theta_{minK} \leq \theta_2, \theta_3$, then the circumcenter x_O is outside of the triangle and the largest edge is $\overline{x_L x_{L'}}$. We obtain the upper bound

$$\frac{A}{|T|} \leq \frac{1}{4} (1 + \cot(\theta_{minK})^2). \quad (\text{B.8})$$

3. Let $\pi/2 \leq \theta_3 \leq \pi - 2\theta_{minK}$ and $\theta_{minK} \leq \theta_1, \theta_2$. Then the largest edge is opposite to θ_3 . The length d_{KP} of the line $\overline{x_K x_P}$ is given by $d_{KP} = \frac{d_{KL'}}{2 \cos \theta_1}$ and by sine rule and $\theta_3 = \pi - \theta_1 - \theta_2$ we obtain

$$\frac{d_{KP}}{d_{KL}} = \frac{\sin \theta_2}{2 \cos \theta_1 \sin \theta_3} = \frac{1}{2} \frac{\sec \theta_1 \sin \theta_2}{\sin \theta_1 \cos \theta_2 + \cos \theta_1 \sin \theta_2} \geq \frac{\sec(\theta_{minK})^2}{4}. \quad (\text{B.9})$$

After enlarging the Voronoi box K by the factor $\frac{d_{KP}}{d_{KL}}$, all triangles sharing x_K lie inside the Voronoi box K and we obtain

$$|K| \geq \frac{\sec(\theta_{minK})^4}{8} |T|. \quad (\text{B.10})$$

There is a special case where the circumcenter x_O of the triangle $T = (x_K, x_L, x'_L)$ is contained in the l -th neighboring triangle of T sharing x_K as common node and all these l neighboring triangles have the same circumcenter x_O , i.e., all nodes of the triangles lie on a circle. Then the similitude radius is given by the largest edge of the l neighboring triangles and the intersection of line $\overline{x_Q x_O}$. Using sine rule and knowing that x_K is a common node of all these triangles, the radio can be determined analogously to the case 3.

We make the following observation:

Lemma B.2. *Let Ω be an open, bounded, polygonal subset of \mathbb{R}^2 and $\mathcal{M} = (\mathcal{P}, \mathcal{V}, \mathcal{E})$ a given Voronoi mesh and $(\mathcal{P}, \mathcal{T})$ its dual Delaunay triangulation fulfilling (B.6). The following statements are true:*

1. For all $w_h \in X_{\mathcal{V}}(\mathcal{M})$, $w_D \in X_D(\mathcal{M})$ and $p \in \mathbb{N}$, $w_h \geq 0$ the estimate

$$\underline{C}_D^{1/p} \|w_h\|_{L^p} \leq \|w_D\|_{L^p} \leq \overline{C}_D^{1/p} \|w_h\|_{L^p} \quad (\text{B.11})$$

is valid.

2. Furthermore there exist two constants $\underline{C}_{FEM,p} = \frac{6}{(p+1)(p+2)} \leq 1$ and $\overline{C}_{FEM} = 1$ such that for $w_D \in X_D(\mathcal{M})$ and $w_T \in P_1(\mathcal{M})$, $w_D \geq 0$

$$\underline{C}_{FEM,p}^{1/p} \|w_D\|_{L^p} \leq \|w_T\|_{L^p} \leq \overline{C}_{FEM} \|w_D\|_{L^p} \quad (\text{B.12})$$

holds.

Proof. For simplifying the notation, Assume that w_K is nonnegative for all $K \in \mathcal{V}$. The proof of (B.11) follows immediately from (B.6) together with

$$\|w_D\|_{L^p} = \left(\sum_{K \in \mathcal{V}} |D_K| w_K^p \right)^{1/p} \leq \left(\sum_{K \in \mathcal{V}} \overline{C}_D |K| w_K^p \right)^{1/p} = \overline{C}_D^{1/p} \|w_h\|_{L^p}.$$

Similar, the estimate from above is derived. For the second assertion (B.12) we observe together with $|D_K \cap T| = 1/3|T|$ that

$$\|w_D\|_{L^p}^p = \sum_{K \in \mathcal{V}} \sum_{T \in \mathcal{N}_T(K)} \frac{|T|}{3} w_K^p = \sum_{T \in \mathcal{T}} \frac{|T|}{3} (w_K^p + w_L^p + w_M^p)$$

with $T = (x_K, x_L, x_M)$. It remains to show that for every triangle the identities

$$\frac{\underline{C}_{FEM,p}|T|}{3} (w_K^p + w_L^p + w_M^p) \leq \int_T w_T(\mathbf{x})^p d\mathbf{x} \leq \frac{\overline{C}_{FEM,p}|T|}{3} (w_K^p + w_L^p + w_M^p)$$

with constants $\underline{C}_{FEM,p}$ and $\overline{C}_{FEM,p}$ hold. In a first step the triangle $T = (\mathbf{x}_K, \mathbf{x}_L, \mathbf{x}_M)$ is transformed to the reference triangle $\widehat{T} = (P_1, P_2, P_3)$ with vertices $P_1 = (0, 0)$, $P_2 = (1, 0)$, $P_3 = (0, 1)$ using

$$\mathbf{x} = \mathbf{x}_K + (\mathbf{x}_L - \mathbf{x}_K)\xi + (\mathbf{x}_M - \mathbf{x}_K)\rho$$

with $\xi \in [0, 1]$ and $\rho \in [0, \xi]$. The transformed function reads as

$$\begin{aligned} \widehat{w}(\xi, \rho) &= w_T(\mathbf{x}_K + (\mathbf{x}_L - \mathbf{x}_K)\xi + (\mathbf{x}_M - \mathbf{x}_K)\rho) \\ &= w_K + (w_L - w_K)\xi + (w_M - w_K)\rho \end{aligned}$$

and its functional determinant is given by $\det(J) = 2|T|$. The integral over the triangle T can be written in the form

$$\begin{aligned} I &= \int_T (w_T(\mathbf{x}))^p d\mathbf{x} = 2|T| \int_0^1 \int_0^{1-\xi} (\widehat{w}(\xi, \rho))^p d\rho d\xi \\ &= \frac{2|T|}{(p+1)(w_M - w_K)} \int_0^1 (w_M + (w_L - w_M)\xi)^{p+1} - (w_K + (w_L - w_K)\xi)^{p+1} d\xi \\ &= \frac{2|T|}{(p+2)(p+1)(w_M - w_K)} \left(\frac{w_M^{p+2} - w_L^{p+2}}{w_M - w_L} - \frac{w_L^{p+2} - w_K^{p+2}}{w_L - w_K} \right). \end{aligned}$$

To continue the estimate, we use $0 \leq (v^a - 1)(v^b - 1)$ for all $v \geq 1$ and $a, b \geq 0$. Let be $v = x/y$ with nonnegative real numbers x, y . Using

$$x^a y^b + y^a x^b \leq x^{a+b} + y^{a+b} \tag{B.13}$$

we can bound the cyclic sum

$$\sum_{cycl} x^i y^j z^k := x^i (y^j z^k + y^k z^j) + y^i (x^j z^k + x^k z^j) + z^i (x^j y^k + y^k x^j)$$

by $\sum_{cycl} x^i y^j z^k \leq 2(z^{i+j+k} + x^{i+j+k} + y^{i+j+k})$.

The binomial theorem yields for $p \in \mathbb{N}$ the identity

$$\frac{w_M^{p+2} - w_L^{p+2}}{w_M - w_L} = \sum_{i=0}^{p+1} w_M^i w_L^{p+1-i}$$

and we conclude

$$\begin{aligned} I &= \frac{2|T|}{(p+1)(p+2)} \sum_{\substack{i,j,k \geq 0 \\ i+j+k=p}} w_M^i w_L^j w_K^k = \frac{2|T|}{(p+1)(p+2)} \sum_{\substack{i,j,k \geq 0 \\ i+j+k=p}} \frac{1}{6} \sum_{cycl} w_M^i w_L^j w_K^k \\ &\leq \frac{4|T|}{6(p+1)(p+2)} (w_M^p + w_L^p + w_K^p) \sum_{\substack{i,j,k \geq 0 \\ i+j+k=p}} 1 = \frac{|T|}{3} (w_M^p + w_L^p + w_K^p). \end{aligned}$$

In the last line we used

$$\sum_{\substack{i,j,k \geq 0 \\ i+j+k=p}} 1 = \sum_{i=0}^p \left(1 + \sum_{j=0}^{i-1} 1 \right) = \sum_{i=0}^p i + 1 = \frac{(p+2)(p+1)}{2}.$$

Since the term $w_M^p + w_L^p + w_K^p, w_M, w_L, w_K > 0$ is contained in $\sum_{\substack{i,j,k \geq 0 \\ i+j+k=p}} w_M^i w_L^j w_K^k$, the lower bound of I follows. \square

Theorem B.3 (Discrete Gagliardo-Nirenberg inequality). *Under the assumptions of Lemma B.2 the following statements are true:*

- For every $p \in [1, \infty)$ and let $g(p) = \frac{\max(\overline{C}_D, \overline{C}_D^{(p+1)/2})}{\underline{C}_D \underline{C}_{FEM,p}}$. Then there exist constants $c_p > 0$ and

$$C_{\mathcal{M},1,p} = \begin{cases} g(p), & \text{if } p \in \mathbb{N}, \\ \max(g(r), g(r+1)), & \text{if } p \in (r, r+1) \end{cases}$$

such that

$$\|w_h\|_{L^p} \leq c_p C_{\mathcal{M},1,p}^{1/p} \|w_h\|_{L^1}^{1/p} \|w_h\|_{H^1, \mathcal{M}}^{1-1/p} \quad \forall w_h \in X_{\mathcal{V}}(\mathcal{M}). \quad (\text{B.14})$$

Especially, for p from a compact interval $[p_1, p_2]$ the constant c_p can be bounded by

$$c_p \leq \max(c_{p_1}, c_{p_2}, 1)^{1/p_1}.$$

- Additionally, for every $\varepsilon > 0$ and $p \in (1, \infty)$ there exist constants $c_{\varepsilon,p} > 0$ such that for all $w_h \in X_{\mathcal{V}}(\mathcal{M})$

$$\begin{aligned} \|w_h\|_{L^p}^p &\leq \varepsilon C_{\mathcal{M},2,p} \|w_h \ln |w_h|\|_{L^1} \|w_h\|_{H^1, \mathcal{M}}^{p-1} \\ &\quad + c_{\varepsilon,p} C_{\mathcal{M},3,p} \|w_h\|_{L^1} \end{aligned} \quad (\text{B.15})$$

with the two constants

$$C_{\mathcal{M},2,p} = \begin{cases} g(p), & \text{if } p \in \mathbb{N}, \\ \max(g(r), g(r+1)), & \text{if } p \in (r, r+1), \end{cases}$$

$$C_{\mathcal{M},3,p} = \begin{cases} \frac{\bar{C}_D}{C_D C_{FEM,p}}, & \text{if } p \in \mathbb{N}, \\ \max\left(\frac{\bar{C}_D}{C_D C_{FEM,r}}, \frac{\bar{C}_D}{C_D C_{FEM,r+1}}\right), & \text{if } p \in (r, r+1). \end{cases}$$

The constants $C_{\mathcal{M},i,p}$, $i = 1, 2, 3$, are the only difference to the continuous version in [63].

Proof. The proof takes advantage of the fact that the discrete Voronoi finite volume gradient seminorm in $2d$ coincides with the gradient seminorm of continuous piecewise linear finite elements. From [32, Sec. 3.4.2] we obtain $|w_h|_{H^1, \mathcal{M}} = |w_T|_{H^1(\Omega)}$.

First let $p \in \mathbb{N}$. As a consequence of Lemma B.2 we can use the continuous version to prove the discrete inequalities. From Lemma B.2 we conclude the estimates

$$\|w_h\|_{L^p} \leq \frac{1}{(C_D C_{FEM,p})^{1/p}} \|w_T\|_{L^p},$$

$$\|w_T\|_{L^p} \leq \bar{C}_D^{1/p} \|w_h\|_{L^p},$$

$$\|w_T\|_{H^1, \mathcal{M}} \leq \max(1, \bar{C}_D^{1/2}) \|w_h\|_{H^1, \mathcal{M}}.$$

Using the continuous Gagliardo-Nirenberg inequality [63, (1.8)] the assertion

$$\|w_h\|_{L^p} \leq c_p C_{\mathcal{M},1,p}^{1/p} \|w_h\|_{L^1}^{1/p} \|w_h\|_{H^1, \mathcal{M}}^{1-1/p},$$

results with a constant $C_{\mathcal{M},1,p}$ defined above. Together with $r \in [p, p+1]$, $p \in \mathbb{N}$ and

$$1/r = \theta/p + (1-\theta)/(p+1), \quad \theta \in [0, 1]$$

we conclude

$$\begin{aligned} \|w_h\|_{L^r} &\leq \|w_h\|_{L^p}^\theta \|w_h\|_{L^{p+1}}^{1-\theta} \\ &\leq c_p^\theta c_{p+1}^{1-\theta} C_{\mathcal{M},1,p}^{\theta/p} C_{\mathcal{M},1,p+1}^{(1-\theta)/(p+1)} \|w_h\|_{L^1}^{\theta/p + (1-\theta)/(p+1)} \|w_h\|_{H^1}^{1-\theta/p - (1-\theta)/(p+1)}. \end{aligned} \quad (\text{B.16})$$

We mention that $x \ln x$, $x \geq 0$ is convex and therefore the linear interpolation of $w_T \ln w_T$ denoted by $(w_T \ln w_T)_T$ is greater or equal $w_T \ln w_T$. Hence,

$$\int_T w_T \ln w_T \, d\mathbf{x} \leq \int_T (w_T \ln w_T)_T \, d\mathbf{x}.$$

Using the continuous Gagliardo-Nirenberg inequality [63, (1.9)] we get

$$\begin{aligned} \|w_h\|_{L^p}^p &\leq \varepsilon \frac{1}{\underline{C}_D \underline{C}_{FEM,p}} \|(w_T \ln w_T)_T\|_{L^1} \|w_T\|_{H^1}^{p-1} \\ &\quad + c_{\varepsilon,p} \|w_T\|_{L^1} \\ &\leq \varepsilon C_{\mathcal{M},2,p} \|w_h \ln |w_h|\|_{L^1} \|w_h\|_{H^1,\mathcal{M}}^{p-1} + c_{\varepsilon,p} C_{\mathcal{M},3,p} \|w_h\|_{L^1}. \end{aligned}$$

□

A discrete *Sobolev-Poincaré inequality*, i.e., for every $p \in [1, \infty)$ exist constants $c = c(\Omega, p)$ such that

$$\|f_h - m_\Omega(f_h)\|_{L^p} \leq \frac{c_p}{(\underline{C}_D \underline{C}_{FEM,p})^{1/p}} \|f_h\|_{H^1,\mathcal{M}}, \quad m_\Omega(f_h) = \frac{1}{|\Omega|} \int_\Omega f_h dx \quad (\text{B.17})$$

for every $f_h \in X_{\mathcal{V}}(\mathcal{M})$ holds, can be proven in the same fashion as the first assertion in Theorem B.3.

A.2.2 2. Approach

In the following, we recapitulate from [10] a discrete version of the Gagliardo-Nirenberg inequality [108] and a discrete Sobolev-Poincaré inequality on *admissible finite volume meshes* [32]. The proof uses functions of bounded variation and is done under the general assumption on the mesh, that

$$\exists \xi > 0 : \quad \text{dist}(x_K, \sigma) \geq \xi |x_K - x_L| \quad \forall \sigma = K|L \in \mathcal{E}_{int} \cap \mathcal{E}_K \quad \forall x_K \in \mathcal{P}, \quad (\text{B.18})$$

see [10, p. 5.2.1]. In the case of Voronoi meshes this condition fulfilled with $\xi = \frac{1}{2}$.

Lemma B.4 (see [10, Th. 5.3.1 and Th. 5.3.2]). *Let Ω be an open bounded polyhedral domain of \mathbb{R}^d , $d \geq 2$. Let $\mathcal{M} = (\mathcal{P}, \mathcal{V}, \mathcal{E})$ be a given admissible finite volume mesh which satisfies (B.18).*

- *Then, there exists a constant $C > 0$ only depending on d and Ω (but not on \mathcal{M}) such that*

$$\|w_h\|_{L^p} \leq C_{gn,p} \|w_h\|_{L^1}^{1-\theta} \|w_h\|_{H^1,\mathcal{M}}^\theta, \quad C_{gn,p} := \frac{C}{\xi^{\theta/2}}, \quad \forall w_h \in X_{\mathcal{V}}(\mathcal{M}), \quad (\text{B.19})$$

where

$$\theta = \frac{2d(p-1)}{(d+2)p}, \quad p \in \begin{cases} [1, \infty) & \text{if } d = 2, \\ [1, 2d/(d-2)] & \text{if } d > 2. \end{cases} \quad (\text{B.20})$$

- Let $d \leq 2$. Then for all $p \in [1, \infty)$ there exists a constant $C > 0$ only depending on p, d and Ω (but not on \mathcal{M}) such that

$$\|w_h\|_{L^p} \leq \frac{C}{\xi^{1/2}} \|w_h\|_{H^1, \mathcal{M}} \quad \forall w_h \in X_{\mathcal{V}}(\mathcal{M}). \quad (\text{B.21})$$

- Let $d > 2$. Then for all $1 \leq p \leq \frac{2d}{d-2}$ there exists a constant $C > 0$ only depending on p, d and Ω (but not on \mathcal{M}) such that

$$\|w_h\|_{L^p} \leq \frac{C}{\xi^{1/2}} \|w_h\|_{H^1, \mathcal{M}} \quad \forall w_h \in X_{\mathcal{V}}(\mathcal{M}). \quad (\text{B.22})$$

Remark B.5. An inspection of the proof shows that $C_{gn,p}$ depends continuously on p and can therefore be estimated uniformly for $p \in [2, 3]$. The results of [10, Th. 5.3.1 and Th. 5.3.2] allow us to handle all Voronoi meshes, including arbitrarily anisotropic grids.

In the proof of Theorem 3.8.3 we need a special case of the discrete Gagliardo-Nirenberg inequality:

Corollary B.6. Let Ω be an open bounded polyhedral domain of \mathbb{R}^d , $d \geq 2$. Let $\mathcal{M} = (\mathcal{P}, \mathcal{V}, \mathcal{E})$ be a given admissible finite volume mesh which satisfies (B.18). For any $\varepsilon > 0$ and any $p \in (1, 3]$ there exists a constant $c_{\varepsilon,p} > 0$ such that

$$\|w_h\|_{L^p}^p \leq \frac{\varepsilon}{\xi^{(p-1)/2}} \|w_h \ln |w_h|\|_{L^1}^{p(1-\theta)} \|w_h\|_{H^1, \mathcal{M}}^{p\theta} + c_{\varepsilon,p} \|w_h\|_{L^1}. \quad (\text{B.23})$$

(For $\varepsilon \rightarrow 0$ it follows $c_{\varepsilon,p} \rightarrow \infty$). Especially for $d = 2$ we obtain the estimate

$$\|w_h\|_{L^p}^p \leq \frac{\varepsilon}{\xi^{(p-1)/2}} \|w_h \ln |w_h|\|_{L^1} \|w_h\|_{H^1, \mathcal{M}}^{p-1} + c_{\varepsilon,p} \|w_h\|_{L^1}. \quad (\text{B.24})$$

Proof. For $N > 1$ we define the function

$$\chi(s) := \begin{cases} 0, & \text{for } |s| \leq N, \\ 2(|s| - N), & \text{for } |s| \in (N, 2N], \\ |s|, & \text{for } |s| > 2N. \end{cases}$$

Adding and subtracting $\chi(w_h)$ we obtain

$$\|w_h\|_{L^p}^p \leq c_p \left(\| |w_h| - \chi(w_h) \|_{L^p}^p + \|\chi(w_h)\|_{L^p}^p \right). \quad (\text{B.25})$$

The first term of (B.25) can be estimated by

$$\| |w_h| - \chi(w_h) \|_{L^p}^p = \sum_{\substack{K \in \mathcal{V} \\ |w_K| \leq 2N}} |K| (|w_K| - \chi(w_K))^p \leq (2N)^{p-1} \|w_h\|_{L^1}.$$

Applying (B.19) for the second term of (B.25) we obtain

$$\|\chi(w_h)\|_{L^p}^p \leq \frac{C^p}{\xi^{p\theta/2}} \|\chi(w_h)\|_{L^1}^{p(1-\theta)} \|\chi(w_h)\|_{H^1, \mathcal{M}}^{p\theta}.$$

We estimate

$$\|\chi(w_h)\|_{L^1} \leq \sum_{\substack{K \in \mathcal{V} \\ |w_K| > N}} |K| |w_K| \leq \frac{1}{\ln N} \sum_{\substack{K \in \mathcal{V} \\ |w_K| > N}} |K| |w_K| \ln |w_K| \leq \frac{1}{\ln N} \|w_h \log |w_h|\|_{L^1}$$

and since $\left| \frac{\chi(w_L) - \chi(w_K)}{w_L - w_K} \right| \leq 2$ and $|\chi(w_h)| \leq |w_h|$ we deduce

$$\|\chi(w_h)\|_{H^1, \mathcal{M}}^2 = \|\chi(w_h)\|_{L^2}^2 + \|\chi(w_h)\|_{H^1, \mathcal{M}}^2 \leq \|w_h\|_{L^2}^2 + 4\|w_h\|_{H^1, \mathcal{M}}^2 \leq 4\|w_h\|_{H^1, \mathcal{M}}^2$$

and therefore $\|\chi(w_h)\|_{H^1, \mathcal{M}} \leq 2\|w_h\|_{H^1, \mathcal{M}}$. All together we find

$$\|w_h\|_{L^p}^p \leq c_\varepsilon \|w_h\|_{L^1} + \frac{\varepsilon}{\xi^{p\theta/2}} \|w_h \ln |w_h|\|_{L^1}^{p(1-\theta)} \|w_h\|_{H^1, \mathcal{M}}^{p\theta}$$

with $c_{\varepsilon, p} = c_p (2N)^{p-1}$ and $\varepsilon = \frac{c_p C^p 2^{p\theta}}{(\ln N)^{(1-\theta)p}}$. Especially for $d = 2$ it results $\theta = (p-1)/p$ and the inequality

$$\|w_h\|_{L^p}^p \leq c_\varepsilon \|w_h\|_{L^1} + \frac{\varepsilon}{\xi^{(p-1)/2}} \|w_h \ln |w_h|\|_{L^1}^p \|w_h\|_{H^1, \mathcal{M}}^{p-1}$$

follows. □

A.3 Technical lemmas

A.3.1 Global existence proof

In this part we collect some auxiliary results, which we use in Section 3.8 and in Section 3.10.

Lemma C.1. *Let $x, y, p \in \mathbb{R}$, $x, y > 0$.*

1. *For $p \geq 2$ the following inequalities hold:*

$$\frac{4(p-1)}{p^2} (x^{p/2} - y^{p/2})^2 \leq (x-y)(x^{p-1} - y^{p-1}) \leq (x^{p/2} - y^{p/2})^2. \quad (\text{C.26})$$

2. *For $p \geq 1$, we have*

$$\frac{1}{p}(x^p - y^p) \leq x^{p-1}(x - y). \quad (\text{C.27})$$

3. *Finally, for $p \geq 2$ the inequalities*

$$\frac{2}{p^2}(x^{p/2} - y^{p/2})^2 \leq (x^{p-2} + y^{p-2})(x - y)^2 \leq 2(x^{p/2} - y^{p/2})^2 \quad (\text{C.28})$$

are fulfilled.

Proof. 1. For $z \geq 1$, we consider the function

$$f(z) = (z-1)(z^{p-1} - 1) - \frac{4(p-1)}{p^2}(z^{p/2} - 1)^2.$$

The first and second derivatives of f are given by

$$\begin{aligned} \frac{d}{dz}f(z) &= \frac{(p-2)^2}{p}z^{p-1} - (p-1)z^{p-2} + \frac{4(p-1)}{p}z^{p/2-1} - 1, \\ \frac{d^2}{dz^2}f(z) &= \frac{(p-2)(p-1)}{p} \left(((p-2)z - p)z^{p-3} + 2z^{p/2-2} \right). \end{aligned}$$

It is easy to see that $f(1) = 0$, $f'(1) = 0$ and $f''(1) = 0$. Further, we deduce from $f''(z) > 0$ for $z > 1$ that $f'(z) > 0$ and $f(z) > 0$. W.l.o.g. let $x \geq y > 0$ otherwise swap the variables x and y . Setting $z = x/y$, we find

$$0 \leq f(x/y) = (x-y)(x^{p-1} - y^{p-1}) - \frac{4(p-1)}{p^2} (x^{p/2} - y^{p/2})^2$$

and finally by using Muirhead's inequality

$$\left(x^{p/2} - y^{p/2} \right)^2 - (x-y)(x^{p-1} - y^{p-1}) = x^{p-1}y + xy^{p-1} - 2x^{p/2}y^{p/2} \geq 0$$

holds. The case $y = 0$ is trivial.

2. For the second statement we consider for $z \geq 1$ the function

$$f(z) = \frac{p-1}{p}z^p - z^{p-1} + \frac{1}{p}.$$

Since $f(1) = 0$, the first derivative of f

$$\frac{d}{dz}f(z) = (p-1)z^{p-2}(z-1) \geq 0$$

implies $f(z) \geq 0$. Setting $z = x/y$, $x \geq y > 0$ we find

$$0 \leq y^p f(x/y) = \left(\frac{p-1}{p} \frac{x^p}{y^p} - \frac{x^{p-1}}{y^{p-1}} + \frac{1}{p} \right) y^p = x^{p-1}(x-y) - \frac{1}{p}(x^p - y^p).$$

For $x \leq y$ it results

$$\frac{1}{p}(x^p - y^p) \leq (x^p - y^p) \leq x^{p-1}(x - y(y/x)^{p-1}) \leq x^{p-1}(x - y).$$

3. We use the function

$$f(z) = (z^{p-2} + 1)(z-1)^2 - \frac{2}{p^2}(z^{p/2} - 1)^2.$$

The first and second derivatives are given by

$$\begin{aligned} \frac{d}{dz}f(z) &= 2(z-1)(z^{p-2} + 1) + (p-2)(z-1)^2z^{p-3} - \frac{2}{p}(z^{p/2} - 1)z^{p/2-1}, \\ \frac{d^2}{dz^2}f(z) &= \frac{1}{p} \left(2p + (p-2)z^{p/2-2} + z^{p-4}g(z) \right) \end{aligned}$$

with

$$\begin{aligned} g(z) &= (p-3)(p-2)p - 2(p-2)(p-1)pz + (p-1)(p^2-2)z^2, \\ g'(z) &= 2(p-1)((p^2-2)z - (p-2)p), \\ g''(z) &= 2(p-1)(p^2-2). \end{aligned}$$

From $g''(z) > 0$ for $z > 1$ and $p \geq 2$ we see that $g(z)$ is a convex function. Furthermore $f(z)$ is convex since $g'(1) = 4(p-1)^2 > 0$, $g(1) = 2$ and $f''(z) > 0$. Using $f'(1) = f(1) = 0$ we get $f(z) > 0$ and with $z = x/y$, $x \geq y > 0$ the first inequality of (C.28).

The last assertion follows from

$$(x^{p/2} - y^{p/2})^2 - \frac{1}{2}(x^{p-2} + y^{p-2})(x-y)^2 = \frac{1}{2}(x^p + y^p - x^{p-2}y^2 - x^2y^{p-2}) + (x^{p-1}y + xy^{p-1} - 2x^{p/2}y^{p/2})$$

together with Muirhead's inequality for the term

$$x^p + y^p \geq x^{p-2}y^2 + x^2y^{p-2}, \quad x^{p-1}y + xy^{p-1} \geq 2x^{p/2}y^{p/2}.$$

□

Lemma C.2. *Let x be a real number. Then*

$$f(x) = \frac{(e^x - 1)(e^{-x} + 1)}{2x} \geq 1, \quad g(x) = \frac{(e^x - 1)(e^{-x} - 1)}{x^2} \leq -1$$

hold. We define the value of the functions at $x = 0$ as limit $x \rightarrow 0$.

Proof. Using the power series of $\sinh(x)$ and $\cosh(x)$ we obtain

$$f(x) = 1 + \sum_{i=1}^{\infty} \frac{x^{2i}}{(2i+1)!}, \quad g(x) = -1 - 2 \sum_{i=2}^{\infty} \frac{x^{2i-2}}{(2i)!}.$$

□

A.3.2 Convergence proof

Lemma C.3. *Let $\sup_{\mathcal{D}} \|a_l\|_{L^2(S, H^1(\Omega))} < \infty$ and the assumptions of Theorem 3.12.3 be fulfilled. Then*

$$\|a_l - a_h\|_{L^2(S, L^2)} \rightarrow 0 \tag{C.29}$$

as size $\mathcal{D} \rightarrow 0$.

Proof. We decompose the integral over the whole domain into the intersected area of the control volume K and the kite D_σ (see Subsection 3.12.2) over the edge $\sigma = K|L$

$$\begin{aligned} T_1 &:= \|a_l - a_h\|_{L^2(S, L^2(\Omega))}^2 = \int_S \int_{\Omega} (a_l(x) - a_h(x))^2 dx dt \\ &= \int_S \sum_{K \in \mathcal{V}} \sum_{L \in \mathcal{N}_{\mathcal{V}}(K)} \int_{K \cap D_\sigma} (a_l(x) - a_K)^2 dx dt. \end{aligned}$$

Using the Taylor expansion of a_l around x , i.e. $a_l(x_K) - a_l(x) = \nabla a_l(x)(x_K - x)$ and the boundedness of the gradient we get

$$\begin{aligned} T_1 &\leq \text{size}(\mathcal{M})^2 \int_S \sum_{K \in \mathcal{V}} \sum_{L \in \mathcal{N}_{\mathcal{V}}(K)} \int_{K \cap D_\sigma} |\nabla a_l|^2 dx dt \\ &\leq \text{size}(\mathcal{M})^2 \|a_l\|_{L^2(S, H^1(\Omega))} \rightarrow 0 \end{aligned}$$

as $\text{size}(\mathcal{D}) \rightarrow 0$. □

We repeat the definition of the weak gradient reconstruction operator $\nabla_w : X_{\mathcal{V}}(\mathcal{M}) \rightarrow L^2(\Omega)^2$ (see (3.12.67)), i.e.

$$\nabla_w a_l(\mathbf{x}) = 2 \frac{a_L - a_K}{d_\sigma} \mathbf{n}_\sigma \quad \text{for } \mathbf{x} \in D_\sigma, \sigma = K|L.$$

Lemma C.4. *Let \mathcal{D} be a sequence of discretization of $Q = S \times \Omega$ and a_l in $L^2(S, H^1(\Omega))$ a corresponding sequence of piecewise linear functions with $\|a_l\|_{L^2(S, H^1(\Omega))} \leq C$. Then a subsequence (also called a_l) can be extracted which converges to $\hat{a} \in L^2(S, H^1(\Omega))$ in the following senses*

$$a_l \rightharpoonup \hat{a} \text{ in } L^2(S, H^1(\Omega)), \quad \nabla_w a_l \rightharpoonup \nabla \hat{a} \text{ in } L^2(S, L^2(\Omega)^2).$$

Proof. Since $\|a_l\|_{L^2(S, H^1)}$ is uniformly bounded, the weak gradient reconstruction of a_l is also uniformly bounded in $L^2(S, L^2(\Omega)^2)$. Therefore there exist a subsequence and a $\hat{g} \in L^2(S, L^2(\Omega)^2)$ such that $\nabla_w a_l \rightharpoonup \hat{g}$ in $L^2(S, L^2(\Omega)^2)$. We show that $\hat{g} = \nabla \hat{a}$ holds.

In the following we use arbitrary test functions $\chi \in C_0^\infty(S)$ and $w \in C_0^\infty(\Omega)^2$. The set $\{g : g = \sum_{i=1}^n \chi_i w_i \text{ with } w_i \in C_0^\infty(\Omega)^2, \chi_i \in C_0^\infty(S)\}$ is dense in $L^2(S, L^2(\Omega)^2)$. According to the Helmholtz decomposition, the set $\{w : w = \nabla \varphi_c + \text{curl } \varphi_r \text{ with } \varphi_c, \varphi_r \in C_0^\infty(\Omega)\}$ is dense in $L^2(\Omega)^2$.

For all $\psi \in C_0^\infty(\Omega)$ we denote by ψ_l the Lagrange interpolant of ψ consisting in a continuous piecewise affine function such that for all nodes x_K in the mesh one obtains $\psi_l(x_K) = \psi(x_K)$. Due to the regularity of the mesh, one gets from classical FEM theory that $\psi_l \rightarrow \psi$ in $H^1(\Omega)$. Moreover, we introduce a strong gradient interpolation operator by

$$\nabla_s \psi(\mathbf{x}) := \frac{\psi(x_L) - \psi(x_K)}{d_\sigma} \mathbf{n}_\sigma + \frac{\psi(\mathbf{x}_{T_{\sigma K}}) - \psi(\mathbf{x}_{T_{\sigma L}})}{m_\sigma} \mathbf{t}_\sigma \quad \text{for } \mathbf{x} \in D_\sigma, \sigma = K|L,$$

where $\mathbf{x}_{T_{\sigma K}}$ and $\mathbf{x}_{T_{\sigma L}}$ denote the circumcenters of the triangles to the left and to the right of $\overline{x_K, x_L}$. Furthermore, the continuous curl operator is defined by a $(-\frac{\pi}{2})$ rotation of the gradient, i.e., one has

$$\text{curl } \psi = \begin{pmatrix} \psi_y \\ -\psi_x \end{pmatrix} = - \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix}^\perp = -(\nabla \psi)^\perp.$$

Therefore, the strong curl interpolator is defined by

$$\operatorname{curl}_s \psi(\mathbf{x}) := -\frac{\psi(x_L) - \psi(x_K)}{d_\sigma} \mathbf{t}_\sigma + \frac{\psi(\mathbf{x}_{T_{\sigma K}}) - \psi(\mathbf{x}_{T_{\sigma L}})}{m_\sigma} \mathbf{n}_\sigma \quad \text{for } \mathbf{x} \in D_\sigma, \sigma = K|L.$$

Using the Helmholtz decomposition we obtain

$$\widehat{g}(t) = \nabla g_r(t) + \operatorname{curl} g_c(t) \quad g_r(t), g_c(t) \in H^1(\Omega) \text{ f.a.a. } t \in S.$$

Irrotational part: First we show that the irrotational part of \widehat{g} in the sense of Helmholtz decomposition equals $\nabla \widehat{a}$. Due to the orthogonality of \mathbf{t}_σ and \mathbf{n}_σ we have

$$\begin{aligned} S &= \int_S \chi \int_\Omega \nabla a_l \cdot \nabla \varphi_{r,l} dx dt \\ &= \int_S \chi \sum_{\sigma=K|L \in \mathcal{E}_{int}} \frac{m_\sigma}{d_\sigma} (a_K - a_L) (\varphi_r(x_K) - \varphi_r(x_L)) dt \\ &= \int_S \chi \sum_{\sigma=K|L \in \mathcal{E}_{int}} |D_\sigma| \left(2 \frac{a_K - a_L}{d_\sigma} \right) \left(\frac{\varphi_r(x_K) - \varphi_r(x_L)}{d_\sigma} \right) dt \\ &= \int_S \chi \int_\Omega \nabla_w a_l \cdot \nabla_s \varphi_{r,l} dx dt. \end{aligned}$$

By weak-strong convergence, we obtain in the first and the last integral

$$\int_S \chi \int_\Omega \nabla \widehat{a} \cdot \nabla \varphi_r dx dt = \int_S \chi \int_\Omega \widehat{g} \cdot \nabla \varphi_r dx dt$$

and therefore $\nabla \widehat{g}_r(t, x) = \nabla \widehat{a}(t, x)$ f.a.a. $(t, x) \in Q$.

Divergence-free part: It remains to show that the divergence-free part of \widehat{g} vanishes. In [107] it is shown that covolume discretizations on Delaunay-Voronoi grids fulfill the continuous property $\nabla \cdot \operatorname{curl} \psi \equiv 0$ in a discrete sense, which will be used in the following. For the sake of completeness,

we present the necessary arguments. Obviously, we obtain for the strong curl interpolator on regular meshes that $\text{curl}_s \psi \rightarrow \text{curl} \psi$ in $L^2(\Omega)^2$. Now we compute

$$\begin{aligned}
 & \int_Q \chi \nabla_w a_l \cdot \text{curl}_s \varphi_c \, dx \, dt \\
 &= \int_S \chi \sum_{\sigma=K|L \in \mathcal{E}_{int}} |D_\sigma| \left(2 \frac{a_L - a_K}{d_\sigma} \right) \left(\frac{\varphi_c(\mathbf{x}_{T_{\sigma K}}) - \varphi_c(\mathbf{x}_{T_{\sigma L}})}{m_\sigma} \right) dt \\
 &= \int_S \chi \sum_{\sigma=K|L \in \mathcal{E}_{int}} (a_L - a_K) (\varphi_c(\mathbf{x}_{T_{\sigma K}}) - \varphi_c(\mathbf{x}_{T_{\sigma L}})) dt \\
 &= \int_S \chi \sum_{\substack{L \in \mathcal{V} \\ x_L \notin \partial\Omega}} \sum_{K \in \mathcal{N}_V(L)} a_L (\varphi_c(\mathbf{x}_{T_{\sigma K}}) - \varphi_c(\mathbf{x}_{T_{\sigma L}})) dt \\
 &= 0.
 \end{aligned}$$

The last integral is always identically 0, since the term $\varphi_c(\mathbf{x}_{T_{\sigma K}})$ is added two times with opposite sign. By weak-strong convergence of the first integral, we finally obtain

$$\int_Q \chi \hat{g} \cdot \text{curl} \varphi_c \, dx \, dt = 0,$$

i.e., the divergence-free part of \hat{g} vanishes, and we have indeed $\hat{g}(t, x) = \nabla \hat{a}(t, x)$ f.a.a. $(t, x) \in Q$. \square

Lemma C.5. *Let Ω be an open, bounded, polyhedral subset of \mathbb{R}^d . Moreover, let $\{\mathcal{M}\}$ be a series of Voronoi finite volume meshes with size $(\mathcal{M}) \rightarrow 0$. Let be $u \in L^\infty(\Omega)$ and $u(x) \geq 0$ f.a.a. $x \in \Omega$. Then the piecewise constant functions u_h with $u_h(x) = u_K$ if $x \in K$,*

$$u_K = \frac{1}{|K|} \int_K u(x) \, dx \quad \forall K \in \mathcal{V},$$

converge to u in $L^p(\Omega)$ for all $p \in [1, \infty)$ as $\text{size}(\mathcal{M}) \rightarrow 0$.

Proof. We show the result for $p = 1$, the assertion for $p \in (1, \infty)$ then results from

$$\|u_h - u\|_{L^p} \leq (2\|u\|_{L^\infty})^{\frac{p-1}{p}} \|u - u_h\|_{L^1}^{1/p}.$$

Let be $\varepsilon > 0$. Luzin's theorem [128, p. 1013] guarantees the existence of a closed set $M_\varepsilon \subset \Omega$ such that

$$u|_{M_\varepsilon} \text{ is continuous and } |\Omega \setminus M_\varepsilon| < \frac{\varepsilon}{8\|u\|_{L^\infty}}.$$

Since M_ε is compact there exists a $\delta > 0$ such that

$$|u(x) - u(y)| < \frac{\varepsilon}{2|\Omega|} \quad \forall x, y \in M_\varepsilon, |x - y| < \delta.$$

Let $\text{size}(\mathcal{M}) < \delta$ which implies $|x - y| < \delta$. Introducing the set

$$K \setminus M_\varepsilon := \{x \in \mathbb{R}^N : x \in K, x \notin M_\varepsilon\}$$

and by using the identity $K = (K \setminus M_\varepsilon) \cup (K \cap M_\varepsilon)$ we estimate the integral in the L^1 -norm by considering integrals on subsets

$$\begin{aligned} \|u - u_h\|_{L^1} &\leq \sum_{K \in \mathcal{V}} \frac{1}{|K|} \int_K \int_K |u(x) - u(y)| dy dx \\ &= \sum_{K \in \mathcal{V}} \frac{1}{|K|} \int_{K \cap M_\varepsilon} \int_{K \setminus M_\varepsilon} |u(x) - u(y)| dy dx \\ &\quad + \sum_{K \in \mathcal{V}} \frac{1}{|K|} \int_{K \cap M_\varepsilon} \int_{K \cap M_\varepsilon} |u(x) - u(y)| dy dx \\ &\quad + \sum_{K \in \mathcal{V}} \frac{1}{|K|} \int_{K \setminus M_\varepsilon} \int_K |u(x) - u(y)| dy dx \\ &< \frac{\varepsilon}{2|\Omega|} |\Omega| + 2\|u\|_{L^\infty} \frac{\varepsilon}{8\|u\|_{L^\infty}} + 2\|u\|_{L^\infty} \frac{\varepsilon}{8\|u\|_{L^\infty}} = \varepsilon. \end{aligned}$$

□

We recapitulate the Assumption A7: Let $f \in L^\infty(\Omega)$ with $f(x) \geq 0$ f.a.a. $x \in \Omega$ and let \mathcal{I} be a finite index set. Furthermore, let $\Omega \subset \mathbb{R}^2$ be a polygonal domain and let $\Omega = \cup_{I \in \mathcal{I}} \Omega_I$ be a finite disjoint union of subdomains such that the discontinuities of f coincide with the subdomain boundaries. Let the over all one-dimensional measure of all internal subdomain boundaries be bounded by θ . There exists some $\gamma \in (0, 1]$ such that $f \in C^{0,\gamma}(\Omega_I) := \{w|_{\Omega_I}, w \in C^{0,\gamma}(\mathbb{R}^2)\}$, $I \in \mathcal{I}$.

Lemma C.6. We assume Assumption A7. Let $\mathcal{D} = (\mathcal{M}, (t_n)_{n=1}^N)$ be a sequence of discretizations of $Q = S \times \Omega$ and $a_h \rightarrow \hat{a}$ in $L^2(S, Y)$ as $\text{size}(\mathcal{D}) \rightarrow 0$ with $\|a_h\|_{L^\infty(S, L^\infty(\Omega, \mathbb{R}^m))} \leq R$. Moreover, let $f : \Omega \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ satisfy the Carathéodory condition and there exist $0 < \underline{c}, \bar{c} < \infty$ such that $\underline{c} \leq f(x, y) \leq \bar{c}$, f.a.a. $x \in \Omega, \forall y \in \mathbb{R}^m$. Additionally, there exists a constant $\gamma \in (0, 1]$ such that $f|_{\Omega_I} \in C^{0,\gamma}(\Omega_I \times \mathcal{B}_{\mathbb{R}^m}(0, R))$ for all $I \in \mathcal{I}$, where $\mathcal{B}_{\mathbb{R}^m}(0, R)$ denotes a ball in \mathbb{R}^m centered at 0 with radius R . Then

$$\|f_h(\cdot, a_h) - f(\cdot, \hat{a})\|_{L^1(Q)} \rightarrow 0 \quad \text{as } \text{size}(\mathcal{D}) \rightarrow 0.$$

Proof. For any given discretization \mathcal{M} , let be

$$\Xi := \{x \in \Omega : x \in K \in \mathcal{V} \text{ with } |K \cap \Omega_I| \neq 0 \wedge |K \cap \Omega_J| \neq 0 \text{ for } I \neq J, I, J \in \mathcal{I}\}$$

with $|\Xi| \leq 2\theta \text{ size}(\mathcal{M})$ for all discretizations \mathcal{D} (see also Assumption A7). In order to estimate

$$\begin{aligned} S_1 &= \|f_h(\cdot, a_h) - f(\cdot, \widehat{a})\|_{L^1(\Omega)} = \sum_{K \in \mathcal{V}} \int_K \left| \frac{1}{|K|} \int_K f(y, a_h(x)) dy - f(x, \widehat{a}(x)) \right| dx \\ &\leq \sum_{K \in \mathcal{V}} \int_K S_2(x) dx \end{aligned}$$

with

$$S_2(x) = \frac{1}{|K|} \int_K |f(y, a_h(x)) - f(x, \widehat{a}(x))| dy$$

we consider the cases:

- On $x \in \Omega \setminus \Xi$, we obtain

$$S_2(x) \leq \frac{c}{|K|} \int_K (|a_h(x) - \widehat{a}(x)|^\gamma + |x - y|^\gamma) dy \leq c|a_h(x) - \widehat{a}(x)|^\gamma + c \text{ size}(\mathcal{M})^\gamma.$$

- On $x \in \Xi$, we find by using the bound \bar{c} of f the estimate $S_2(x) \leq 2\bar{c}$.

Therefore we deduce with some constant c_1 the following estimate

$$\begin{aligned} \int_S S_1 dt &\leq \int_S \left(\int_{\Omega \setminus \Xi} S_2 dx + \int_{\Xi} S_2 dx \right) dt \\ &\leq c_1 \text{ size}(\mathcal{M})^\gamma + c_1 \int_Q |a_h - \widehat{a}|^\gamma dx dt. \end{aligned}$$

Since $\gamma \in (0, 1]$, we apply Hölder's inequality with $q = 2/\gamma$ and $p = 2/(2 - \gamma)$ and conclude

$$\|f_h(\cdot, a_h) - f(\cdot, \widehat{a})\|_{L^1(Q)} \leq c_1 \text{ size}(\mathcal{M})^\gamma + c_1 (T|\Omega|)^{1-\gamma/2} \|a_h - \widehat{a}\|_{L^2(S,Y)} \rightarrow 0$$

as $\text{size}(\mathcal{D}) \rightarrow 0$. □

A.4 Convex analysis

This section rephrases results of convex analysis, which we need in Section 3.6 and 3.10. The results are mainly taken from [115, Sec. 1.1.3, Sec. 8.10], [27, 128] and [77, Sec. 0.3.2., Sec. 4.2.].

In this section let V be a (real) linear space.

A set A in a linear space is called *convex* if $\lambda u + (1 - \lambda)v \in A$ whenever $u, v \in A$ and $\lambda \in [0, 1]$. A function $f : V \rightarrow \mathbb{R}$ is called *convex* if $f(\lambda u + (1 - \lambda)v) \leq \lambda f(u) + (1 - \lambda)f(v)$ for any $u, v \in V$ and $\lambda \in [0, 1]$. If $u \neq v$ and $\lambda \in]0, 1[$ imply a strict inequality then f is called *strictly convex*. A function $f : V \rightarrow \mathbb{R}$ is convex iff its *epigraph* $\text{epi}(f) := \{(v, a) \in V \times \mathbb{R} : a \geq f(v)\}$ is a convex subset of $V \times \mathbb{R}$. A function $f : V \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ called *proper* iff its *domain* $\text{dom}(f) := \{v \in V : f(v) < +\infty\} \neq \emptyset$ and $f(x) > -\infty$ for all $x \in V$. A function $f : V \rightarrow \overline{\mathbb{R}}$ is said to be *lower (resp. upper) semicontinuous* if

$$\forall u \in V, u_k \rightarrow u : \quad f(u) \leq \liminf_{k \rightarrow \infty} f(u_k) \quad \left(\text{resp. } f(u) \geq \limsup_{k \rightarrow \infty} f(u_k) \right).$$

Let $f : V \rightarrow \overline{\mathbb{R}}$ be a proper convex function. The *subdifferential* ∂f of the function f is defined by

$$\partial f(u) := \{v^* \in V^* : f(v) - f(u) \geq \langle v^*, v - u \rangle \quad \forall v \in V\}. \quad (\text{C.30})$$

An element $v^* \in V^*$ satisfying $v^* \in \partial f(u)$ is said to be a *subgradient* of the function f at a point u . If f is a proper convex function on \mathbb{R}^n then f is subdifferentiable in every relative interior point, see [77, Sec. 4.2.3]. Let $f : V \rightarrow \overline{\mathbb{R}}$ be convex. The *conjugate functional* $f^* : V^* \rightarrow \overline{\mathbb{R}}$ of f is defined by

$$f^*(v^*) = \sup_{v \in V} \{\langle v^*, v \rangle - f(v)\}. \quad (\text{C.31})$$

The transformation $f \mapsto f^*$ is called the *Legendre transformation* in the smooth case, or the *Legendre-Fenchel transformation* in the general case. From the definition (C.31) it follows that f^* is convex, lower semicontinuous and $f^*(v^*) + f(v) \geq \langle v^*, v \rangle$ which is called *Fenchel's inequality*. Furthermore $f^{**} = f$ iff f lower semicontinuous. If f is a proper and lower semicontinuous function, then

$$v^* \in \partial f(v) \quad \Leftrightarrow \quad v \in \partial f^*(v^*) \quad \Leftrightarrow \quad f(v) + f^*(v^*) = \langle v^*, v \rangle.$$

Bibliography

- [7] B. Andreianov, M. Bendahmane, and M. Saad. “Finite volume methods for degenerate chemotaxis model.” In: *J. Comput. Appl. Math.* 235.14 (2011), pp. 4015–4031.
- [8] D. M. Beazley. “SWIG: an easy to use tool for integrating scripting languages with C and C++.” In: *Proceedings of the 4th conference on USENIX Tcl/Tk Workshop, 1996 - Volume 4. TCLKT’96*. Monterey, California: USENIX Association, 1996, pp. 15–15.
- [9] M. Bessemoulin-Chatard. “A finite volume scheme for convection–diffusion equations with nonlinear diffusion derived from the Scharfetter–Gummel scheme.” In: *Numer. Math.* 121.4 (2012), pp. 637–670.
- [10] M. Bessemoulin-Chatard, C. Chainais-Hillairet, and F. Filbet. “On discrete functional inequalities for some finite volume schemes.” In: *Preprint arXiv:1202.4860v2* (2014). Accepted for publication in *IMA J. Numer. Anal.*
- [11] D. Bothe and D. Hilhorst. “A reaction–diffusion system with fast reversible reaction.” In: *J. Math. Anal. Appl.* 286.1 (2003), pp. 125–135.
- [12] D. Bothe and M. Pierre. “Quasi-steady-state approximation for a reaction-diffusion system with fast intermediate.” In: *J. Math. Anal. Appl.* 368.1 (2010), pp. 120–132.
- [13] D. Bothe and M. Pierre. “The instantaneous limit for reaction-diffusion systems with a fast irreversible reaction.” In: *Discrete Contin. Dyn. Syst. Ser. S* 5.1 (2012), pp. 49–59.
- [14] D. Bothe, M. Pierre, and G. Rolland. “Cross-Diffusion Limit for a Reaction-Diffusion System with Fast Reversible Reaction.” In: *Commun. Partial Differ. Equations* 37.11 (2012), pp. 1940–1966.
- [15] N. Bouillard, R. Eymard, R. Herbin, and P. Montarnal. “Diffusion with dissolution and precipitation in a porous medium: mathematical analysis and numerical approximation of a simplified model.” In: *ESAIM, Math. Model. Numer. Anal.* 41.6 (2007), pp. 975–1000.
- [16] G. E. Briggs and J. B. S. Haldane. “A note on the kinetics of enzyme action.” In: *Biochem. J.* 19.2 (1925), p. 338.
- [17] A. J. Brown. “Enzym action.” In: *J. Chem. Soc.* 81 (1902), pp. 373–386.

-
- [18] M. Bukal, E. Emmrich, and A. Jüngel. *Entropy-stable and entropy-dissipative approximations of a fourth-order quantum diffusion equation*. Version 1. Aug. 27, 2012.
- [19] J. Carberry. *Chemical and Catalytic Reaction Engineering*. Dover Books on Chemistry Series. Dover, 2001.
- [20] M. Chipot, K. D. D. Hilhorst, and M. Olech. “Contraction in L^1 for a system arising in chemical reactions and molecular motors.” In: *Differ. Equ. Appl.* 1.1 (2009), pp. 139–151.
- [21] S. Chou and S. Tang. “Conservative $P1$ conforming and nonconforming Galerkin FEMs: effective flux evaluation via a nonmixed method approach.” In: *SIAM J. Numer. Anal.* 38.2 (2000), pp. 660–680.
- [22] L. Desvillettes and K. Fellner. “Exponential decay toward equilibrium via entropy methods for reaction-diffusion equations.” In: *J. Math. Anal. Appl.* 319.1 (2006), pp. 157–176.
- [23] L. Desvillettes, K. Fellner, M. Pierre, and J. Vovelle. “Global existence for quadratic systems of reaction-diffusion”. Anglais. In: *advanced nonlinear studies* 7.3 (2007), pp. 491–511.
- [24] P. Deuffhard and F. Bornemann. *Gewöhnliche Differentialgleichungen*. Numerische Mathematik Bd. 2. Germany: de Gruyter, 2008.
- [25] R. Doe. *Intel pushes 193-nm litho down to 15-nm*. June 2009. url: http://www.eetimes.com/document.asp?doc_id=1171202.
- [26] M. Ederer and E. Gilles. “Thermodynamically Feasible Kinetic Models of Reaction Networks”. In: *Biophys J.* 92.6 (2007), pp. 1846–1857.
- [27] I. Ekeland and R. Temam. *Convex analysis and variational problems*. North-Holland, Amsterdam, 1976.
- [28] E. Emmrich. *Gewöhnliche und Operator-Differentialgleichungen. Eine integrierte Einführung in Randwertprobleme und Evolutionsgleichungen für Studierende*. Wiesbaden Vieweg, 2004.
- [29] E. Emmrich. “Convergence of the variable two-step BDF time discretisation of nonlinear evolution problems governed by a monotone potential operator.” English. In: *BIT* 49.2 (2009), pp. 297–323.
- [30] A. Erdmann, F. Shao, J. Fuhrmann, A. Fiebach, G. P. Patis, and P. Trefonas. “Modeling of double patterning interactions in litho-cure-litho-etch (LCLE) processes”. In: *Proc. of SPIE* 7640.1 (2010), 76400B.
- [31] R. Eymard, J. Fuhrmann, and K. Gärtner. “A finite volume scheme for nonlinear parabolic equations derived from one-dimensional local Dirichlet problems.” In: *Numer. Math.* 102.3 (2006), pp. 463–495.
- [32] R. Eymard, T. Gallouët, and R. Herbin. “The Finite Volume Method”. In: *Handbook of Numerical Analysis*. Ed. by P. Ciarlet and J.-L. Lions. North Holland, 2000, pp. 713–1020.

- [33] R. Eymard, A. Handlovicova, R. Herbin, K. Mikula, and O. Stašová. *Applications of approximate gradient schemes for nonlinear parabolic equations*. Tech. rep. 2012.
- [34] R. Eymard, D. Hilhorst, H. Murakawa, and M. Olech. “Numerical approximation of a reaction-diffusion system with fast reversible reaction”. In: *Chin. Ann. Math., Ser. B* (2010), pp. 1–24.
- [35] A. Fiebach, A. Glitzky, and A. Linke. *Convergence of an implicit Voronoi finite volume method for reaction-diffusion problems*. Preprint 1827. WIAS, 2013.
- [36] A. Fiebach, A. Glitzky, and A. Linke. “Uniform global bounds for solutions of an implicit Voronoi finite volume method for reaction–diffusion problems”. In: *Numer. Math.* 128.1 (2014), pp. 31–72.
- [37] P. C. Franzone, P. Deuffhard, B. Erdmann, J. Lang, and L. F. Pavarino. “Adaptivity in space and time for reaction-diffusion systems in electrocardiology.” English. In: *SIAM J. Sci. Comput.* 28.3 (2006), pp. 942–962.
- [38] T. Fühner, T. Schnattinger, G. Ardelean, and A. Erdmann. “Dr.LiTHO – a development and research lithography simulator”. In: *Proc. of SPIE* 6520 (2007), 65203F–1.
- [39] J. Fuhrmann, A. Fiebach, A. Erdmann, and P. Trefonas. “Acid diffusion effects between resists in freezing processes used for contact hole patterning”. In: *Microelectron. Eng.* 87.5–8 (2010), pp. 951–954.
- [40] J. Fuhrmann, A. Fiebach, and G. P. Patsis. “Macroscopic and stochastic modeling approaches to pattern doubling by acid catalyzed cross-linking”. In: *Proc. of SPIE* 7639, 76392I (2010), p. 76392I.
- [41] J. Fuhrmann, A. Fiebach, M. Uhle, A. Erdmann, C. R. Szmanda, and C. Truong. “A model of self-limiting residual acid diffusion for pattern doubling”. In: *Microelectron. Eng.* 86.4–6 (2009), pp. 792–795.
- [42] H. Gajewski. “On the time discretization of the carrier transport equations”. In: ed. by H. Gajewski, P. Deuffhard, and P. A. Markowich. 1991, pp. 67–68.
- [43] H. Gajewski and K. Gärtner. “On the discretization of van Roosbroeck’s equations with magnetic field.” In: *Z. Angew. Math. Mech.* 76.5 (1996), pp. 247–264.
- [44] H. Gajewski and K. Gärtner. “A dissipative discretization scheme for a nonlocal phase segregation model”. In: *ZAMM, Z. Angew. Math. Mech.* 85.11 (2005), pp. 815–822.
- [45] H. Gajewski and K. Gröger. “Initial-boundary value problems modelling heterogeneous semiconductor devices”. In: *Surveys on analysis, geometry and mathematical physics*. Vol. 117. Teubner-Texte Math. Leipzig: Teubner, 1990, pp. 4–53.

- [46] H. Gajewski and K. Gröger. “Reaction—Diffusion Processes of Electrically Charged Species”. In: *Math. Nachr.* 177.1 (1996), pp. 109–130.
- [47] H. Gajewski and I. V. Skrypnik. “Existence and uniqueness results for reaction-diffusion processes of electrically charged species”. In: *Nonlinear Elliptic and Parabolic Problems (Zurich 2004)*. Vol. 64. Progress in Nonlinear Differential Equations and Their Applications. Basel: Birkhäuser, 2005, pp. 151–188.
- [48] K. Gärtner. “Existence of bounded discrete steady-state solutions of the van Roosbroeck system on boundary conforming Delaunay grids.” In: *SIAM J. Sci. Comput.* 31.2 (2009), pp. 1347–1362.
- [49] K. Gärtner. “Charge Transport in Semiconductors and a Finite Volume Scheme”. In: *Finite Volumes for Complex Applications VI Problems & Perspectives*. Ed. by J. Fořt, J. Fürst, J. Halama, R. Herbin, and F. Hubert. Vol. 4. Springer Proceedings in Mathematics. Springer Berlin Heidelberg, 2011, pp. 513–522.
- [50] R. S. Ghaida, K. B. Agarwal, L. W. Liebmann, S. R. Nassif, and P. Gupta. “A novel methodology for triple/multiple-patterning layout decomposition”. In: *Proc. of SPIE* 8327 (2012), pages.
- [51] V. Giovangigli. *Multicomponent flow modeling*. Modeling and simulation in science, engineering & technology. Birkhäuser, 1999.
- [52] V. Girault and P. A. Raviart. *Finite element methods for Navier-Stokes equations*. Vol. 5. Springer Series in Computational Mathematics. Theory and algorithms. Berlin: Springer-Verlag, 1986, pp. x+374.
- [53] A. Glitzky. *Elektro-Reaktions-Diffusionssysteme mit nichtglatten Daten*. Logos Verlag Berlin, 2001.
- [54] A. Glitzky. “Exponential decay of the free energy for discretized electro-reaction-diffusion systems.” In: *Nonlinearity* 21.9 (2008), pp. 1989–2009.
- [55] A. Glitzky. “Energy estimates for electro-reaction-diffusion systems with partly fast kinetics.” English. In: *Discrete Contin. Dyn. Syst.* 25.1 (2009), pp. 159–174.
- [56] A. Glitzky. “Uniform exponential decay of the free energy for Voronoi finite volume discretized reaction-diffusion systems.” In: *Math. Nachr.* 284.17-18 (2011), pp. 2159–2174.
- [57] A. Glitzky and K. Gärtner. “Energy estimates for continuous and discretized electro-reaction-diffusion systems.” In: *Nonlinear Anal., Theory Methods Appl., Ser. A* 70.2 (2009), pp. 788–805.
- [58] A. Glitzky and K. Gärtner. “Existence of bounded steady state solutions to spin-polarized drift-diffusion systems.” In: *SIAM J. Math. Anal.* 41.6 (2010), pp. 2489–2513.

- [59] A. Glitzky and J. Griepentrog. “Discrete Sobolev–Poincaré Inequalities for Voronoi Finite Volume Approximations”. In: *SIAM J. Numer. Anal.* 48.1 (2010), pp. 372–391.
- [60] A. Glitzky and J. A. Griepentrog. “Discrete Sobolev-Poincaré inequalities for Voronoi finite volume approximations.” English. In: *SIAM J. Numer. Anal.* 48.1 (2010), pp. 372–391.
- [61] A. Glitzky, K. Gröger, and R. Hünlich. “Free energy and dissipation rate for reaction diffusion processes of electrically charged species.” In: *Appl. Anal.* 60.3-4 (1996), pp. 201–217.
- [62] A. Glitzky and R. Hünlich. “Energetic estimates and asymptotics for electro-reaction-diffusion systems.” In: *Z. Angew. Math. Mech.* 77.11 (1997), pp. 823–832.
- [63] A. Glitzky and R. Hünlich. “Global estimates and asymptotics for electro-reaction-diffusion systems in heterostructures.” In: *Appl. Anal.* 66.3-4 (1997), pp. 205–226.
- [64] A. Glitzky and R. Hünlich. *Electro–reaction–diffusion systems in heterostructures*. 19. WIAS-Report, 2000.
- [65] A. Glitzky and R. Hünlich. “Electro–reaction–diffusion systems including cluster reactions of higher order”. In: *Math. Nachr.* 216 (2000), pp. 95–118.
- [66] A. Gorban and G. Yablonsky. “Extended Detailed Balance for Systems with Irreversible Reactions”. In: *Chem. Eng. Sci.* 66 (2011), pp. 5388–5399.
- [67] K. Gröger. *Free energy estimates and asymptotic behaviour of reaction–diffusion processes*. Preprint 20. Institut für Angewandte Analysis und Stochastik im Forschungsverbund Berlin e.V., 1992.
- [68] S. R. d. Groot and P. Mazur. *Non-equilibrium thermodynamics*. Amsterdam, 1969.
- [69] C. Grossmann and H. Roos. *Numerical treatment of partial differential equations*. Universitext. Translated and revised from the 3rd (2005) German edition by Martin Stynes. Berlin: Springer, 2007, pp. xii+591.
- [70] E. Hairer, S. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff problems*. Springer Ser. Comp. Mathem. Bd. 1. Germany: Springer, 1993.
- [71] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*. Springer Ser. Comp. Mathem. Bd. 2. Germany: Springer, 2004.
- [72] R. N. Hall. “Electron-Hole Recombination in Germanium”. In: *Phys. Rev.* 87 (2 1952), pp. 387–387.
- [73] W. Heineken. “Adaptive Verfahren zur numerischen Berechnung von Reaktions-Diffusions-Systemen”. PhD thesis. Otto-von-Guericke-Universität Magdeburg, 2005.
- [74] W. Henke and M. Torkler. “Modeling of edge roughness in ion projection lithography”. In: *J. Vac. Sci. Technol. B* 17.6 (1999), pp. 3112–3118.

- [75] D. Hilhorst and M. Vohralík. “A posteriori error estimates for combined finite volume–finite element discretizations of reactive transport equations on nonmatching grids”. In: *Comput. Methods Appl. Mech. Eng.* 200.5–8 (2011), pp. 597–613.
- [76] *International technology roadmap for semiconductors*. 2012. url: <http://www.itrs.net>.
- [77] A. Ioffe and V. Tihomirov. *Theory of extremal problems*. Studies in Mathematics and its Applications, Vol. 6. Amsterdam, New York, Oxford: North-Holland Publishing Company. XII, 460, 1979.
- [78] Jost, J. *Partial differential equations. 3rd revised and expanded ed.* Graduate Texts in Mathematics 214. New York, NY: Springer. xiii, 2013.
- [79] A. Jüngel and D. Matthes. “Entropiemethoden für nichtlineare partielle Differentialgleichungen”. In: *Int. Math. Nachr., Wien* 209 (2008), pp. 1–14.
- [80] L. Kalachev, H. Kaper, T. Kaper, N. Popovic, and A. Zagaris. “Reduction for Michaelis-Menten-Henri kinetics in the presence of diffusion.” In: *Electron. J. Differ. Equ.* Conference 16 (2007), pp. 155–184.
- [81] T. Koprucki and K. Gärtner. “Discretization scheme for drift-diffusion equations with a generalized Einstein relation”. In: *WIAS Preprint* 1738 (2012).
- [82] S. Kräutle. “General Multi-Species Reactive Transport Problems in Porous Media: Efficient Numerical Approches and Existence of Global Solutions”. PhD thesis. University of Erlangen-Nueremberg, 2008.
- [83] A. Kremling. *Kompndium Systembiologie Mathematische Modellierung und Modellanalyse*. Wiesbaden: Vieweg+Teubner Verlag, 2012.
- [84] A. Kufner, O. John, O. John, and S. Fučík. *Function spaces*. Monographs and textbooks on mechanics of solids and fluids: Mechanics, analysis. Noordhoff International Publishing, 1977.
- [85] O. Ladyzhenskaya, V. Solonnikov, and N. Ural'tseva. *Linear and quasi-linear equations of parabolic type*. Translated from the Russian by S. Smith. Translations of Mathematical Monographs. 23. Providence, RI: American Mathematical Society (AMS). XI, 648 p., 1968.
- [86] L. D. Landau and E. M. Lifschitz. *Lehrbuch der theoretischen Physik: Statistische Physik*. Berlin: Akademie-Verlag, 1984.
- [87] J. Lang. *Adaptive multilevel solution of nonlinear parabolic PDE systems. Theory, algorithm, and applications*. English. Berlin: Springer, 2001, pp. xii + 157.
- [88] J. Lang. “Two-Dimensional Fully Adaptive Solutions of Reaction-Diffusion Equations”. In: *Appl. Numer. Math* 18 (1995), pp. 223–240.

- [89] J. Lang. “High-resolution self-adaptive computations on chemical reaction-diffusion problems with internal boundaries”. In: *Chem. Eng. Sci.* 51.7 (1996), pp. 1055–1070.
- [90] M. Liero and A. Mielke. “Gradient structures and geodesic convexity for reaction-diffusion systems”. In: *WIAS Preprint* 1701 (2012).
- [91] J. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Études mathématiques. Gauthier-Villars, 1969.
- [92] E. H. C. Lubich and G. Wanner. *Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations. Reprint of the second 2006 ed.* Reprint of the second 2006 ed. Berlin: Springer, 2010, pp. xviii + 644.
- [93] K. Lucas, C. Cork, B. Yu, P. D., G. Luk-Pat, A. Miloslavsky, and A. Painter. *Triple patterning in 10nm node metal lithography*. 26 November 2012. url: <http://spie.org/x91233.xml>.
- [94] K. Lucas, C. Cork, B. Yu, G. Luk-Pat, B. Painter, and D. Z. Pan. “Implications of triple patterning for 14nm node design and patterning”. In: *Proc. of SPIE* 8327 (2012), pages.
- [95] C. Mack. *Fundamental Principles of Optical Lithography : the Science of Microfabrication*. Chichester, West Sussex, England: Wiley, 2007.
- [96] D. Matiut, A. Erdmann, B. Tollkuehn, and A. Semmler. “New models for the simulation of post-exposure bake of chemically amplified resists”. In: *Proc. of SPIE* 5039 (2003), pp. 1132–1142.
- [97] G. Matthies and L. Tobiska. “Mass conservation of finite element methods for coupled flow-transport problems”. In: *Int. J. Comput. Sci. Math.* 1.2-4 (2007), pp. 293–307.
- [98] H. Meyer, W. Klein, and A. Weiss. “Kinetische untersuchung der reversiblen dimerisierung von o-phenylendioxidimethylsilan”. In: *J. Organomet. Chem.* 177.2 (1979), pp. 323–328.
- [99] L. Michaelis and M. Menten. “Die Kinetik der Invertinwirkung”. In: *Biochem. Zeitsch.* 49 (1913), pp. 333–369.
- [100] A. Mielke. “A gradient structure for reaction–diffusion systems and for energy-drift-diffusion systems”. In: *Nonlinearity* 24.4 (2011), p. 1329.
- [101] A. Mielke. “Thermomechanical modeling of energy-reaction-diffusion systems, including bulk-interface interactions”. In: *Discrete Contin. Dyn. Syst* 6(2) (2013), pp. 479–499.
- [102] A. Mielke, J. Haskovec, and P. A. Markowich. “On uniform decay of the entropy for reaction-diffusion systems”. In: *WIAS Preprint* 1768 (2013).
- [103] J. Morgan. “Global existence for semilinear parabolic systems”. In: *SIAM J. Math. Anal.* 20.5 (1989), pp. 1128–1144.

- [104] J. Moser. “A new proof of de Giorgi’s theorem concerning the regularity problem for elliptic differential equations”. In: *Comm. Pure Appl. Math.* 13.3 (1960), pp. 457–468.
- [105] R. F. Muirhead. “Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters.” In: *Edinb. M. S. Proc.* 21 (1903), pp. 143–157.
- [106] D. L. Nelson and M. M. Cox. *Lehninger Principles of Biochemistry, Fourth Edition*. Ed. by Freeman. Fourth Edition. 2004.
- [107] R. A. Nicolaides. “Direct discretization of planar div-curl problems”. In: *SIAM J. Numer. Anal.* 29.1 (1992), pp. 32–56.
- [108] L. Nirenberg. “An extended interpolation inequality.” In: *Ann. Sc. Norm. Super. Pisa, Sci. Fis. Mat., III. Ser.* 20 (1966), pp. 733–737.
- [109] OpenMP Architecture Review Board. *OpenMP Application Program Interface Version 3.0*. May 2008.
- [110] P. P. Érdi and J. Tóth. *Mathematical models of chemical reactions. Theory and applications of deterministic and stochastic models*. Manchester: Manchester University Press, 1989, pp. xxiv + 259.
- [111] M. Pierre. “Global existence in reaction-diffusion systems with control of mass: a survey”. In: *Milan J. Math.* 78.2 (2010), pp. 417–455.
- [112] A. Pomp, O. Schenk, and W. Fichtner. “An ILU preconditioner adapted to diffusion processes in semiconductors”. In: 1999, pp. 133–141.
- [113] G. Rolland. “Global existence and fast-reaction limit in reaction-diffusion systems with cross effects”. PhD thesis. École normale supérieure de Cachan, 2012.
- [114] G. van Rossum et al. *Python Language Website*. Mar. 2010. url: <http://www.python.org/>.
- [115] T. Roubíček. *Nonlinear partial differential equations with applications*. Vol. 153. International Series of Numerical Mathematics. Basel: Birkhäuser Verlag, 2005, pp. xviii+405.
- [116] M. Schäfer. *Computational Engineering - Introduction to Numerical Methods*. Springer, 2006.
- [117] O. Schenk and K. Gärtner. “Solving unsymmetric sparse systems of linear equations with PARDISO”. In: *Computational science—ICCS 2002, Part II (Amsterdam)*. Vol. 2330. Lecture Notes in Comput. Sci. Berlin: Springer, 2002, pp. 355–363.
- [118] T. Schnattinger. “Mesoscopic Simulation of Photoresist Processing in Optical Lithography”. PhD thesis. Erlangen: Universität Erlangen-Nürnberg, 2007.
- [119] L. A. Segel and M. Slemrod. “The quasi-steady state assumption: a case study in perturbation”. In: *SIAM Rev.* 31.3 (Sept. 1989), pp. 446–477.

- [120] J. R. Shewchuk. “Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator”. In: *Applied Computational Geometry: Towards Geometric Engineering*. Ed. by M. C. Lin and D. Manocha. Vol. 1148. Lecture Notes in Computer Science. From the First ACM Workshop on Applied Computational Geometry. Springer-Verlag, May 1996, pp. 203–222.
- [121] W. Shockley and W. T. Read. “Statistics of the Recombinations of Holes and Electrons”. In: *Phys. Rev.* 87 (5 1952), pp. 835–842.
- [122] H. Si, K. Gärtner, and J. Fuhrmann. “Boundary conforming Delaunay mesh generation”. In: *Comput. Math. Math. Phys.* 50.1 (2010), pp. 38–53.
- [123] T. Streckenbach, J. Fuhrmann, H. Langmach, and M. Uhle. *Pdelib- a software toolbox for numerical computations*. 2009. url: <http://pdelib.org>.
- [124] A. M. Stuart and A. R. Humphries. *Dynamical Systems and Numerical Analysis*. Cambridge University Press, 1996.
- [125] T. Suzuki. *Free energy and self-interacting particles*. English. Boston, MA: Birkhäuser, 2005.
- [126] R. Wegscheider. “Über simultane Gleichgewichte und die Beziehungen zwischen Thermodynamik und Reaktionskinetik homogener Systeme”. In: *Monatshefte für Chemie und verwandte Teile anderer Wissenschaften* 32.8 (1911), pp. 849–906.
- [127] Y. Wei and R. Brainard. *Advanced Processes for 193-nm Immersion Lithography*. Press Monographs. SPIE, 2009.
- [128] E. Zeidler and L. Boron. *Nonlinear functional analysis and its applications: II/B Nonlinear monotone operators*. Nonlinear Functional Analysis and Its Applications Bd. 2. Springer New York, 1990.

Kurzzusammenfassung

Gegenstand der Arbeit sollen abgeschlossene Systeme von parabolischen partiellen Differentialgleichungen sein, welche die Umwandlung von chemischen, elektrisch neutralen Spezies in heterogenen Materialien beschreiben, welche durch Diffusion und reversiblen Reaktionen getrieben werden. Insbesondere die Modellierung von Prozessen der chemischen Verfahrenstechnik, wie sie zum Beispiel bei der Herstellung von Halbleitern in der Fotolithografie auftreten, führen auf Reaktions-Diffusionssysteme, die durch nicht glatte Daten im Orte gekennzeichnet sind.

Der Beitrag dieser Arbeit betrifft die örtliche Voronoi Finite-Volumen und die zeitliche implizite Euler Diskretisierung solcher dissipativen Reaktions-Diffusionssysteme. Für den Beweis von Aussagen über die Stabilität von Lösungen und von a priori-Abschätzungen für die diskreten Lösungen werden Techniken in die diskrete Welt übertragen, die sich bereits beim Studium des kontinuierlichen Problems als zielführend herausgestellt haben. Den Ausgangspunkt bilden dabei Abschätzungen der diskreten Freienenergie, die entlang von Trajektorien des diskreten Problems monoton und exponentiell gegen ihren Gleichgewichtswert fällt.

Unter zusätzlichen Annahmen an die Reaktionsordnung (Reaktionen mit maximal quadratischen Quelltermen) konnte die uniforme zeitlich globale Existenz von positiven Lösungen gezeigt werden. Ausdrücklich sei darauf hingewiesen, dass in beiden Fällen die auftretenden Konstanten unabhängig von der Güte des Voronoi-Gitters und vom Zeitschritt sind, womit auch anisotrope Gitter nicht ausgeschlossen sind. Die lokale Existenz von Lösungen des diskreten Problems konnte unter den natürlichen Voraussetzungen der Quasipositivität der Reaktionsterme und der Erhaltung der Atomzahl bewiesen werden. Zusammenfassend konnten alle qualitativen Eigenschaften des kontinuierlichen Systems (thermodynamisches Gleichgewicht, monotones und exponentielles Fallen der freien Energie, globale obere und untere Schranken) auch für das diskretisierte Problem nachgewiesen werden. Ein weiteres neues Resultat betrifft die Konvergenz des Schemas.

Für das Beispiel der Michaelis-Menten-Henry-Kinetik wird eine prototypische Implementierung des Schemas aufgezeigt, bei der großen Wert auf die Erhaltung der theoretisch erzielten Eigenschaften über große Zeitintervalle gelegt wird. Um die Stabilität und Anwendbarkeit der Methode auf reale Problem zu demonstrieren werden verschiedene Beispiele mit Michaelis-Menten-Henry-Kinetik betrachtet, welche sich speziellen Eigenschaften der implementierten Methode widmen. In allen Beispielen sind während der Rechnungen die physikalischen Eigenschaften des Schemas bis auf Rundungsfehler auch über lange Zeitintervalle erhalten. Darüber hinaus löst das analysierte Verfahren die verschiedenen Zeitskalen des Systems auf, wenn auch wandernde Reaktionsfronten oder starke Gradienten, in der Nachbarschaft von Materialübergängen nur grob durch das Gitter approximiert sind.

Curriculum Vitae

The curriculum vitae is not published in the digital version of the thesis due to privacy restrictions.

Eidesstattliche Erklärung

Gemäß §7 (4), der Promotionsordnung des Fachbereichs Mathematik und Informatik der Freien Universität Berlin versichere ich hiermit, dass ich alle Hilfsmittel und Hilfen angegeben und auf dieser Grundlage die Arbeit selbständig verfasst habe. Des Weiteren versichere ich, dass ich diese Arbeit nicht schon einmal zu einem früheren Promotionsverfahren eingereicht habe.

Berlin, 18. Dezember 2013

André Fiebach