

FAKULTÄT FÜR WIRTSCHAFTSINFORMATIK
UND WIRTSCHAFTSMATHEMATIK
UNIVERSITÄT MANNHEIM

SIMULATION AND OPTIMIZATION OF
COUPLED GAS AND POWER NETWORKS

INAUGURALDISSERTATION ZUR ERLANGUNG DES AKADEMISCHEN
GRADES EINES DOKTORS DER NATURWISSENSCHAFTEN DER
UNIVERSITÄT MANNHEIM

VORGELEGT VON

EIKE FOKKEN

GEBOREN IN FRANKFURT AM MAIN

MANNHEIM, MÄRZ 2022

Dekan:	Dr. Bernd Lübcke
Erstgutachterin:	Prof. Dr. Simone Göttlich
Zweitgutachter:	Prof. Dr. Michael Herty
Tag des Fachvortrages:	7. Juni 2022

SIMULATION AND OPTIMIZATION OF COUPLED GAS AND POWER NETWORKS

THIS DOCTORAL THESIS WAS WRITTEN BY

EIKE FOKKEN

AT THE

CHAIR OF SCIENTIFIC COMPUTING

AT THE UNIVERSITY OF MANNHEIM

UNDER SUPERVISION OF

PROF. DR. SIMONE GÖTTLICH

ABSTRACT

In this thesis, the coupling of gas and power networks is examined. The isentropic Euler equations are employed as a model for gas flow in a pipeline network, whereas the power flow equations are used as a model for an electrical grid.

We find a set of conditions on pressure functions that guarantee well-posedness of the Riemann problem of the isentropic Euler equations in the sub-sonic regime. Coupling conditions for modeling gas-fired power plants and power-to-gas plants, both of which couple gas networks to power networks, are proposed and their well-posedness is shown.

Solutions of the power flow equations are derived as certain time-periodic solutions to the physically more accurate Telegrapher's equations and the exponential stability of these solutions is examined. In addition, a numerical method for solving the Telegrapher's equations is introduced, which mimics the exponential stability of the analytical solution.

For the power network, both deterministic and stochastic power demand is considered and the influence of the demand on the pressure (and flux) in the gas network is examined. Our findings are applied to gas and power networks of realistic size and corresponding simulation and optimization problems are solved. This also showcases the software *grazer*, that was written for this thesis and is capable of efficiently solving time-dependent problems that are defined on networks. It is released under the open source license "MIT License" to enable future use, for example in further research.

ZUSAMMENFASSUNG

Diese Arbeit befasst sich mit der Kopplung von Gas- und Stromnetzen. Die isentropen Eulergleichungen dienen hierbei als Modell für den Gasfluss in einem Pipeline-Netz, während die Lastflussgleichungen (auch *Power-Flow*-Gleichungen) den Stromfluss in einem Stromnetz modellieren.

Für das Gasmodell werden Bedingungen gefunden, die von Druckfunktionen der isentropen Eulergleichungen erfüllt werden müssen, um die Wohlgestelltheit des Riemann-Problems ebendieser Gleichungen für Unterschall-Anfangsbedingungen zu garantieren. Außerdem werden Kopplungsbedingungen für die Modellierung von Gaskraftwerken und Power-to-Gas-Anlagen, die für die Kopplung von Gas- und Stromnetzen verantwortlich sind, aufgestellt und ihre Wohlgestelltheit gezeigt.

Lösungen der Lastflussgleichungen werden als spezielle, zeitperiodische Lösungen der Telegraphengleichungen hergeleitet. Die exponentielle Stabilität dieser Lösungen wird untersucht und ein numerisches Verfahren vorgestellt, das ebenfalls exponentielle Stabilität aufweist und so das Verhalten der analytischen Lösung imitiert.

Im Stromnetz werden sowohl deterministische als auch stochastische Nachfrageprofile berücksichtigt und ihr Einfluss auf Druck (und Fluss) im Gasnetz untersucht. All diese Resultate werden auf Gas- und Stromnetze realistischer Größe angewandt und zugehörige Simulations- und Optimierungsprobleme werden gelöst. Dies führt gleichzeitig die Fähigkeiten der Software *grazer* vor, die für diese Arbeit geschrieben wurde, und in der Lage ist, zeitabhängige Probleme, die auf Graphen definiert sind, effizient zu lösen. *Grazer* ist unter der Open-Source-Lizenz "MIT Lizenz" veröffentlicht, um die Weiterverwendung, etwa in zukünftiger Forschung, zu ermöglichen.

ACKNOWLEDGMENTS

Here I would like to express my gratitude to several people that in one way or another contributed to this thesis.

I would like to thank Simone Göttlich for guiding me in my research, always making time for discussions on it and in addition being both a firm and considerate supervisor. Working with her was a pleasure. Thanks also go to Michael Herty, who helped my understanding of hyperbolic balance laws and especially the isentropic Euler equations tremendously. I am very grateful to Oliver Kolb for sharing his expertise on implementing *ANACONDA*, which in turn led to the creation of *grazer*. All three also co-authored papers with me and the collaboration was a very pleasant experience.

It was always a pleasure to meet my colleagues at the Chair of Scientific Computing: Thomas Schillinger, Jan Friedrich and Jennifer Weißen and the former colleagues Thomas Jung, Marion Triplet, Kathinka Hameister and Stephan Knapp. I thank all of them for many fruitful discussions and also the times spent together outside of work. Thomas Schillinger also proof-read parts of the manuscript, which I am very grateful for.

Special thanks go to Jürgen Jaap, who was very helpful in all technical troubles I had over the years and Sabine Braak, who knows the ins and outs of all administrative details at the university and managed a lot of what is going on in the background of the Chair.

I also thank my parents Christine and Fokke Fokken, who always encouraged me, especially to pursue my doctoral studies, and also took over baby sitting in a pinch.

The biggest thanks go to my wife Anna Schilling, who encouraged me when I needed it, tacitly overlooked my occasional mental absence and did the lion's share of proofreading. Especially in the last few weeks she also watched our children for more than her fair share.

Lastly I thank the BMBF project "ENets" (05M18VMA) and the DFG project "Multiscale Control Concepts for Transport-Dominated Problems" (SPP1962) for their financial support as well as the University of Mannheim for a scholarship.

CONTENTS

1	Introduction and outline	1
2	Mathematical background	7
2.1	Balance laws	7
2.1.1	Hyperbolic balance laws	9
2.1.2	Riemann problems	11
2.1.3	Entropy solutions	16
2.1.4	Front-tracking	17
2.1.5	Boundary conditions	18
2.1.6	Junctions and networks	21
2.2	Numerical methods for hyperbolic balance laws	24
2.2.1	Schemes for linear hyperbolic balance laws	26
2.2.2	Treatment of source terms and splitting	29
2.2.3	IBOX scheme	30
2.2.4	CWENO3 scheme	31
2.3	Optimization	33
2.4	Semigroup theory	35
3	The relation of Telegrapher's equations and the Power flow model	39
3.1	Solution structure of the Telegrapher's equations	40
3.2	From Telegrapher's equation to power flow equations	47
3.2.1	Periodic solutions to the Telegrapher's equations	47
3.2.2	Power coupling conditions	52
3.3	Numerical scheme	53
3.3.1	The scheme in the interior of a line	54
3.3.2	Numerical coupling conditions	59
3.4	Numerical examples	61
3.5	Summary and outlook	65
4	The isentropic Euler equations and gas-power-coupling	67
4.1	Gas networks	67
4.1.1	Well-posedness of the Riemann problem for the isentropic Euler equation	70
4.1.2	A note on the vacuum	80
4.1.3	Extension to junctions	81
4.1.4	Additional constraints for consistency	82
4.2	Coupling to an external gas sink	84
4.3	Numerical results	86
4.3.1	Validation	87
4.3.2	Different pressure laws	88
4.4	Summary	89
5	Application to coupled gas-power networks	91
5.1	The combined model	92
5.2	A simple network	93
5.2.1	Model	93
5.2.2	Discretizations	97

5.2.3	Network properties and numerical results	99
5.3	A realistic network	101
5.3.1	A refined model for realistic networks	101
5.3.2	Discretization	106
5.3.3	Network properties	106
5.3.4	Numerical results	110
5.4	A network with volatile power demand	116
5.4.1	Stochastic power demand	116
5.4.2	Discretization	117
5.4.3	Network properties	118
5.4.4	Numerical results	121
5.5	Summary and outlook	124
6	Optimization	127
6.1	Optimization with control and state variables	128
6.2	Derivatives of time-dependent systems	130
6.2.1	The discretized direct method	133
6.2.2	The discretized adjoint method	134
6.3	Additional models for optimal control problems	137
6.3.1	Compressors and control valves	137
6.3.2	Pressure-constrained gas nodes	137
6.4	Discretization	138
6.5	A small optimal control problem	138
6.5.1	Network properties	138
6.5.2	Numerical results	139
6.6	A large optimal control problem	139
6.6.1	Network properties	140
6.6.2	Numerical results	140
6.7	Summary and outlook	141
7	Summary and Conclusion	143
	Bibliography	145
A	Appendix: On the implementation of grazer	157
A.1	Overview	157
A.2	Installation	157
A.3	Usage	158
A.4	A rough overview of the inner workings	159

INTRODUCTION AND OUTLINE

Overview

In recent years, due to the need to restructure energy systems to incorporate more renewable energy sources, renewed interest into energy systems has sparked. Power generation is changing drastically with more and more available energy coming from renewable sources. The current transformation of the energy system is driven by at least three trends: decarbonization and defossilization of energy supply, growing concerns about the climate crisis, and transformative political decisions, e.g. the phase-out of nuclear energy [AtG] and scheduled phase-out of coal and lignite in Germany. Phase out of all fossil energies is, in principle, doable as averaged over long time spans renewable sources (wind, solar, etc.) provide sufficient amounts of energy to achieve decarbonization. Yet, renewable generation and demand are not synchronized in time and space and thus energy storage as well as energy distribution and transport are of crucial importance. For example, the International Energy Agency (IEA) predicts a global need for substantial growth of energy storage capacity [IEA19; Zab19]. Currently, neither a readily and widely usable storage technology to buffer the quantities of electrical energy required for decarbonization exists, nor is there scientific consensus on which large-scale storage technologies will be available in near-to-mid future. Hence, it comes at no surprise that the coupling of energy domains and sectors is gaining increasing research attention. For example, the economic viability of future power-to-X pathways (where X can be Hydrogen, Methane, or synthetic bio-fuels, for instance) has been investigated in several studies, see for example [Bro+18; HBO16; Sch+15]. These investigations are driven by the fact that natural gas can be stored in sufficient quantities in dedicated installations and to a certain extent directly in the gas grid itself. In other words, coupling of electrical grids and gas networks is currently considered a promising road towards a high share of renewables. Due to their flexibility in comparison to other power plants, gas turbines and gas engine power plants are often proposed as a means to balance power demands that cannot be met with renewable power sources at a given time. The repercussions on the gas pipeline networks by this balancing have been studied in [CBL15] and also joint optimal control of gas and power networks are investigated, for instance in [Zen+16; Zlo+16]. However, only steady-state or restricted gas dynamics have been investigated on networks so far. For example, the model introduced in [Zlo+16] ignores complex dynamics such as shocks or rarefaction waves. Although ruling out full dynamics is often sufficient, the — in comparison to power networks — slow signal speed allows for short term effects to linger long enough to be resolved, which is impossible to achieve with steady state methods.

While [HMS19] addresses the nonlinear dynamics, i.e. the full compressible Euler equations, the analysis in there only considers a single gas power plant without any network interdependence.

The aim of this work is therefore to investigate the interaction of gas and power networks on realistic scales and showcase the capabilities of the software suite *grazer*¹, which the author implemented for this purpose. It is capable of solving time-dependent simulation and optimization problems that are defined on graphs, like the energy networks considered in this work and was used for most of the numerical computations of coupled gas and power networks in this thesis.

For electrical grids several models can be found in the literature, see for example [And15; Bie15; FR16; GHM21] for an overview. All modeling approaches rely on a graph structure where generators of electrical powers and consumers at nodes are connected by transmission lines as seems natural for an entity that also in reality consists of junctions connected by (transmission) lines. Physically, voltage and current are transported along the lines while power loss may occur due to resistances. In many applications, the so-called power flow equations, see [Bie15; HDS13], provide a well-established model to analyze the behavior of electrical grids over time. Mathematically, the resulting nonlinear system of equations is typically solved via Newton’s method. Another approach, suitable for studying not only the temporal resolution of power flow but also the spatial resolution in transmission lines, are the spatially one-dimensional Telegrapher’s equations. These equations are a coupled system of linear hyperbolic balance laws for voltage and current. Unfortunately the signal velocity of the Telegrapher’s equations is so big, that capturing any meaningful time-dependent behavior is numerically challenging, or — for medium and large sized networks — outright impossible to do with reasonable resources. Therefore the use of the power flow equations is desirable to ignore these extremely short-lived effects for reducing numerical load, but should be justified by comparing to the more accurate solution of the Telegrapher’s equations. Even though the power flow equations represent a great simplification over the Telegrapher’s equations, solving them for large networks also involves challenges. Research into these includes distributed optimization [Eng+19; KBoo; Mol+17], convex relaxations [Low14a; Low14b], and the consideration of stochastic disturbances [BCH14; Müh+19].

Gas networks have been studied for a long time and have sparked renewed interest in the last two decades, when coupling conditions have been analyzed from a theoretical point of view, see for example [BGH11; BHK06a; BHK06b; Bre+14; CG08; HHW20; Rei14; Rei15]. In principle similar to electrical grids, gas pipeline networks can be described by a coupled system of balance laws, as done in the references just given. However, due to the compressibility of gas, the arising models differ from those of power networks as one has to consider nonlinear effects in the gas flow through pipelines². Since

¹ <https://github.com/eike-fokken/grazer>

² Note that already nowadays so-called linepack flexibility is used for different purposes by operators of gas grids, see for example [NTS19].

these dynamics allow for discontinuous solutions, a careful theoretical and numerical treatment is needed to master the challenges of simulation and optimization for complex networks. Finite volume or finite difference methods are typically applied to solve the network problem numerically [GZ19; KLB10; Pfe+15], an approach we follow as well.

Historically, however, the critical energy system infrastructure for gas and power grids has been separated in terms of operation and control. Hence, there do not exist established standards for joint operation and control of multi-energy grids. In turn multi-energy systems arising from sector-coupling pose control and optimization challenges, many of which are still open or are subject to ongoing research efforts, see [CBL15; Zen+16; Zlo+17], or more recently [FGK19; FGK20; HMS19; OMa+20].

In view of the changing energy demands and supplies the combined and intertwined energy networks are expected to play a prominent role in the future. Here, the unpredictable and volatile energy sources need to be complemented possibly additional large-scale storage [IEA19; Zab19]. Following previous approaches [CBL15; Zen+16; Zlo+17] we consider energy storage through coupling power networks to gas networks. A major concern when coupling gas and power networks is guaranteeing a stable operation even at times of stress due to (uncertain) heavy loads. The propagation of possible uncertain loads on the power network and its effect on the gas network has been subject to recent investigation and we refer to [CBL15] and the references therein. Contrary to the cited reference [CBL15] in this work we are interested in a full simulation of both the gas and power network as well as a simulation of the stochastic demand, respectively supply. Our approach allows the prediction of gas pressure and gas flux at each point in a pipeline as well as nodes of the network. Both quantities are relevant to assess possible stability issues as well as allow for coupling towards the electricity network. Other research in this area includes [Aïd+09; Baro2; KSBo9; LSo2; SSoo; Wag14] and the monograph [BBKo8] that prescribe the electricity demand as Ornstein–Uhlenbeck processes.

Contributions

The key contributions of this thesis are threefold:

The first is Proposition 4.10. It provides a concise list of conditions on the pressure function — relating gas density and pressure — used in the isentropic Euler equations (see Chapter 4) for corresponding Riemann problems (Section 2.1.2) to be well-posed for arbitrary initial values, whose flow speeds are less than the local speed of sound within the gas, so-called *sub-sonic* initial conditions.

A corollary of this proposition is Proposition 4.2, which extends the usual notion of pressure functions known as γ -laws in Proposition 4.2. γ -laws are a family of pressure functions indexed by $\gamma \in \mathbb{R}$, which usually are restricted to $1 \leq \gamma \leq 3$, as inspired by physics, where these values are related to the number of degrees of freedom in a gas. We show that on the one hand, the Riemann problem is well-posed

in addition for $-1 < \gamma < 1$, and that there are ill-posed Riemann problems with sub-sonic initial conditions for every $\gamma \notin [-1, 3]$. We do not make a claim about $\gamma = -1$.

A second contribution is the proposal of a combined gas-power network model and a comparison of pressure and Bernoulli coupling conditions as well as an approach to modeling uncertain power demand for realistically sized gas and power networks with hundreds of nodes and connections.

The last major contribution is the software suite *grazer* itself with simulation and optimization capability for time-dependent problems defined on graphs, extensive automated software tests and of course its open source license³, which guarantees the right to use *grazer* free of charge, analyze the source code and redistribute modified and unmodified versions. Let us emphasize that our approach and also *grazer* is not limited to the particular application of gas and power networks but could eventually be applied to problems of traffic flow and supply chain dynamics on networks.

A minor contribution is the analysis with regards to entropy stability of a well-known second order scheme for the Telegrapher's equations.

Other software suites

Finally, we point to other existing numerical software with possibly similar objective. For example, in [Kol11] the solver *ANACONDA* for a gas network based on (general) hyperbolic balance laws has been introduced. This implementation, or rather its description in [Kol11] serves as a foundation for our own software suite *grazer*, introduced in this thesis. Our tool has the advantage of easier extensibility, an open source license and a modern design approach featuring for example extensive software testing. In addition we also include stochastic power demands in the power flow network setting. In [ATK21] uncertainty in power flow is computed relying on approaches based on neural networks. A difference to the presented approach is the restriction to linear power flow problems and the absence of coupling to gas networks. A further concept called *planqres*, see [Beu+19] has been presented to also address general renewable energy sources as well as energy distribution based on discrete optimization approaches, which focuses on energy system modeling in more generality. Another software suite in the area of energy system simulation and optimization is *PyPSA* [BHS18], which models whole energy systems but doesn't model pipelines. Furthermore, there exists a software suite by Fraunhofer SCAI called *MYNTS*, see [Cle+16], that also includes simulation and optimization of inter-connected grid operations with a focus on the design of suitable networks.

Outline

The material of this thesis is mostly taken from the author's publications, namely

³ MIT License, see <https://opensource.org/licenses/MIT>

- Eike Fokken, Simone Göttlich: *On the relation of power flow and Telegrapher's equations: continuous and numerical Lyapunov stability*, arxiv preprint, 2021 [FG21],
- Eike Fokken, Simone Göttlich, Oliver Kolb: *Modeling and simulation of gas networks coupled to power grids*, Journal of Engineering Mathematics, 2019, Springer Netherlands [FGK19],
- Eike Fokken, Tillmann Mühlpfordt, Timm Faulwasser, Simone Göttlich: *Modeling and Simulation of Sector-Coupled Energy Networks: A Gas-Power Benchmark*, in: Mathematical Modeling, Simulation and Optimization for Power Engineering and Management, 2021, Springer International Publishing [Fok+21],
- Eike Fokken, Simone Göttlich, Michael Herty: *Efficient simulation of coupled gas and power networks under uncertain demands*, European Journal of Applied Mathematics, 2022, Cambridge University Press, [FGH22],
- Eike Fokken, Simone Göttlich, Oliver Kolb: *Optimal Control of Compressor Stations in a Coupled Gas-to-Power Network*, in: Advances in Energy System Optimization, 2020, Springer International Publishing [FGK20].

In Chapter 2 we will provide the necessary mathematical preliminaries for this thesis.

Afterwards, in Chapter 3, based on [FG21], we analyze both Telegrapher's and power flow equations and the relation between them. We show the extent to which one can trust the less physically accurate but numerically more tractable power flow equations to yield a solution that is close to the one that theoretically could be obtained by solving the Telegrapher's equations directly. In this chapter we also analyze a second order scheme applicable for the Telegrapher's equations on networks, and show its entropy-stability, which is similar to that of the analytical solution. The analytical results in this chapter have already been proved for the most part, for example in [Nic16], and are presented to give an understanding of the underlying concepts.

In Chapter 4, which is based on the publication [FGK19], we examine the isentropic Euler equations and find conditions on its pressure function, that guarantee well-posedness of the corresponding Riemann problems. We also prove the well-posedness and show properties of our proposed gas-power coupling model.

In Chapter 5, which is based on [Fok+21] and [FGH22], we introduce uncertainty of power demand and present simulation results showing the applicability of our models to networks, whose sizes are a dozen of nodes up to hundreds of nodes.

Chapter 6, based in part on [FGK20] and [FGH22], finally presents the analysis of two common approaches to extracting derivative information from systems of equations and the solution of optimal control problems of a gas-power network.

Lastly we summarize the work in Chapter 7.

To facilitate future use of *grazer*, Appendix A gives a brief overview of the structure and use of the software suite.

MATHEMATICAL BACKGROUND

2.1 BALANCE LAWS

Both the dynamics of gas flow and evolution of voltage and current can most appropriately be described by certain partial differential equations, so-called hyperbolic balance laws. To define them, let us first declare some nomenclature. We use $\mathcal{T} \subset \mathbb{R}_0^+$ as the space of time points, $X \subset \mathbb{R}$ as the space of positions and $\mathcal{S} = \mathbb{R}^m$ as the so-called state space. Note also that we usually write u_t for $\frac{\partial u}{\partial t}$ and u_x for $\frac{\partial u}{\partial x}$.

Definition 2.1 (balance law).

- Let $d = \dim(X) = 1$, $f \in C^3(\mathcal{S}, \mathcal{S})$, $S \in C^0(\mathcal{S}, \mathcal{S})$ and $u : X \times \mathcal{T} \rightarrow \mathcal{S}$. Then a *balance law in one space dimension* is an equation of the form

$$\begin{aligned} \partial_t u(x, t) + \partial_x f(u(x, t)) &= S(u(x, t)), \\ u(x, 0) &= u_0(x), \end{aligned} \quad (2.1)$$

for all $x \in X, t \in \mathcal{T}$. Naturally, u has to satisfy some properties to make sense of the balance law, there must be at least a notion of derivatives of u .

- Any $u \in C^1(X \times \mathcal{T}, \mathcal{S})$ satisfying the Balance law (2.1) is called a *classical solution* of the balance law.
- An initial value problem of the form (2.1) is also called a *Cauchy problem*.

In the case $\dim(\mathcal{S}) = 1$, Equation (2.1) is often called a *scalar balance law*. Other naming conventions include calling what we just termed “scalar balance law” simply a *balance law* and what we termed a “balance law” in Definition 2.1 instead a *system of balance laws*. We will use balance law for all of the above and sometimes use “scalar balance law” or “system of balance laws”, when the distinction is necessary.

The name *balance law* is better understood in the integral form. Integrating (2.1) over an interval $I = [x_0, x_1] \subset \mathbb{R}$ leads to

$$\int_{x_0}^{x_1} \partial_t u \, dx + \int_{x_0}^{x_1} \partial_x (f(u)) \, dx = \int_{x_0}^{x_1} S(u) \, dx.$$

Using our assumption that u and f are continuously differentiable, we can apply the fundamental theorem of calculus to get

$$\int_{x_0}^{x_1} \partial_t u \, dx + f(u(x_1)) - f(u(x_0)) = \int_{x_0}^{x_1} S(u) \, dx,$$

Again, because u is continuously differentiable, we can interchange integration and differentiation to get

$$\partial_t \int_{x_0}^{x_1} u \, dx + f(u(x_1)) - f(u(x_0)) = \int_{x_0}^{x_1} S(u) \, dx. \quad (2.2)$$

By setting $U = \int_{x_0}^{x_1} u \, dx$ we get

$$\partial_t U + f(u(x_1)) - f(u(x_0)) = \int_{x_0}^{x_1} S(u) \, dx. \quad (2.3)$$

Here U is some quantity (e.g. mass), u its spatial density, f its flux and S some *balancing function*. This formulation can be interpreted as: *The change of U inside the interval I is equal to its flux over the boundary of I plus some balance encoded in (the integral of) S .*

Hence the name balance law. The balancing function models some way of creating or destroying units of U and is therefore called *source* or *sink*, although in mathematical literature only *source* is used and sinks are just negative sources. The concept of a balance law also makes sense in more than one space dimension ($\dim(X) > 1$), where a volume takes the place of the interval I and instead of the fundamental theorem of calculus, the Gauß divergence theorem must be applied. Yet, in this work, no more than one space dimension is needed.

A balance law with vanishing source is called a *conservation law*, because in that case, the quantity U can only decrease (increase) by leaving (entering) the interval I and no unit of U can be created or destroyed, therefore U is conserved.

In deriving Equation (2.3), we used a strong assumption on u , namely that its derivative exists and is continuous. A great drawback of this assumption is that for most interesting balance laws, no classical solutions exist. Even if they exist, they can often not be defined for more than a small time interval. We can drop the assumption of continuous derivatives to find more general classes of solutions and will do so from now on. Although much of the following theory can also be derived for balance laws in more than one space dimension, we will from now on restrict ourselves to $d = 1$. First we state our notion of a test function.

Definition 2.2 (test function). A function $\phi \in C_0^\infty(\mathbb{R} \times \mathcal{T}, \mathcal{S})$, that is a smooth function with compact support, is called a *test function*.

With the help of test functions we can define a sufficiently general solution type:

Definition 2.3 (weak solution, [GR14, Definition 2.1]). A function $u \in L_{\text{loc}}^\infty(\mathbb{R} \times \mathbb{R}_0^+, \mathcal{S})$ that satisfies

$$\int_{\mathbb{R}_0^+} \int_{\mathbb{R}} (u \cdot \partial_t \phi + f(u) \cdot \partial_x \phi + S(u) \cdot \phi) \, dx \, dt + \int_{\mathbb{R}} u_0(x) \cdot \phi \, dx = 0$$

for all test functions ϕ is called a *weak solution* of the Balance law (2.1).

It is easy to verify that classical solutions are also weak solutions, using integration by parts and the fact that boundary terms vanish because of the compact support of ϕ . Note that weak solutions need not be differentiable, continuous or unique.

2.1.1 Hyperbolic balance laws

In order to find solutions we need a key assumption on our balance laws, namely hyperbolicity.

Definition 2.4 (hyperbolic balance law).

- A balance law of the form (2.1) is called *hyperbolic*, if the Jacobian f' (sometimes written as \mathcal{J}_f) of f is diagonalizable everywhere.
- If in addition f' has pairwise distinct eigenvalues everywhere, the balance law is called *strictly hyperbolic*.

The solution theory of hyperbolic balance laws is often developed without a source term, that is, for conservation laws, see [CG10] for example. Results on actual balance laws can then be obtained from those of conservation laws by a splitting technique, treating the source term separately from the flux, which we will also employ in Chapter 3. The drawback of the splitting is that it needs strong assumptions on the source term to be applicable, see [HR02, Section 4.5] for details. The existence of solutions of (systems of) conservation laws can be established with the front tracking technique, briefly described in Section 2.1.4, although only under strong assumptions on the initial conditions. It amounts to constructing certain approximate solutions to the conservation law in question, where the approximation error is governed by some parameter δ and making sure that the approximate solutions converge to a weak solution of the conservation law for $\delta \rightarrow 0$. More details than in Section 2.1.4 can be found in [Daf09, Section 16] and [HR02]. A derivation for coupling conditions in networks is found in [Gug+12, Section 3].

To study the structure of solutions to hyperbolic conservation laws, let us first consider a scalar conservation law. For differentiable u we can transform the conservation law

$$\partial_t u + \partial_x f(u) = 0 \quad (2.4)$$

into

$$\partial_t u + f'(u) \partial_x u = 0,$$

which is called the *quasi-linear form* of the conservation law.

Assume that there is a solution u . We try to decouple this solution at different points x . We do so by defining so-called characteristic curves.

Definition 2.5. A straight line $x(t) = x_0 + f'(u_0(x_0))t$ is called a *characteristic curve* through the point $x = x_0, t = 0$ of the scalar conservation law (2.4).

Using the definition of characteristic curves, it is easily computed that $\frac{du(x(t), t)}{dt} = 0$ for a solution u of (2.4) for all $0 < t < T$ for some $T \in \mathbb{R}^+$. This means that solutions of (2.4) are constant along characteristic curves:

$$u(x(t), t) = u_0(x_0) \text{ for } 0 < t < T. \quad (2.5)$$

Vice versa, as long as no two characteristic curves meet up to time T , we can define a solution for every $(x, t) \in X \times [0, T)$ by (2.5). T can be chosen as the minimal time for which there is a point in which some characteristic curve meet, or $T = \infty$, if no two characteristic curves ever meet. In a point where characteristic curves meet, the solution is necessarily non-differentiable. This solution method is called the *method of lines*.

A simple example, already exhibiting much of the general behavior of hyperbolic balance laws is given by the advection equation:

Example 2.6. Let $\mathcal{S} = \mathbb{R}$, $a \in \mathbb{R}$ and $u_0 : \mathbb{R} \rightarrow \mathbb{R}$. The hyperbolic balance law

$$\begin{aligned} u_t + au_x &= 0, \\ u(x, 0) &= u_0(x) \end{aligned}$$

is known as the *advection equation*. It is solved by

$$u(x, t) = u_0(x - at),$$

which can be found by the method of lines. As all characteristic

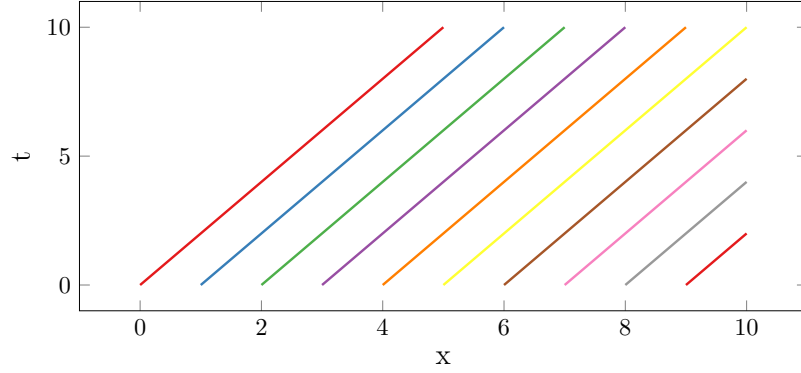


Figure 2.1: Characteristic curves of the advection equation for $a = 2$.

curves are parallel, see Figure 2.1, we can choose $T = \infty$ for the advection equation. Here we can see that a point in the solution $u(x, t)$ of the advection equation can only be influenced by points in the initial condition that are connected to x by a line of “velocity” — that is slope — a . Note also that the information travels to the right, whenever $a > 0$ (as in Figure 2.1), and to the left when $a < 0$.

Let us now come to systems of conservation laws, that is $\dim \mathcal{S} > 1$. In this case, information in solutions travels at most with a velocity given by an eigenvalue of f' . For the following discussion we restrict ourselves to strictly hyperbolic systems, which are the only ones relevant for the gas and power networks considered here. Remember also that we demanded $f \in C^3(\mathcal{S}, \mathcal{S})$ in Definition 2.1, which is important for the smoothness claims in the following. It is useful to order the eigenvalues of $f'(u)$ from smallest to greatest: $\lambda_1(u) < \dots < \lambda_m(u)$. Their corresponding (right) eigenvectors will be denoted by $r_1(u), \dots, r_m(u)$. Note that the i -th eigenvector actually defines a vector field on the state space \mathcal{S} in the differential geometry sense (and the i -th eigenvalue defines a scalar field). For linear systems of conservation laws, the scalar theory can be carried over, by a change of basis:

Example 2.7. Let $M \in \mathbb{R}^{m \times m}$ be diagonalizable. Let $\mathcal{S} = \mathbb{R}^m$ and consider the system of conservation laws

$$u_t + Mu_x = 0. \quad (2.6)$$

We can switch to another basis of \mathcal{S} , given by the eigenvectors $R = (r_1, \dots, r_m)$ of M , thereby diagonalizing M . Then we arrive at a new system

$$w_t + \text{diag}(\lambda_1, \dots, \lambda_m)w_x = 0,$$

which is decoupled and therefore can be solved by solving the m different advection equations, each with their own method of lines from above. This means that the system (2.6) decomposes into different non-interacting waves along the eigenvectors r_i of the system. Each wave travels with a velocity given by its corresponding eigenvalue λ_i .

2.1.2 Riemann problems

For non-linear systems, different waves may interact, and therefore the method of lines can usually not be used. The eigenvalues and eigenvectors still play a major role and will help us solve so-called Riemann problems, which in turn are used in front-tracking for showing well-posedness of hyperbolic conservation laws for certain well-behaved initial conditions.

Definition 2.8 (Riemann problem). A Cauchy problem (2.1) is called a *Riemann problem*, if the initial condition has at most one discontinuity and is piecewise constant. Concretely, for some $u_l, u_r \in \mathcal{S}$, a Riemann problem is given by

$$\begin{aligned} \partial_t u(x, t) + \partial_x f(u(x, t)) &= 0, \\ u(x, 0) &= \begin{cases} u_l & \text{for } x < 0 \\ u_r & \text{for } x \geq 0. \end{cases} \end{aligned} \quad (2.7)$$

For Riemann problems of strictly hyperbolic conservation laws with C^3 flux function, there are two important solution types, namely shocks and rarefaction waves, which we will characterize in the following. But first we introduce integral curves.

Definition 2.9 (integral curve). Let $I \subset \mathbb{R}$ be an open interval and $0 \in I$. An *integral curve* of a differentiable vector field $r : \mathcal{S} \rightarrow \mathcal{S}$ is a differentiable path $q : I \rightarrow \mathcal{S}$, that fulfills

$$q'(\xi) = \alpha(\xi)r(q(\xi)), \quad (2.8)$$

for some differentiable $\alpha : I \rightarrow \mathbb{R}$ and for all $\xi \in I$.

Of course one can re-parameterize any such q to get $\alpha = 1$. These curves play an important role in the following definition distinguishing two important classes of eigen fields (λ_i, r_i) .

Definition 2.10 (genuine non-linearity and linear degeneracy). In the following, let q be an integral curve of the i -th eigenvector field r_i .

- The i -th field (λ_i, r_i) is called *genuinely non-linear* if

$$(\nabla_s \lambda_i)(q(\xi)) \cdot r_i(q(\xi)) \neq 0 \quad (2.9)$$

holds for all $\xi \in I$.

- The i -th field (λ_i, r_i) is called *linearly degenerate* if

$$(\nabla_s \lambda_i)(q(\xi)) \cdot r_i(q(\xi)) = 0 \quad (2.10)$$

holds for all $\xi \in I$.

Equation (2.9) and Equation (2.10) define the evolution of eigenvalues along integral curves. Equation (2.10) means that an eigenvalue remains constant along the curve, while Equation (2.9) means that an eigenvalue is either strictly increasing or strictly decreasing along the curve. For example the linear system (2.6) is linearly degenerate, because the eigenvalues of a constant matrix are constant. Genuine non-linearity and linear degeneracy are useful properties when searching solutions of a Riemann problem.

There are two important solution types of Riemann problems, namely shocks and rarefaction waves, which we will now characterize in Definitions 2.13 and 2.15.

Proposition 2.11. [LeVo2, Section 13.8.5] *Consider a Riemann problem (2.7), whose initial condition fulfills*

$$\lambda_i(u_l) < \lambda_i(u_r) \quad (2.11)$$

for the i -th eigenvalue field of the underlying conservation law (but possibly not for λ_j , $j \neq i$). Let further (λ_i, r_i) be genuinely non-linear and assume there is an integral curve q of r_i , such that

$$\begin{aligned} q(\lambda_i(u_l)) &= u_l, \\ q(\lambda_i(u_r)) &= u_r. \end{aligned}$$

Then this problem has a (weak) solution given by

$$u(x, t) = \begin{cases} u_l & \text{for } \frac{x}{t} < \lambda_i(u_l), \\ q\left(\frac{x}{t}\right) & \text{for } \lambda_i(u_l) \leq \frac{x}{t} \leq \lambda_i(u_r), \\ u_r & \text{for } \lambda_i(u_r) < \frac{x}{t}. \end{cases} \quad (2.12)$$

Finding an integral curve q through one point amounts to solving the ordinary differential equation Equation (2.8). The right-hand side of it is given by the i -th eigenvector as a function of the i -th eigenvalue. Dependent on this function being continuous or locally Lipschitz continuous, existence and local uniqueness may be provable. For the only conservation law in this work that has genuinely non-linear eigenvector fields, namely the isentropic Euler equations detailed in Chapter 4, these fields are locally Lipschitz continuous in their respective eigenvalues in the considered domains.

For sufficiently smooth, strictly hyperbolic fluxes, there is always — at least locally — an integral curve through any given point:

Proposition 2.12 ([Daf09, Theorem 7.6.5]). *Consider a strictly hyperbolic conservation law with flux function $f \in C^3(\mathcal{S}, \mathcal{S})$ and a point $\hat{u} \in \mathcal{S}$. For a genuinely non-linear field (λ, r) of this conservation law there is a C^3 -path q in \mathcal{S} , defined in a neighborhood I of $\lambda(\hat{u})$, such that*

$$\begin{aligned} q'(\xi) &= r, \\ q(\lambda(\hat{u})) &= \hat{u}, \end{aligned} \tag{2.13}$$

As can be seen from Proposition 2.11, for each point $q(\xi)$, $\xi \in I$, the Riemann problem with initial data $u_l = \hat{u}$, $u_r = q(\xi)$ has a solution, whose structure is given in that proposition.

Definition 2.13 (rarefaction wave).

- The solution of a Riemann problem given in Proposition 2.11 is called an *i-th rarefaction wave*.
- The integral curve q from Proposition 2.11 through u_l is called an *i-th rarefaction curve*.

A crucial condition in Proposition 2.11 is $\lambda_i(u_l) < \lambda_i(u_r)$. Apparently the left border of the rarefaction wave travels with speed $\lambda_i(u_l)$, while the right border travels with speed $\lambda_i(u_r)$. Yet, initially they are at the same point, namely at $x = 0$. Therefore for these borders to separate at all, the speed of the left border must be less than that of the right border. Luckily, also in the opposite case there may be solutions as we will see now.

Proposition 2.14 (Rankine-Hugoniot condition [HR02, Eq. 1.19]). *Consider a Riemann problem and assume that there is some $s \in \mathbb{R}$ such that*

$$s(u_r - u_l) = f(u_r) - f(u_l). \tag{2.14}$$

Then

$$u(x, t) = \begin{cases} u_l & \text{for } x < st, \\ u_r & \text{for } x \geq st \end{cases}$$

is a weak solution. Condition (2.14) is called the Rankine-Hugoniot condition.

Definition 2.15 (shock). The solution as detailed in Proposition 2.14 is called a *shock* and s is called the *shock speed*.

A shock with $s = \lambda_i(u_l) = \lambda_i(u_r)$ is called an *i-th contact discontinuity*.

Of course, contact discontinuities cannot appear for genuinely non-linear fields.

Note that for fixed u_l , Equation (2.14) is a (non-linear) system of equations for u_r and s . Therefore we have $\dim(\mathcal{S})$ equations for $\dim(\mathcal{S}) + 1$ variables and it is plausible that there is a whole path of solutions for (2.14). Indeed, there holds:

Proposition 2.16 ([Dafo9, Theorem 8.2.1]). *For a genuinely non-linear field (λ_i, r_i) of a conservation law with flux $f \in C^3(\mathcal{S}, \mathcal{S})$ there is a neighborhood $I \subset \mathbb{R}$ of 0 and both a path $q \in C^3(I, \mathcal{S})$ and a function $\sigma \in C^2(I, \mathbb{R})$, such that for each $\tau \in I$, $\hat{u} = q(\tau)$ and $s = \sigma(\tau)$, there is a shock between $\hat{u} = q(0)$ and \hat{u} with shock speed s . In addition there holds $\sigma(0) = \lambda_i(\hat{u})$ and $q'(0) = r_i$.*

We have now two kinds of weak solutions for a Riemann problem. Unfortunately usually there are multiple weak solutions, as can already be seen noting that there may be multiple solutions for the Rankine-Hugoniot condition in Proposition 2.14. Therefore we have to single out physically plausible ones. To this end the following definition is useful.

Definition 2.17 (Lax entropy condition). A shock with left state u_l , right state u_r and shock speed s is said to satisfy the *Lax entropy condition* if there is $i \in \{1, \dots, \dim(\mathcal{S})\}$ such that

$$\begin{aligned} \lambda_i(u_l) &> s > \lambda_i(u_r), \\ \lambda_j(u_l), \lambda_j(u_r) &< s && \text{for all } j < i, \\ s < \lambda_j(u_l), \lambda_j(u_r) &&& \text{for all } j > i. \end{aligned}$$

Such a shock is called an *i-th shock*.

This means in particular that an admissible shock can only occur between states u_l and u_r , where $\lambda_i(u_l) > \lambda_i(u_r)$ holds, so in the opposite situation of possible rarefaction solutions.

Actually there is a relation between *i*-th shocks and *i*-th rarefaction waves as the following propositions show.

Definition 2.18 (Lax curves). Consider a strictly hyperbolic balance law with flux function $f \in C^3(\mathcal{S}, \mathcal{S})$ and genuinely non-linear *i*-th field (λ_i, r_i) and a state $\hat{u} \in \mathcal{S}$. Let

$$(\nabla_{\mathcal{S}} \lambda_i)(\hat{u}) \cdot r_i(\hat{u}) > 0.$$

Then the curve

$$L_i^+(\cdot; \hat{u}) : I \rightarrow \mathcal{S}, \quad \xi \mapsto q(\xi),$$

given by (2.13) for $\xi \geq \lambda_i(\hat{u})$ and by the curve q of Proposition 2.16 for $\xi < \lambda_i(\hat{u})$ is called the *i-th forward Lax curve* of \hat{u} .

The curve

$$L_i^-(\cdot; \hat{u}) : I \rightarrow \mathcal{S}, \quad \xi \mapsto q(\xi),$$

given by (2.13) for $\xi \leq \lambda_i(\hat{u})$ and by the curve q of Proposition 2.16 for $\xi > \lambda_i(\hat{u})$ is called the *i-th backward Lax curve* of \hat{u} .

For the case of

$$(\nabla_{\mathcal{S}} \lambda_i)(\hat{u}) \cdot r_i(\hat{u}) < 0,$$

the notion of forward and backward are reversed.

Proposition 2.19 ([Dafo9, Theorem 8.2.2]). *Let $L : I \rightarrow \mathcal{S}$ be any of the (forward or backward) Lax curves of Definition 2.18. Then L is twice continuously differentiable.*

Proposition 2.20. *Consider a conservation law and a state $\hat{u} \in \mathcal{S}$, such that the i -th field of the conservation law is genuinely non-linear at \hat{u} . If a state \tilde{u} lies on the i -th forward Lax-curve of \hat{u} , then the Riemann problem with initial data*

$$u(x, 0) = \begin{cases} \hat{u} & \text{for } x < 0 \\ \tilde{u} & \text{for } x > 0, \end{cases}$$

can be solved by an i -th shock (if \tilde{u} is in the shock part of the Lax curve) or by an i -th rarefaction wave (otherwise).

If \tilde{u} is on the i -th backward Lax curve of \hat{u} , instead the Riemann problem with initial data

$$u(x, 0) = \begin{cases} \tilde{u} & \text{for } x < 0 \\ \hat{u} & \text{for } x > 0, \end{cases}$$

can be solved in the analogous way.

Proof. This follows from the construction of the Lax curves. Note that a shock of the “wrong” Lax curve (e.g. backward instead of forward in the first case above) is also a weak solution of the Riemann problem, but the hypothetical rarefaction wave from the wrong curve is not, because condition (2.11) is then not fulfilled. \square

As just stated, also the “wrong shocks” constitute weak solutions to a Riemann problem. To discard them we use the Lax entropy condition (Definition 2.17).

Proposition 2.21 ([GR14, Theorem 5.2]). *Consider a Riemann problem with initial data*

$$u(x, 0) = \begin{cases} u_l & \text{for } x < 0 \\ u_r & \text{for } x > 0. \end{cases}$$

Let the i -th field (λ_i, r_i) be genuinely nonlinear and let u_r lie in the shock part of the forward Lax curve through u_l , namely $u_r = L_i^+(\tau; u_l)$ for some $\tau \in \mathbb{R}$. If τ is small enough, then the corresponding shock between u_l and u_r satisfies the Lax entropy condition (Definition 2.17). If instead u_r lies in the shock part of the i -th backward Lax curve of u_l (and τ is small enough), the corresponding shock solution violates the Lax entropy condition.

Under some conditions, the collection of i -th Lax curves for all $i \in \{1, \dots, m\}$ can be used to construct solutions of Riemann problems, even if the initial states of the problem do not lie on the same Lax curves.

Theorem 2.22 ([HR02, Theorem 5.17, Laxe’s theorem], [Daf09, Theorem 9.4.1]). *Consider a strictly hyperbolic conservation law*

$$u_t + f(u)_x = 0,$$

where each component of the flux function $f_j, j \in \{1, \dots, m\}$ fulfills $f_j \in C^2(\mathcal{S})$ and all characteristic fields are either genuinely non-linear or linearly

degenerate in \mathcal{S} . Then for $u_l \in \mathcal{S}$ there exists a neighborhood D of u_l such that for all $u_r \in D$ the corresponding Riemann problem, given by

$$u(x, 0) = \begin{cases} u_l & \text{for } x < 0 \\ u_r & \text{for } x \geq 0, \end{cases}$$

has a unique solution in D consisting of up to m elementary waves, i.e., rarefaction waves, admissible shocks, or contact discontinuities.

The structure of this solution is as follows. There are constant states $u_l = u_0, \dots, u_m = u_r$ and when the i -th characteristic family is linearly degenerate, u_i is joined to u_{i-1} by an i -th contact discontinuity, while when the i -th characteristic family is genuinely non-linear, u_i is joined to u_{i-1} by either an i -th rarefaction wave or an admissible i -th shock.

For certain systems and initial conditions, the neighborhood D can be quite large. In Chapter 4 we will examine under which conditions the Riemann problem of the isentropic Euler equations can be solved in the way of Theorem 2.22 and find that the neighborhood D can be chosen rather large under mild assumptions.

2.1.3 Entropy solutions

As was already alluded to, weak solutions to a Riemann problem (and hence to a Cauchy problem) are not unique. One way to single out physically relevant solutions is to discard weak solutions that exhibit unphysical behavior. For example, one such unphysical behavior may be the production of energy out of thin air. Yet, mathematically this is a more general concept that can be encoded with the help of a so-called entropy-entropy-flux pair.

Definition 2.23 (entropy-entropy-flux pair). Let $\eta, \psi : \mathcal{S} \rightarrow \mathbb{R}$ be smooth and let η be convex. Consider a conservation law

$$u_t + f(u)_x = 0,$$

and let there hold

$$\eta'(u)f'(u) = \psi'(u).$$

Then (η, ψ) is called an *entropy-entropy-flux pair* of the conservation law.

With the help of entropy-entropy-flux pairs we can now single out certain weak solutions, that behave more physical than others.

Definition 2.24 (entropy solution). A (weak) solution u of the Cauchy problem is called an *entropy solution*, if it satisfies (in the weak sense)

$$\eta(u)_t + \psi(u)_x \leq 0, \tag{2.15}$$

for all entropy-entropy-flux pairs (η, ψ) of the conservation law.

A classical (and hence differentiable) solution actually fulfills (2.15) with equality. This means that entropy is constant along a classical

solution. This cannot hold for all weak solutions anymore, so it is relaxed to demand only that entropy doesn't increase¹. Unfortunately even entropy solutions are not unique when dealing with systems of conservation laws.

It is noteworthy (and not surprising, considering the name) that shocks fulfilling the Lax entropy condition (Definition 2.17) are entropy solutions of their Riemann problems.

2.1.4 Front-tracking

In this section we will cite an existence theorem for a Cauchy problem. Before we do so we must introduce a bit more nomenclature.

Definition 2.25. [total variation, bounded variation]

- Let $g : I \rightarrow \mathbb{R}^n$ be a function. The *total variation* $TV(g, I) \in \mathbb{R}_0^+ \cup \{\infty\}$ of g on an interval $I \subset \mathbb{R}$, is given by

$$TV(g, I) = \sup_{(x_i) \in P(I)} \sum \|g(x_i) - g(x_{i-1})\|_{L^1}, \quad (2.16)$$

where $P(I)$ is the set of all finite, strictly increasing tuples in I .

- g is said to be of *bounded variation* on I , if $TV(g, I) < \infty$.

We will not go into detail of the front-tracking technique and refer the reader to [Daf09, Chapter 14] and of course to the book [HR02]. Disregarding any details, we just name an existence and a uniqueness result for Cauchy problems with small total variation in the initial condition.

Theorem 2.26 ([HR02, Part of Theorem 6.6]). *Consider the strictly hyperbolic system of equations*

$$\begin{aligned} u_t + f(u)_x &= 0, \\ u(x, 0) &= u_0(x), \end{aligned}$$

and assume that $f \in C^2(\mathcal{S}, \mathcal{S})$ is such that each characteristic wave family is either linearly degenerate or genuinely nonlinear. If $TV(u_0, X)$ is sufficiently small, there exists a global weak solution u to this initial value problem. This solution may be constructed by the front-tracking algorithm described in [HR02, Section 6.1.]

The front-tracking technique consists of

- replacing the initial condition by a piecewise constant approximation whose deviation from u_0 is governed by some parameter $\delta > 0$,
- solving the Riemann problems arising between the constant pieces,

¹ Note that the physicist's notion of entropy corresponds to a negative entropy in the sense of Definition 2.23, see [Eva10, Remark on page 121], so that the physicist would instead say that entropy never decreases.

- replacing any rarefaction wave again by piecewise constant solutions (again, of deviation at most δ), which now are inadmissible shocks,
- and dealing with colliding discontinuities of all shocks generated either by the original Riemann problems or by the rarefaction approximation.

The weak solution whose existence is attested in Theorem 2.26 is then found as the limit of the approximate solution as $\delta \rightarrow 0$.

As we are dealing with balance laws, the preceding theorem is not sufficient for our goals. Yet, instead of presenting general results on balance laws, we will provide an existence result for the isentropic Euler equations at the end of the following section, where both boundary conditions and a source term are incorporated.

2.1.5 Boundary conditions

Up to now we always considered balance laws on $X = \mathbb{R}$, that is, the full real line. In applications one has boundaries of the domain of the balance law and junctions between multiple domains of hyperbolic balance laws. A general introduction is beyond the scope of this work and therefore we only supply some guidelines and refer to the literature on this topic, for example [LeVo2, Sections 3.11, 7, 21.8], [Dafo9, Sections 4.7, 5.6, 6.9] or [GR14, Chapter 5].

We will discuss boundary conditions for hyperbolic balance laws that have only non-zero eigenvalues and have constant characteristic structure, because the balance laws we investigate have this structure. This means that the number of negative and positive eigenvalues is constant. Let in the following $R \in \mathbb{N}$ be the number of positive eigenvalues (these correspond to right-moving waves, hence the name). This implies that $L = \dim(\mathcal{S}) - R$ is the number of negative eigenvalues because of hyperbolicity and our assumption that no eigenvalue vanishes. The balance laws we examine in this work do of course fulfill this requirement. With this we come to Cauchy problems on bounded domains.

Definition 2.27 (initial-boundary problem, [LeVo2, Section 3.11]). Let $\Psi : \mathcal{S} \rightarrow \mathbb{R}^R$, $\Pi : \mathcal{S} \rightarrow \mathbb{R}^L$, $X = [l, r] \subset \mathbb{R}$, $l < r$ and consider the following system of equations:

$$\begin{aligned} \partial_t u + \partial_x f(u) &= S(u), \\ u(x, 0) &= u_0(x), \\ \Psi(u(l, t)) &= 0, \\ \Pi(u(r, t)) &= 0. \end{aligned}$$

Such a problem is called an *initial-boundary problem*.

A slightly different but important version of the initial-boundary problem includes interaction between the boundaries.

Definition 2.28 (initial-boundary problem with feedback).

Let $\Theta : \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}^{\dim(\mathcal{S})}$, $X = [l, r] \subset \mathbb{R}$, $l < r$ and

$$\begin{aligned}\partial_t u + \partial_x f(u) &= S(u), \\ u(x, 0) &= u_0(x), \\ \Theta(u(l, t), u(r, t)) &= 0.\end{aligned}$$

Such a problem is called an *initial-boundary problem with feedback boundary conditions*.

The key difference to the initial-boundary problem without feedback is the possibility of different boundaries to interact. For example periodic boundary conditions are an instance of feedback boundary conditions. We will need feedback boundary conditions, when we introduce hyperbolic balance laws on networks in Section 2.1.6.

For weak solutions the boundary values $u(l, t)$ and $u(r, t)$ are not well-defined. Yet, restricting to suitable function spaces (e.g. $\mathcal{H}^1((l, r))$ or subspaces thereof) makes it possible to define the boundary values almost everywhere via the trace operator. We will not delve into this here, refer to [Daf09, Section 4.7] instead and assume enough regularity to have well-defined boundary conditions from here on out.

To get a feel for boundary conditions we revisit the advection equation.

Example 2.29 (advection equation with boundary conditions). Let $X = [l, r]$, $\mathcal{T} = \mathbb{R}_0^+$, $a > 0$, $\mathcal{S} = \mathbb{R}$, $u_0 : [l, r] \rightarrow \mathcal{S}$, $w : \mathcal{T} \rightarrow \mathcal{S}$, $w(0) = u_0(l)$ and

$$\begin{aligned}u_t + au_x &= 0, \\ u(x, 0) &= u_0(x), \\ \Psi(u(l, t)) &= u(l, t) - w(t) = 0.\end{aligned}$$

Also revisiting Figure 2.1, it seems natural to extend the method of lines to the boundary and define additional characteristic curves that don't start at $t = 0$ but at some later time t_0 . These do not start at the initial condition but at the boundary. Setting

$$x(t) = l + a(t - t_0)$$

leads to the solution

$$u(x, t) = \begin{cases} u_0(x - at) & \text{for } x - at \in (l, r), \\ w(t_0) & \text{for } x - a(t - t_0) = l \end{cases}$$

which is showcased in Figure 2.2.

This figure also suggests that for suitable solutions the boundary and initial conditions should agree where they meet, which is called a *consistency condition*. In addition, the figure makes clear why no initial condition at the right boundary was given. Characteristic curves starting at the right boundary immediately leave the domain and are hence irrelevant. The situation is reversed for $a < 0$. Here the characteristic curves are left-moving and only the right boundary condition have to

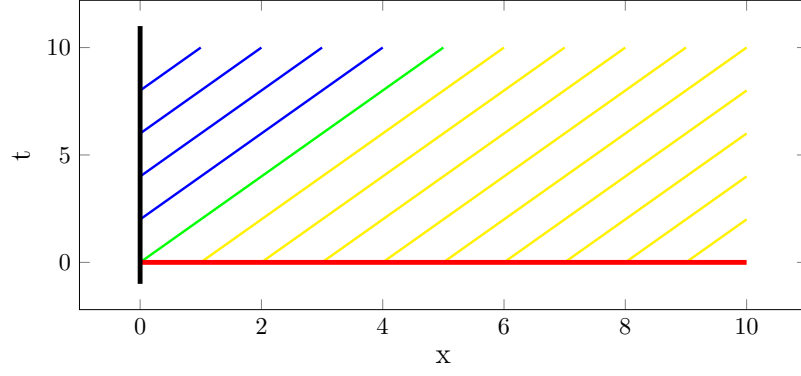


Figure 2.2: Different characteristic curves for $a = 2$. Blue curves start at the (black) boundary condition, yellow curves start at the (red) initial condition and the green one starts at their intersection.

be specified. For systems the situation is more complicated because both positive and negative eigenvalues can appear at the boundary.

To ease notation we consider a balance law on a half-space, so that we can forego a right boundary condition. We also only present a result for a non-scalar but linear conservation law. Nevertheless we will use notation that can be applied for non-linear balance laws.

Proposition 2.30 ([GR14, Chapter V, Lemma 1.1]). *Let $\dim(\mathcal{S}) = m$, $X = \mathbb{R}^+$, $0 < k \leq m$, let $\Psi : \mathcal{S} \rightarrow \mathbb{R}^{m-k}$ and $g : \mathcal{T} \rightarrow \mathbb{R}^{m-k}$. Consider the initial-boundary problem*

$$\begin{aligned} u_t + f(u)_x &= 0, \\ u(x, 0) &= u_0(x), \\ \Psi(u(0, t)) &= g(t), \end{aligned} \tag{2.17}$$

where the functions f and Ψ are linear. Let the eigenvalues λ_i of f' fulfill the inequalities

$$\lambda_1(u) \leq \dots \leq \lambda_k(u) < 0 < \lambda_{k+1}(u) \leq \dots \leq \lambda_m(u)$$

for all $u \in \mathcal{S}$. Let r_1, \dots, r_m be the corresponding eigenvector fields. Consider the decomposition of u at the boundary

$$u(0, t) = a_1(u)r_1(u) + \dots + a_m(u)r_m(u),$$

and partition the coefficient functions like

$$\begin{aligned} A_{\text{out}} &= (a_1, \dots, a_k), \\ A_{\text{in}} &= (a_{k+1}, \dots, a_m). \end{aligned}$$

Then (2.17) is well-posed if $\frac{\partial \Psi(A_{\text{out}}, A_{\text{in}})}{\partial A_{\text{in}}}$ is invertible. This can equivalently be expressed as

$$\det(\Psi'(u) \cdot r_{k+1}, \dots, \Psi'(u) \cdot r_m) \neq 0. \tag{2.18}$$

It is plausible to assume a similar condition also for non-linear systems, as this proposition can be interpreted to mean that the boundary

condition must fix the coefficients of the $m - k$ waves that enter the domain uniquely. This seems plausible also in the nonlinear case and indeed there are related results, see for example [Ama97, Theorems 1.1 and 1.2] or [CG10, Theorem 2.2]. But note that—just as without boundaries—an initial condition of small total variation must be assumed. A corresponding result from [Gug+12] for a junction in a network is presented in the next section as Proposition 2.34.

2.1.6 Junctions and networks

Instead of just boundary conditions, we also want to couple different balance laws together at junctions between them. This is needed in order to define balance laws whose spatial domains form a graph with coupling conditions at the nodes of the graph. The coupling conditions generalize the boundary conditions of plain balance laws. All coupling conditions we consider model some kind of flow balance, which means that the flow of some quantity over the junction must be conserved, while some other quantity is continuous over the junction.

The general setting is as follows. We consider a finite, directed and connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with vertex set \mathcal{V} and edge set \mathcal{E} . An edge $e \in \mathcal{E}$ consists of an interval $[0, \ell_e]$ together with a hyperbolic balance law. To a vertex $v \in \mathcal{V}$ we assign a virtual point x_v and for an edge s starting at v we identify the point $0 \in [0, \ell_s]$ with x_v . For an edge f ending in v we identify $\ell_f \in [0, \ell_f]$ with x_v (here f means “final point”). The reasoning behind this identification is to interpret the graph as a subset of some \mathbb{R}^n , where the start and endpoint of an edge are actually the points where the vertices are located. This image is most natural for planar graphs, where \mathbb{R}^2 can be used without intersection of edges outside of vertices.

For each vertex v let \mathcal{E}_v be the set of all edges connected to v and let $\mathcal{E}_v^s, \mathcal{E}_v^f \subset \mathcal{E}_v$ be the edges starting and ending in v . The purpose of the vertices is to supply coupling conditions for the edges $e \in \mathcal{E}_v$. The full problem on the graph \mathcal{G} can be defined as follows:

Definition 2.31 (graph problem). The *initial-boundary problem on a graph* is given by a collection of balance laws, defined on the interval $(0, \ell_e)$, for each edge $e \in \mathcal{E}$,

$$u_t^e + f^e(u^e)_x = S^e(u^e), \quad u^e(x, 0) = u_0^e(x),$$

together with a coupling condition for each vertex $v \in \mathcal{V}$,

$$\Theta \left((u^s(0, t))_{s \in \mathcal{E}_v^s}, u^f(\ell_e, t) \right)_{f \in \mathcal{E}_v^f} = 0.$$

The coupling conditions must of course fulfill conditions for well-posedness that are related to those of boundary conditions. Luckily, instead of working them out explicitly for coupling conditions, we can transform the coupling conditions into boundary conditions. To see that, we use two lemmas.

Lemma 2.32. Let $[l_0, r_0], [l_1, r_1] \subset \mathbb{R}$ be intervals, let $\mu = \frac{r_1 - l_1}{r_0 - l_0}$ be the quotient of their lengths and consider the linear diffeomorphisms

$$a) \phi_a : [l_0, r_0] \rightarrow [l_1, r_1], x \mapsto \mu(x - l_0) + l_1,$$

$$b) \phi_b : [l_0, r_0] \rightarrow [l_1, r_1], x \mapsto \mu(r_0 - x) + l_1.$$

Then there is an equivalence (to be explained in the proof) of (feedback) initial-boundary problems (Definition 2.28) on $[l_0, r_0]$ and $[l_1, r_1]$ and this equivalence is induced by $\phi_{a/b}$.

Proof. Consider a feedback initial-boundary problem on $[l_1, r_1]$, given by

$$\begin{aligned} u_t + f(u)_x &= 0, \\ u(x, 0) &= u_0(x), \\ \Theta(u(l_0, t), u(r_0, t)) &= 0. \end{aligned} \tag{2.19}$$

For ϕ_a define the corresponding problem on $[l_0, r_0]$ by

$$\begin{aligned} y_t + \frac{1}{\mu} f(y)_x &= 0, \\ y(x, 0) &= u_0(\phi_a(x)), \\ \Theta(y(l_0, t), y(r_0, t)) &= . \end{aligned} \tag{2.20}$$

For ϕ_b define instead

$$\begin{aligned} y_t - \frac{1}{\mu} f(y)_x &= 0, \\ y(x, 0) &= u_0(\phi_b(x)), \\ \Theta(y(r_0, t), y(l_0, t)) &= 0. \end{aligned} \tag{2.21}$$

Note the interchange of the boundary conditions in (2.21). Note also that the eigenvalues of the Jacobian change sign under the transformation ϕ_b . A solution u of (2.19) induces solutions y_a, y_b of (2.20) and (2.21) by $y_{a/b}(y, t) = u(\phi_{a/b}(x), t)$. \square

This lemma allows us to scale the spatial domain of a balance law as well as revert its direction. The next lemma instead allows us to stack balance laws that share the same domain.

Lemma 2.33. *Let*

$$\begin{aligned} u_t^a + f^a(u^a)_x &= 0, \\ u^a(x, 0) &= u_0^a(x), \\ \Theta^a(u^a(l, t), u^a(r, t)) &= 0, \end{aligned} \tag{2.22}$$

and

$$\begin{aligned} u_t^b + f^b(u^b)_x &= 0, \\ u^b(x, 0) &= u_0^b(x), \\ \Theta^b(u^b(l, t), u^b(r, t)) &= 0, \end{aligned} \tag{2.23}$$

be two initial-boundary problems with $\dim(\mathcal{S}^a) = m$ and $\dim(\mathcal{S}^b) = n$. Then the combined system for $w = (w_1, w_2)$,

$$\begin{aligned} w_t + \begin{pmatrix} f^a(w_1) \\ f^b(w_2) \end{pmatrix}_x &= 0, \\ w_1(x, 0) &= u_0^a(x), \\ w_2(x, 0) &= u_0^b(x), \\ \Theta^a(w_1(l, t), w_1(r, t)) &= 0, \\ \Theta^b(w_2(l, t), w_2(r, t)) &= 0, \end{aligned} \tag{2.24}$$

is an initial-boundary problem of dimension $m + n$. Genuine non-linearity or linear degeneracy of a characteristic pair of (2.22) or (2.23) carry over to (2.24) and pairs of solutions (u^a, u^b) to (2.22) and (2.23) form a solution $w = (u^a, u^b)^T$ of (2.24) and vice versa.

Proof. Immediate from the definition. \square

With these lemmas we can rewrite a coupling condition at a vertex v in a graph problem (Definition 2.31) as a boundary condition, by first folding over all edges ending in v and then stacking all edge balance laws into a combined balance law. Then the coupling condition becomes a boundary condition.

In addition, the lemmas even allow us to treat the whole of a graph problem as a hyperbolic balance law with feedback boundary conditions. To this end, we first transform all balance laws on the edges to have the domain $X^e = [0, 1]$. Then we stack all of them together. As the coupling conditions are allowed to couple the start and end of the domains, it is no problem if a vertex in the graph problem lies at the start of one edge and the end of another. This interpretation of the graph problem is attractive from a theoretical point of view, because it makes the theory of ordinary initial-boundary problems applicable. Yet, in practice it is usually better to consider vertices and their corresponding coupling conditions individually, as for example properties like that of Proposition 2.30 are more tractable in this setting.

To close this section, we mention a well-posedness result for the isentropic Euler equations defined on a graph of a single node, to which a number of edges is connected. This is of course relevant to this work because we are interested in simulating gas dynamics in a network. The cited theorem is much more extensive than stated here and not only shows existence of a solution.

Proposition 2.34 (part of [Gug+12, Theorem 3.8]). *Let $m = \dim(\mathcal{S}) = 2$ and consider $n \geq 2$ hyperbolic balance laws, each defined on $\mathcal{T} = [0, T]$ and $X = \mathbb{R}_0^+$ that are coupled at a junction via a coupling function $\Psi \in C^1(\mathbb{R}^{2n}, \mathbb{R}^n)$:*

$$\begin{aligned} u_t^l + f^l(u^l)_x &= g^l(u^l), \\ u^l(x, 0) &= u_0^l(x), \\ \Psi(u(0, t)) &= \Xi, \end{aligned} \tag{2.25}$$

with $l \in \{1, \dots, n\}$ and $\Xi \in \mathbb{R}$. Let there be a constant solution $\bar{u} = (\bar{u}^l)_{l=1, \dots, n}$ of (2.25) that fulfills $\Psi(\bar{u}) = \Xi$.

Let for each l hold $f^l \in C^4(\mathcal{S}, \mathcal{S})$ and let $(f^l)'(\bar{u})$ have one negative and one positive eigenvalue and let each characteristic field be either genuinely nonlinear or linearly degenerate.

Let also $G = (g^l)_l$ be Lipschitz continuous in u and let $TV(G \circ y, X) \leq L$ for all $y \in \mathcal{Y}_\delta$ with

$$\mathcal{Y}_\delta = \left\{ y \mid y - \bar{u} \in L^1(X, \mathcal{S}^n), TV(y, X) \leq \delta \right\},$$

and some $L > 0$.

Lastly let Ψ fulfill Condition (2.18).

Then there is a solution of (2.25) of small total variation.

2.2 NUMERICAL METHODS FOR HYPERBOLIC BALANCE LAWS

As it is impossible to find closed expressions for solutions of (systems of) hyperbolic balance laws in all but the easiest cases, we need algorithms to compute numerical approximations of solutions. While there is much to be said about this topic, we will again only introduce what is needed later on and refer to text books [GR14], [LeV92] and most notably [LeV92]. Again we are only interested in the case of one space dimension.

Here we briefly introduce finite volume methods to show the relevant challenges arising in the numerical treatment of hyperbolic balance laws. As we did in the last section, we only consider conservation laws at first.

To compute an approximate solution of a system of an initial-boundary problem with a finite volume method over a time span $[0, T]$ on a bounded interval $[l, r]$ we introduce a time step size $\Delta t = \frac{T}{N}$ for some $N \in \mathbb{N}$ and a spatial step size $\Delta x = \frac{r-l}{J}$ for some $J \in \mathbb{N}$. As discretization points we then choose $(t^n = n\Delta t)_{n=0, \dots, N}$ and $(x_j = l + j\Delta x)_{j=0, \dots, J}$. This divides the interval $[l, r]$ into J different cells I_j , given by

$$I_j = [x_{j-1}, x_j], \text{ for } j = 1, \dots, J. \quad (2.26)$$

To leverage the conservation property of the conservation law (2.2) (with $S = 0$), we choose as discretization of the function u values u_j^n that shall represent cell averages,

$$u_j^n \approx \frac{1}{\Delta x} \int_{I_j} u(x, t^n) dx.$$

An approximate solution in the interior of the domain of an initial-boundary problem

$$\begin{aligned} \partial_t u + \partial_x f(u) &= 0, \\ u(x, 0) &= u_0(x), \\ \Psi(u(l, t)) &= 0, \\ \Pi(u(r, t)) &= 0 \end{aligned} \quad (2.27)$$

can be found by first discretizing the initial condition,

$$u_j^0 = \frac{1}{\Delta x} \int_{I_j} u(x, t^0) dx,$$

and then computing u_j^{n+1} from u_j^n via: Equation (2.2),

$$\begin{aligned} 0 &= \partial_t \int_{x_{j-1}}^{x_j} u(x, t) dx + f(u(x_j, t^n)) - f(u(x_{j-1}, t)) \\ \Rightarrow 0 &= \partial_t u_j^n + \frac{f(u(x_j, t^n)) - f(u(x_{j-1}, t^n))}{\Delta x} \\ \Rightarrow u_j^{n+1} &= u_j^n - \frac{1}{\Delta x} \int_{t^n}^{t^{n+1}} f(u(x_j, t)) - f(u(x_{j-1}, t)) dt \\ &= u_j^n \\ &\quad - \frac{\Delta t}{\Delta x} \left(\frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} f(u(x_j, t)) dt \right. \\ &\quad \left. - \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} f(u(x_{j-1}, t)) dt \right). \end{aligned} \quad (2.28)$$

To use this equation we need to approximate both the time integrals and thereby the evaluations of f , as we can only access the cell averages of u and not the values at the cell boundaries.

Of course it is desirable that an approximate solution converges against a solution of (2.27), usually in the $\|\cdot\|_{L_1}$ -norm, which lends itself naturally to the definition of weak solutions. A finite volume scheme now consists of choices for the numerical approximations of the time derivative and the flux evaluations. Let us call $u^n = (u_j^n)_{j=1, \dots, J}$. An explicit scheme to solve (2.27) is given by a function

$$\begin{aligned} H : \mathbb{R}^{J+1} &\rightarrow \mathbb{R}^{J+1}, \\ u^n &\mapsto u^{n+1} = H(u^n), \end{aligned}$$

which computes approximation values at the next time point from the approximation at the current time point in an explicit way. Usually the scheme function H is extremely sparse in the sense that any u_j^{n+1} depends only on a few coefficients $u_{j-k}^n, \dots, u_{j+k}^n$.

A scheme function of the form derived in Equation (2.28) is said to have conservation form.

Definition 2.35 (conservation form, numerical flux, [LeVo2, Section 4.1]). Consider an explicit scheme for the hyperbolic conservation law $u_t + (f(u))_x = 0$. The scheme is said to be in *conservation form*, if it can be written as

$$u_j^{n+1} = H(u^n)_j = u_j^n - \frac{\Delta t}{\Delta x} \left(F_{j+\frac{1}{2}}^n - F_{j-\frac{1}{2}}^n \right),$$

where

$$F_{j+\frac{1}{2}}^n = F(u_{j-k+1}^n, \dots, u_{j+k}^n)$$

for some function F . F is then called the *numerical flux* of the scheme and the scheme is called a *k-point scheme*.

The next definition will be useful in Section 3.3.

Definition 2.36 (viscous form, numerical viscosity, [LeVo2, Section 4.1]). The explicit scheme H is said to have *viscous form*, if it can be written as

$$H(u^n)_j = u_j^n - \frac{\Delta t}{\Delta x} \left(\frac{f(u_{j+1}^n) - f(u_{j-1}^n)}{2} \right) + \frac{1}{2} \left(Q_{j+\frac{1}{2}}^n (u_{j+1}^n - u_j^n) - Q_{j-\frac{1}{2}}^n (u_j^n - u_{j-1}^n) \right),$$

where

$$Q_{j+\frac{1}{2}}^n = Q(u_{j-k+1}^n, \dots, u_{j+k}^n)$$

for some function Q . Q is then called the *numerical viscosity* of the scheme.

The numerical treatment of boundary conditions can often be done via a so-called ghost cell approach [LeVo2, Chapter 7], where one uses the definition of the boundary conditions (Ψ and Π in Equation (2.27)) to extrapolate values of “ghost cells” outside the domain of the conservation law, so that the scheme can be used near the interior and the boundary conditions are satisfied. We will use this method in Section 3.3.2.

2.2.1 Schemes for linear hyperbolic balance laws

The simplest example of a scheme, suitable for the advection equation is the Upwind scheme:

Example 2.37 (Upwind scheme). Consider the scalar initial-boundary problem on $X = [l, r]$,

$$\begin{aligned} u_t + au_x &= 0, \\ u(x, 0) &= u_0(x), \\ u(l, t) &= w_l(t) \text{ if } a > 0, \\ u(r, t) &= w_r(t) \text{ if } a < 0 \end{aligned}$$

and the discretization $\Delta t, \Delta x$. An explicit scheme is given by the *Upwind scheme*,

$$u_j^{n+1} = H(u^n)_j = \begin{cases} u_j^n - a \frac{\Delta t}{\Delta x} (u_j^n - u_{j-1}^n) & \text{for } a > 0, \\ u_j^n - a \frac{\Delta t}{\Delta x} (u_{j+1}^n - u_j^n) & \text{for } a < 0 \end{cases}$$

for $1 \leq j \leq J-1$. As numeric boundary values one can simply set $u_0^n = w_l(t^n)$ or (depending on the sign of a) $u_J^n = w_r(t^n)$.

Unfortunately it turns out that explicit schemes for hyperbolic balance laws must obey the so-called CFL condition², relating the discretization and analytical properties of the balance law. It was originally stated as

² The CFL condition is named after the mathematicians Richard Courant (08.01.1888 - 27.01.1972), Kurt Friedrichs (28.09.1901 - 31.12.1982) and Hans Lewy (20.10.1904 - 23.08.1988).

Proposition 2.38 (CFL condition, [LeVo2, Section 4.4]). “A numerical method can be convergent only if its numerical domain of dependence contains the true domain of dependence of the PDE, at least in the limit as Δt and Δx go to zero.”

Here the domain of dependence of a value $u(y, t_1)$ of a solution of the (possibly non-scalar) hyperbolic conservation law

$$u_t + f(u)_x = 0$$

is the set of all values $u(x, t_0)$ with $t_0 < t_1$, that can influence the value $u(y, t_1)$. As we know that information travels at most with velocity

$$v = \max \{ \lambda \mid |\lambda| \text{ is an eigenvalue of the Jacobian } \mathcal{J}_f \text{ of } f \},$$

the domain of dependence of a linear hyperbolic balance law can be estimated.

This estimate can be used to phrase necessary conditions for the CFL condition and informs the following definition.

Definition 2.39 (Courant number). For a system of hyperbolic balance laws with eigenvalue fields $\lambda_1, \dots, \lambda_m$ and a discretization with step sizes Δt and Δx the *Courant number*, often denoted by \mathcal{C} , is given by

$$\mathcal{C} = \frac{\max_{u \in \mathcal{S}} (|\lambda_1(u)|, \dots, |\lambda_m(u)|)}{\frac{\Delta x}{\Delta t}}.$$

Note that the definition is only meaningful for hyperbolic balance laws, whose eigenvalue fields are bounded, as is the case for the advection equation.

For balance laws with unbounded eigenvalue fields one can still define a Courant number by taking the maximum only over some subset $B \subset \mathcal{S}$. This Courant number is then only meaningful, if B contains all values that are attained by solutions of the balance law. For gas equations studied in later chapters, we will impose a so-called *sub-sonic condition*, that will effectively constrain solutions to have useful Courant numbers.

Yet, for the Upwind scheme for the advection equation, the CFL condition can be reformulated as

Proposition 2.40. *The Courant number \mathcal{C} of the Upwind scheme must necessarily fulfill*

$$\mathcal{C} \leq 1$$

for the scheme to be stable. This condition is equivalent to

$$a < \frac{\Delta x}{\Delta t} \text{ and } \Delta t < \frac{\Delta x}{a}.$$

This means the “numerical velocity” $\frac{\Delta x}{\Delta t}$ must be not smaller than the “analytical velocity”, constraining the time step size in relation to the spatial step size.

Different explicit schemes have slightly different CFL conditions, yet all of them have this general form, that restricts the time step size by some multiple of the spatial step size.

The Upwind scheme is a first order scheme, meaning that it converges to a weak solution of a linear conservation law for vanishing discretization step size and fulfills

$$u(x, t^n + \Delta t) - H_{\text{upwind}}(u^n) = O((\Delta t)^g). \quad (2.29)$$

with $g = 1$. Another scheme for the advection equation is given by the Lax-Wendroff scheme. It fulfills Equation (2.29) with $g = 2$.

Example 2.41 (Lax-Wendroff scheme). Consider the same situation as for the upwind scheme, namely a scalar linear initial-boundary problem with velocity a . A second order scheme, named the *Lax-Wendroff scheme* is given by

$$u_j^{n+1} = u_j^n - \text{sign}(a)\mathcal{C} \left(u_j^n - u_{j-1}^n \right) - \frac{1}{2}\mathcal{C}(1 - \mathcal{C}) \left(u_{j+1}^n - 2u_j^n + u_{j-1}^n \right),$$

where \mathcal{C} is the Courant number. Unfortunately it suffers from the same CFL condition as the upwind scheme, namely $\mathcal{C} \leq 1$. In addition, it doesn't converge to a weak solution of a conservation law for vanishing step sizes.

This last drawback of the Lax-Wendroff scheme is related to the fact that the scheme fails to stabilize the total variation (see Definition 2.25). In order to correctly mimic extrema and non-smooth points of weak solutions, it is desirable for a scheme to be total variation diminishing.

Definition 2.42 (total variation diminishing, TVD [LeVo2, Definition 6.1]). A scheme, given by its scheme function H , is called *total variation diminishing*, usually abbreviated *TVD*, if it fulfills

$$\text{TV}(H(u^n)) \leq \text{TV}(u^n)$$

for every time step n .

The upwind scheme above is TVD. Yet, the Lax-Wendroff scheme is not. Due to this reason, the Lax-Wendroff scheme is not convergent. We will run into and address this problem in Section 3.3.

The upwind and Lax-Wendroff schemes in the presented form can only be used for linear hyperbolic conservation laws. There are of course also schemes for non-linear systems, for example the so-called CWENO-schemes, one of which is used as a reference in Section 4.3.1. A full description of the methods used for non-linear systems is beyond the scope of this thesis and we refer to [LeVo2] for more details and especially [Kol14] and [NKS18] for the CWENO3 scheme, briefly used in Section 4.3.1 and described in Section 2.2.4.

The last property of numerical schemes mentioned in this section shall be a discrete analog to entropy-entropy-flux pairs.

Definition 2.43 (entropy-stable scheme, [LeVo2, Section 12.11]). Let η, ψ be an entropy-entropy-flux pair of a hyperbolic conservation law. A scheme H is said to be *entropy-stable*, if there is a function $\phi : \mathcal{S}^{2k} \rightarrow \mathbb{R}$ that fulfills

$$\phi(u, \dots, u) = \psi(u) \text{ for all } u \in \mathcal{S}$$

and there holds

$$\begin{aligned} \eta(H(u)_j) &\leq \eta(u_j) \\ &\quad - \frac{\Delta t}{\Delta x} (\phi(u_{j-k+1}, \dots, u_{j+k}) - \phi(u_{j-k}, \dots, u_{j+k-1})). \end{aligned}$$

If equality holds, the scheme is called *entropy-conserving*.

2.2.2 Treatment of source terms and splitting

Before we end our discussion of explicit schemes we will briefly introduce a method to treat actual balance laws, instead of just conservation laws. We do so by retracing the steps of [LeVo2, Chapter 17]. Therefore we must somehow incorporate the source term in Definition 2.1 into a scheme. One way of doing this, is to split the numerical method for the balance law into two parts, one part that advances the underlying conservation law (setting the source term to zero) and one part that exclusively deals with the source term. We will restrict to a linear balance law, as this is the setting we are interested in, later in Chapter 3.

Let therefore $\dim(\mathcal{S}) = m$, $A, F \in \mathbb{R}^{m \times m}$ and consider the linear balance law

$$\begin{aligned} u_t + Au_x &= -Fu, \\ u(x, 0) &= u_0(x). \end{aligned}$$

We rewrite this as

$$u_t = \mathcal{A}u,$$

where $\mathcal{A} = -A\partial_x - F$ is a linear operator. This looks like a ordinary linear differential equation, defined in some function space (and actually it is, see Section 2.4 for an introduction to semigroups, which make this line of thought rigorous) and we can formally write its solution as

$$u(t) = \exp(\mathcal{A}t)u_0 = \exp((-A\partial_x - F)t)u_0.$$

If the operators $-A\partial_x$ and F commuted, we could further rewrite this as

$$u(t) = \exp(\mathcal{A}t)u_0 = \exp(-tF) \exp(-tA\partial_x)u_0.$$

This form suggests solving first the problem involving $A\partial_x$ and then solving the problem involving F .

Without F , the first problem is just the linear conservation law

$$u_t + Au_x = 0,$$

which we can solve with a scheme of our choice. Let us call its scheme function PDE, such that we can write a single step as $u^{n+1} = \text{PDE}(u^n)$. The second problem, without $A\partial_x$ has the form

$$u_t = -Fu,$$

which is an actual ordinary differential equation and can be solved with an appropriate method or even by evaluating the matrix exponential $\exp(-tF)$. Let us call its scheme function ODE.

A scheme for the whole balance law is then given by

$$u^{n+1} = H(u^n) = \text{ODE} \circ \text{PDE}(u^n).$$

Of course the operators in $A\partial_x$ and F usually do not commute and hence this scheme introduces errors. It turns out, that it is still of first order, if both schemes are of first order.

A better way of splitting is Strang splitting, where one first makes a half-step of the ODE scheme, then a full step of the PDE scheme and a half-step of the ODE scheme again. If both individual schemes are at least of second order, then the Strang-splitting scheme is as well. This can be derived by analyzing the power series of the exponential. The derivation can be found in [LeVo2, Chapter 17].

In Section 3.3 We will use Strang-splitting to define certain numerical methods to solve a system of advection equations with linear source term.

2.2.3 IBOX scheme

The CFL condition is very restrictive in both power and gas networks. While we circumvent the problem in power networks in Chapter 3 by replacing the relevant hyperbolic balance law with something else, the problem remains in gas networks. To overcome it, we make use of an implicit scheme, the implicit box scheme introduced in [KLB10]. Its scheme function H is now implicit, meaning, it is a function $H : \mathbb{R}^{J+1} \times \mathbb{R}^{J+1} \rightarrow \mathbb{R}^{J+1}$, $(u^n, u^{n+1}) \mapsto H(u^n, u^{n+1})$ and the next time step u^{n+1} is determined by finding a zero of H . This can be done by a suitable Newton-type method. The precise definition of the box scheme is given as follows.

Definition 2.44 (IBOX). Consider a hyperbolic balance law

$$u_t + f(u)_x = S(u).$$

The *implicit box scheme* on a lattice (t^n, x_j) is given by

$$\begin{aligned} H(u^n, u^{n+1})_j = & \frac{u_{j-1}^{n+1} + u_j^{n+1}}{2} - \frac{u_{j-1}^n + u_j^n}{2} \\ & + \frac{\Delta t}{\Delta x} \left(f(u_j^{n+1}) - f(u_{j-1}^{n+1}) \right) \\ & - \frac{\Delta t}{2} S(u_j^{n+1}) + S(u_{j-1}^{n+1}). \end{aligned}$$

For an implicit scheme existence and uniqueness must be shown. These properties along with results on stability and the TVD property can be found in [KLB10] as well as in [Kol11]. The key property of the implicit box scheme for our purposes is its “inverted” CFL condition. It has the following form

Proposition 2.45 ([Kol11, Proposition 4.2]). *For the implicit box scheme to be stable there must necessarily hold*

$$\Delta t \geq \frac{\Delta x}{2\lambda_{\min}}.$$

Here λ_{\min} denotes the minimum of the set $\{|\lambda| \mid \lambda \text{ is an eigenvalue of } \mathcal{J}_f\}$.

Handling of boundary conditions in the implicit setting is straightforward. Boundary conditions can simply be added as a function of which a zero must be found and hence for each boundary condition the dimension of the space in which a zero must be found is incremented by one.

2.2.4 CWENO3 scheme

In Section 4.3.1, the CWENO3 scheme is used, which is of a different kind to the ones presented so far. Namely, it is not total variation diminishing, although it is convergent and still captures extrema and discontinuities well. We follow along [Kol14] to introduce the method.

The scheme follows the REA-algorithm, short for “reconstruct-evolve-average”, see [LeVo2, Section 4.10]. Here, the cell averages $(u_j^n)_{j=1,\dots,J}$ of the current time step are used to reconstruct a piecewise polynomial representation $\hat{u}(x, t^n)$ of the solution $u(x, t^n)$ to a conservation law at the current time step t^n . This representation is then evolved with some time integration technique to an approximation $\check{u}(x, t^{n+1})$ of the solution $u(x, t^{n+1})$ at the next time step, where it is then again averaged in to the cell averages $(u^{n+1})_{j=1,\dots,J}$ at the next time step. The CWENO3 scheme provides a method for the reconstruction phase of the algorithm. Therefore we consider a function $u : X \rightarrow \mathcal{S}$, because the time-dependency is irrelevant for the reconstruction phase. In the cell I_j the CWENO3 scheme chooses a polynomial $P = P(x)$, which is constructed from the cell averages u_{j-1}, u_j, u_{j+1} . P is in turn a convex combination of three other polynomials P_L, P_C, P_R ,

$$P(x) = \omega_L P_L(x) + \omega_C P_C(x) + \omega_R P_R(x),$$

with $\omega_L, \omega_C, \omega_R \geq 0$ and $\omega_L + \omega_C + \omega_R = 1$, to be determined later.

P_L is the unique linear polynomial that fulfills

$$\int_{I_{j-1}} P_L(x) dx = u_{j-1} \text{ and } \int_{I_j} P_L(x) dx = u_j$$

and P_R is the unique linear polynomial that fulfills

$$\int_{I_j} P_R(x) dx = u_j \text{ and } \int_{I_{j+1}} P_R(x) dx = u_{j+1}.$$

So P_L and P_R preserve the cell averages over two cells.

P_C is determined by yet another polynomial P_{opt} , which is the unique quadratic polynomial, that preserves all three cell averages in I_{j-1} , I_j and I_{j+1} ,

$$\begin{aligned} \int_{I_{j-1}} P_{\text{opt}}(x) dx &= u_{j-1}, \\ \int_{I_j} P_{\text{opt}}(x) dx &= u_j, \\ \int_{I_{j+1}} P_{\text{opt}}(x) dx &= u_{j+1}. \end{aligned}$$

Finally, P_C is determined by

$$P_{\text{opt}} = c_L P_L + c_C P_C + c_R P_R,$$

where $c_L, c_C, c_R \geq 0$, $c_L + c_C + c_R = 1$ must be chosen for the method. We use $c_L = c_R = 0.25$ and hence $c_C = 0.5$ as in [Kol14].

For the weights ω_i , $i \in \{L, C, R\}$, we cannot simply choose c_i , because this would introduce oscillations whenever there are extrema or discontinuities in the function u to be reconstructed. Instead, for $i \in \{L, C, R\}$, they must be chosen as

$$\omega_i = \frac{\alpha_i}{\sum_{m \in \{L, C, R\}} \alpha_m} \text{ and } \alpha_i = \frac{c_i}{(K\Delta x^q + \text{IS}_i(u_{j-1}, u_j, u_{j+1}))^p},$$

for some positive constants K, p, q and so-called smoothness indicators IS_i . These are given by

$$\text{IS}_i(u_{j-1}, u_j, u_{j+1}) = \sum_{k=1}^2 \int_{I_j} \Delta x^{2k-1} \left(\frac{d^k P_i}{dx^k}(x) \right)^2 dx.$$

For these choices we have the following theorem:

Theorem 2.46 ([Kol14, Theorems 2.1 2.2, 3.3 and 3.4]). *For $K > 0$, $p \geq 1$, $q \leq 3$, $pq \leq 2$ and $u \in C^3(X)$ the reconstruction P fulfills*

$$P(x) - u(x) \in O(\Delta x^3).$$

If u has a discontinuity in I_{j-1} but fulfills $u \in C^3(I_j \cup I_{j+1})$, then

$$P(x) - u(x) \in O(\Delta x^2),$$

for $x \in I_j$.

To have a complete scheme for solving a conservation law, we must use the reconstruction to obtain values $u_{j+\frac{1}{2}}^+(t) = P_j(x_{j-1})$ and $u_{j+\frac{1}{2}}^-(t) = P_{j-1}(x_{j-1})$ at the cell boundaries and solve the Riemann problem between them to obtain a flux between cells, which can then be used in Equation (2.28) to evolve u . Finding the analytical solution to the Riemann problem can be difficult and is usually computationally expensive. Therefore one can also choose an approximate solution to the Riemann problem, [Kol14] for example uses the local Lax-Friedrichs flux [LeVo2, Section 12.5]. To keep the third-order accuracy of the reconstruction, a suitable third-order method for the time evolution must be chosen. In [Kol14] the total-variation-diminishing Runge-Kutta-scheme of third order from [GS96] is employed.

Boundary treatment

In order to keep the third order of the CWENO3 scheme also at boundaries, a different reconstruction is chosen in cells, whose boundary coincides with a boundary of the domain, that is, in cells I_1 and I_J from (2.26). The boundary condition itself, for example $\Psi(u(l, t))$ in Equation (2.27), is incorporated into the scheme by using it to compute a boundary state $u_{\frac{1}{2}}^-$, which is then used as the (in this case) left initial condition for a Riemann problem at the boundary. The reconstruction in I_1 is then done by considering new polynomials P_1, P_2, P_3 , that are chosen similarly to P_L, P_C, P_R above. For details see [NKS18]. An advantage of this boundary treatment is the absence of additional ghost cells to be prescribed, while high order convergence is still present.

2.3 OPTIMIZATION

Apart from simulating the evolution of gas and power networks we would also like to optimize them. To this end we employ the principle “first-discretize-then-optimize”, in which optimization is only applied after the dynamics of the networks have been discretized and thereby rendered finite-dimensional. Hence it is enough for our purpose to use finite-dimensional optimization, for which we will now provide some background.

Yet firstly, note that many letters traditionally used for certain quantities in optimization and control theory are also traditionally used in the field of hyperbolic balance laws. To simultaneously keep to the customs of each field but also provide some visual distinction, all optimization quantities are denoted in bold letters, namely we use

- \mathbf{u} for a vector of controls as opposed to the state u of balance laws and
- \mathbf{f} for an objective function as opposed to the flow f of balance laws.

A common finite-dimensional optimization problem is, to find a solution \mathbf{u} for

$$\begin{aligned} \mathbf{u} &= \operatorname{argmin} f(\mathbf{u}) \\ \text{s.t. } \mathbf{g}(\mathbf{u}) &= 0 \\ \mathbf{h}(\mathbf{u}) &\geq 0, \end{aligned} \tag{2.30}$$

where $\mathbf{u} \in \mathbb{R}^{n_u}$, $f : \mathbb{R}^{n_u} \rightarrow \mathbb{R}$, $\mathbf{g} : \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_g}$ and $\mathbf{h} : \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_h}$.

In the case where \mathbf{f}, \mathbf{g} and \mathbf{h} are smooth enough, it is possible to use their derivatives to search for a solution \mathbf{u} . To this end we first introduce some notation.

Definition 2.47 ([NW06, Definitions 12.1, 12.4]). Consider an optimization problem (2.30).

- A point \mathbf{u} in the domain of \mathbf{f} is called *feasible*, if $\mathbf{g}(\mathbf{u}) = 0$ and $\mathbf{h}(\mathbf{u}) \geq 0$.

- A feasible point $\hat{\mathbf{u}}$ is called a *local minimum* of (2.30) if there is a neighborhood \mathcal{U} of $\hat{\mathbf{u}}$, such that

$$\mathbf{f}(\hat{\mathbf{u}}) \leq \mathbf{f}(\mathbf{u})$$

holds for all feasible $\mathbf{u} \in \mathcal{U}$.

Additionally, for constrained optimization problems it is useful to define a Lagrange function, which will be used to state optimality conditions in Proposition 2.50.

Definition 2.48 (Lagrange function, [NWo6, Section 12.3]). Consider an optimization problem (2.30). The function $\mathcal{L} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_g} \times \mathbb{R}^{n_h}$, defined by

$$\mathcal{L}(\mathbf{u}, \lambda, \mu) = \mathbf{f}(\mathbf{u}) - \lambda^T \mathbf{g}(\mathbf{u}) - \mu^T \mathbf{h}(\mathbf{u}),$$

is called the *Lagrange function* of the problem.

We must also introduce some technical properties of the constraints which are needed for the optimality conditions.

Definition 2.49 (active set, LICQ, [NWo6, Definitions 12.1 and 12.4]).

- The *active set* $\mathcal{A}(\mathbf{u})$ of constraints at a feasible point \mathbf{u} is the collection of all indices of equations in \mathbf{g} and indices of all those inequalities \mathbf{h}_i in \mathbf{h} for which equality holds, $\mathbf{h}_i(\mathbf{u}) = 0$. If we gather all constraints in a single vector $\mathbf{c} = (\mathbf{g} \ \mathbf{h})^T$, we have

$$\begin{aligned} \mathcal{A}(\mathbf{u}) &\subset \{1, \dots, n_g + n_h\}, \\ \mathcal{A}(\mathbf{u}) &= \left\{ j \in \{1, \dots, n_g + n_h\} \mid \mathbf{c}_j(\mathbf{u}) = 0 \right\} \\ &= \{1, \dots, n_g\} \cup \{j + n_g \mid \mathbf{h}_j(\mathbf{u}) = 0\}. \end{aligned}$$

- A feasible point \mathbf{u} is said to fulfill *linear independence constraint qualifications*, in short *LICQ*, if the tuple of vectors

$$\left(\frac{\partial \mathbf{c}_i}{\partial \mathbf{u}} \right)_{i \in \mathcal{A}(\mathbf{u})}$$

is linearly independent.

With this in mind we can write down conditions that enable us to distinguish minima of (2.30).

Proposition 2.50 (first-order necessary optimality conditions, [NWo6, Theorem 12.1]). Consider an optimization problem (2.30) where \mathbf{f} , \mathbf{g} and \mathbf{h} are continuously differentiable. Let $\hat{\mathbf{u}}$ be a local minimum of the problem that fulfills LICQ. Then there are vectors $\hat{\lambda} \in \mathbb{R}^{n_g}$, $\hat{\mu} \in \mathbb{R}^{n_h}$, such that

$$\begin{aligned} \partial_{\mathbf{u}} \mathcal{L}(\hat{\mathbf{u}}, \hat{\lambda}, \hat{\mu}) &= 0, \\ \mathbf{g}(\hat{\mathbf{u}}) &= 0, \\ \mathbf{h}(\hat{\mathbf{u}}) &\geq 0, \\ \hat{\mu} &\geq 0, \\ \hat{\mu}^T \mathbf{h}(\hat{\mathbf{u}}) &= 0. \end{aligned} \tag{2.31}$$

The conditions (2.31) are also known as the *Karush–Kuhn–Tucker conditions* or *KKT conditions*.

We will not try to find optima of optimization problems ourselves but instead employ the optimization suite IPOPT [WBo6], which will search for points satisfying the KKT conditions for us.

To this end IPOPT (and other such software) needs to evaluate f, g, h as well as their derivatives f', g' and h' . We will go into more detail on finding these derivatives in Chapter 6.

2.4 SEMIGROUP THEORY

In Chapter 3 we will deal with the Telegrapher's equations, a set of linear hyperbolic balance laws. The linearity can be used to prove strong results on these systems. The main tool for this is the theory of linear operator semigroups, which is explained in detail in [EN01]. The idea of this theory is to view the linear hyperbolic balance law (where $A, F \in \mathbb{R}^{n \times n}$)

$$\begin{aligned} u_t + Au_x &= -Fu, \\ u(x, 0) &= u_0(x) \end{aligned} \tag{2.32}$$

as a ordinary linear differential equation on a Banach space \mathcal{X} ,

$$\begin{aligned} u_t &= \mathcal{A}u, \\ u(x, 0) &= u_0(x) \end{aligned}$$

for some linear operator $\mathcal{A} : D(\mathcal{A}) \rightarrow \mathcal{X}$, whose domain $D(\mathcal{A})$ is a dense subspace of \mathcal{X} . If \mathcal{X} is finite-dimensional, this equation is readily solved by

$$u(t) = \exp(\mathcal{A}t)u_0. \tag{2.33}$$

Yet, in the system (2.32) the operator \mathcal{A} is given by $\mathcal{A}u = -A\partial_x u - Fu$ and \mathcal{X} cannot be finite-dimensional. Still, for well-behaved operators \mathcal{A} the solution has many properties of the expression (2.33). An operator in this context is well-behaved, if it is the generator of a strongly continuous semigroup.

Still, let us start with the definition relevant for finite-dimensional \mathcal{X} . In that case all linear operators on \mathcal{X} are bounded.

Definition 2.51 (uniformly continuous semigroup, [EN01, Definition I.3.2]). A family $(T(t))_{t \in \mathbb{R}_0^+}$ of bounded linear operators on a Banach space \mathcal{X} is called a *uniformly continuous semigroup* on \mathcal{X} , if the map $t \mapsto T(t)$ is continuous in the operator norm and satisfies

$$\begin{aligned} T(s)T(t) &= T(s+t) \text{ for all } s, t \geq 0, \\ T(0) &= \text{Id}_{\mathcal{X}}. \end{aligned}$$

Expressed differently, a uniformly continuous semigroup on \mathcal{X} is a continuous (in the operator topology) monoid homomorphism $T : \mathbb{R}_0^+ \rightarrow \mathcal{B}(\mathcal{X}, \mathcal{X})$, where $\mathcal{B}(\mathcal{X}, \mathcal{X})$ is the set of all bounded linear operators from \mathcal{X} to itself. These semigroups can be uniquely characterized by the following proposition.

Proposition 2.52 ([EN01, Theorem I.3.7]). *Every uniformly continuous semigroup of bounded linear operators is of the form*

$$T(t) = \exp(\mathcal{A}t)$$

for some bounded linear operator \mathcal{A} . Therefore uniformly continuous semigroups are actually Fréchet-differentiable and there holds $\dot{T}(0) = \mathcal{A}$.

Here $\exp(\mathcal{A}t)$ is defined by its power series (which is well-defined, because the bounded linear operators on \mathcal{X} form a Banach algebra). Such an operator \mathcal{A} is called the *generator of the semigroup*.

Unfortunately the operator $\mathcal{A} = -A\partial_x - F$ is not bounded and it is therefore not clear how to define $\exp(\mathcal{A}t)$. Therefore we must work with more general semigroups.

Definition 2.53 (strongly continuous semigroup, [EN01, Definition I.5.1]). A family $(T(t))_{t \in \mathbb{R}_0^+}$ of bounded linear operators on a Banach space \mathcal{X} is called a *strongly continuous semigroup* on \mathcal{X} , if it satisfies

$$\begin{aligned} T(s)T(t) &= T(s+t) \text{ for all } s, t \geq 0, \\ T(0) &= \text{Id}_{\mathcal{X}}, \end{aligned}$$

and if, in addition, for all $x \in \mathcal{X}$ the map

$$\begin{aligned} \xi_x : \mathbb{R}_0^+ &\rightarrow \mathcal{X}, \\ t &\mapsto T(t)x \end{aligned}$$

is continuous (note that here the x in ξ_x is an index, not a derivative).

This can also be formulated differently: A strongly continuous semigroup on \mathcal{X} is a continuous (in the strong topology) monoid homomorphism $T : \mathbb{R}_0^+ \rightarrow \mathcal{B}(\mathcal{X}, \mathcal{X})$. Strongly continuous semigroups cannot be written in the form $t \mapsto \exp(\mathcal{A}t)$ for some bounded \mathcal{A} . Yet, there is still the notion of a generator.

Definition 2.54 (generator of a strongly continuous semigroup, [EN01, Definition II.1.2]). The operator $\mathcal{A} : D(\mathcal{A}) \subset \mathcal{X} \rightarrow \mathcal{X}$, given by

$$\mathcal{A}x = \frac{d}{dt} \xi_x(0) = \lim_{h \searrow 0} \frac{T(h)x - x}{h},$$

where the domain of \mathcal{A} is given by

$$D(\mathcal{A}) = \{ x \in \mathcal{X} \mid \xi_x : \mathbb{R}_0^+ \rightarrow \mathcal{X} \text{ is differentiable} \},$$

is called the *generator* of the semigroup.

A property of strongly continuous semigroups that is crucial for their application in solving systems of the form (2.32), is given in

Proposition 2.55 ([EN01, Lemma II.1.1]). *A function ξ_x given like above by $\xi_x(t) = T(t)x$ is (right) differentiable in $t = 0$, if and only if it is differentiable on \mathbb{R}_0^+ and its derivative is given by*

$$\frac{d}{dt} \xi_x(t) = T(t) \frac{d}{dt} \xi_x(0).$$

With this in mind we can now solve the system (2.32) if \mathcal{A} is the generator of a strongly continuous semigroup $T(t)$ and the initial condition u_0 is in the domain of \mathcal{A} , as then the solution is simply given as

$$u(x, t) = T(t)u_0(x).$$

Yet, as was the case for more general hyperbolic balance laws, when passing to an integral form of the equation, more general solutions are possible.

Definition 2.56 (mild solution, [ENo1, Definition II.6.3]). Let \mathcal{A} generate a strongly continuous subgroup. A function $u : \mathbb{R}_0^+ \rightarrow \mathcal{X}$ which fulfills $\int_0^t u(s) ds \in D(\mathcal{A})$ for all $t \geq 0$, as well as

$$u(t) - u_0 = \mathcal{A} \int_0^t u(s) ds,$$

is called a *mild solution* of (2.32).

It is actually the case that mild solutions of (2.32) are weak solutions in the sense of Definition 2.3, see [Bal77]. It can be shown that if \mathcal{A} generates a strongly continuous semigroup every initial condition admits a mild solution:

Proposition 2.57 ([ENo1, Proposition II.6.4]). Let \mathcal{A} generate a strongly continuous semigroup $T(t)$ on \mathcal{X} and let $u_0 \in \mathcal{X}$. Then $u : \mathbb{R}_0^+ \rightarrow \mathcal{X}$, $t \mapsto T(t)u_0$ is a mild solution of

$$\begin{aligned} \partial_t u &= \mathcal{A}u, \\ u(0) &= u_0 \end{aligned}$$

and this solution is a classical solution whenever $u_0 \in D(\mathcal{A})$.

As was the case for non-linear balance laws, we will need to introduce boundary conditions for our linear balance laws. Luckily these can be reinterpreted as initial conditions following [KMNo3]. We will use the theory of semigroups in Chapter 3.

THE RELATION OF TELEGRAPHER'S EQUATIONS AND THE POWER FLOW MODEL

In this chapter we will determine a valid model for the power network part of the combined gas-power network. We shall show that in so-called alternating current (AC) networks—which are almost all power networks in use in energy systems—the solution often used in engineering, namely the solution to the power flow equations, can be thought of as an exponentially stable solution of the Telegrapher's equations defined on the power network.

The analytical properties of the system of Telegrapher's equations are not new and have also been shown in e.g. [Nic16]. The benefit of our presentation is a more concrete formulation in terms of physical properties like voltage and current of the power system and a more elementary approach, proving certain properties directly, instead of referring to theory.

Additional research in this direction include exponential stability considerations [EK17] as well as open- and closed loop control problems. Open loop problems tackle questions of optimal inflow [GKL19; GPT19; GT18], while closed loop control is concerned with feedback control [Gug14] or boundary stabilization [GPR18]. Similar to the ideas presented in [GHS16], we study the stability of the Telegrapher's equations using the concept of Lyapunov functions [BC16]. As was done in [Nic16] we prove a stability result for the Telegrapher's equations on networks with linear coupling conditions. We then show how such a solution can be thought of as a solution to the power flow equations. Later on we examine certain numerical approximations and prove, that they mimic the Lyapunov stability as for example introduced in [BC16]. In contrast to the already existing literature [GH19; GHS16; GS17], where only first order schemes have been applied, we analyze a well-known numerical scheme of second order with regards to its stability. A numerical analysis shows that the scheme is Lyapunov stable and can be used to numerically study networks of linear hyperbolic balance laws as the simulation results indicate.

This chapter is organized as follows: In Section 3.1 we present the physical meaning of the Telegrapher's equation, define the network setting and prove theoretical properties concerning the Lyapunov stability. Section 3.2 is concerned with the derivation of the well-known power flow equations from the Telegrapher's equation. In Section 3.3, two closely related numerical schemes of second order are introduced to solve the Telegrapher's equation. The Lyapunov stability for the numerical approximation is also discussed in detail. The last Section 3.4 deals with numerical experiments to investigate the performance of the scheme.

3.1 SOLUTION STRUCTURE OF THE TELEGRAPHER'S EQUATIONS

The Telegrapher's equations are a system of linear hyperbolic balance laws in one space dimension as in Section 2.1 and especially Section 2.4. They model the time-dependent evolution of voltage and current on a transmission line and can be written as an initial-boundary value problem on a single line with $x \in [0, \ell]$ and $t \in [0, T]$ as:

$$\begin{aligned} u_t + Au_x + Fu &= 0, \\ u(0, x) &= u_0(x), \\ u^1(t, 0) &= v_0(t), \\ u^1(t, \ell) &= v_\ell(t), \end{aligned} \tag{3.1}$$

where $u : [0, T] \times [0, \ell] \rightarrow \mathbb{R}^2$ and

$$A = \begin{pmatrix} 0 & \frac{1}{C} \\ \frac{1}{L} & 0 \end{pmatrix}, \quad F = \begin{pmatrix} \frac{G}{C} & 0 \\ 0 & \frac{R}{L} \end{pmatrix}. \tag{3.2}$$

Here, $R, L, G, C > 0$ are constants depending on the line, $T > 0$ is some time horizon and ℓ is the length of the line. Therefore in the notation of Section 2.1, we have $\mathcal{T} = [0, T]$, $\mathcal{X} = [0, \ell]$ and $\mathcal{S} = \mathbb{R}^2$. The vector u is composed of the voltage $u^1 = v$ and the current $u^2 = i$. In the engineering literature (e.g. [MMo1, §1.2.2 Lossy Transmission Lines]) the Telegrapher's equations are usually written as

$$\begin{aligned} \frac{\partial v}{\partial x} &= -(Ri + L \frac{\partial i}{\partial t}), \\ \frac{\partial i}{\partial x} &= -(Gv + C \frac{\partial v}{\partial t}). \end{aligned}$$

Yet another form of the Telegrapher's equations is obtained by transformation into characteristic variables ξ^+, ξ^- :

$$\begin{aligned} \sqrt{\frac{C}{L}}v &= \xi^+ - \xi^-, \\ i &= \xi^+ + \xi^-, \end{aligned} \tag{3.3}$$

so that Equation (3.1) turns into

$$\xi_t + \Lambda \xi_x + B\xi = 0, \tag{3.4}$$

with

$$\Lambda = \frac{1}{\sqrt{LC}} \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \quad B = \begin{pmatrix} a & b \\ b & a \end{pmatrix} \tag{3.5}$$

and $a = \frac{1}{2} \left(\frac{R}{L} + \frac{G}{C} \right)$, $b = \frac{1}{2} \left(\frac{R}{L} - \frac{G}{C} \right)$. Note that the eigenvalues of B are $a + b = \frac{R}{L} > 0$ and $a - b = \frac{G}{C} > 0$ and that in addition $a > 0$ holds.

In a power network setting, these equations are defined on the lines and are coupled at the nodes by some conditions. In an AC power network all voltages and currents that serve as boundary conditions are sinusoidal with the same angular frequency ω , that is, they are of the form $p_1 \sin(\omega t) + p_2 \cos(\omega t)$ for real p_1, p_2 and $\omega > 0$.

We start off by examining a time-periodic solution of the Telegrapher's equations on a line with the same angular frequency ω as that

of the boundary conditions. Therefore we write v and i as a Fourier series in time,

$$\begin{aligned} v(x, t) &= \sum_{m \in \mathbb{Z}} V_m(x) e^{im\omega t}, \\ i(x, t) &= \sum_{m \in \mathbb{Z}} I_m(x) e^{im\omega t}, \end{aligned} \quad (3.6)$$

with complex-valued functions V_m, I_m , which we call complex voltage and complex current. For the voltage and current to be real-valued, we impose

$$V_m = V_{-m}^* \text{ and } I_m = I_{-m}^* \text{ for all } m \in \mathbb{Z},$$

and from now on consider only $m \geq 0$. Due to the linear independence of the function family $(z \mapsto e^{cz})_{c \in \mathbb{C}}$ the Telegrapher's equations decompose into a decoupled family of complex linear ordinary differential equations (ODEs), each of which governs only V_m and I_m via

$$\begin{aligned} \frac{\partial V_m}{\partial x} &= -(R + im\omega L) I_m, \\ \frac{\partial I_m}{\partial x} &= -(G + im\omega C) V_m. \end{aligned} \quad (3.7)$$

We now define

$$\begin{aligned} Y_m^0 &= \frac{\sqrt{G + im\omega C}}{\sqrt{R + im\omega L}}, \\ \gamma_m &= \sqrt{(R + im\omega L) \cdot (G + im\omega C)}, \end{aligned}$$

where we choose that branch of the square root that is positive on \mathbb{R}^+ and continuous on $\mathbb{C} \setminus \{r \in \mathbb{R} \mid r < 0\}$.

It is easily verified, that

$$\begin{aligned} V_m(x) &= \frac{1}{\sinh(\gamma_m \ell)} \left[V_m^\ell \sinh(\gamma_m x) + V_m^0 \sinh(\gamma_m(\ell - x)) \right], \\ I_m(x) &= -\frac{Y_m^0}{\sinh(\gamma_m \ell)} \left[V_m^\ell \cosh(\gamma_m x) - V_m^0 \cosh(\gamma_m(\ell - x)) \right] \end{aligned} \quad (3.8)$$

solve equations (3.7). Here V_m^0, V_m^ℓ are the complex voltages at either side of the line, which we choose as boundary conditions for now. It is also possible to prescribe the current at the end of a line as done in the network setting in the following Section 3.1. Note also that Y_m^0 and $\sinh(\gamma_m \ell)$ are never zero because $R, L, G, C, \ell > 0$. Plugging these into Equation (3.6) yields a solution to the Telegrapher's equations. This solution will be the basis for deriving the power flow equations in Section 3.2. Next, we will investigate the stability of the Telegrapher's equations on networks.

Stability of the Telegrapher's equations on networks

We want to model power networks, that is, generators and consumers (loads), that are connected by transmission lines. Hence we use a directed graph as a power network model. The setting is the same as in Section 2.1.6. Let therefore $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a finite directed graph with vertex set \mathcal{V} and edge set \mathcal{E} . An edge e consists of an interval

$[0, \ell_e]$ together with edge parameters R_e, L_e, G_e, C_e . As in Section 2.1.6 we assign a virtual point x_v to a vertex and for an edge s starting at v we identify $0 \in [0, \ell_s]$ with x_v . For an edge e ending in v we identify $\ell_s \in [0, \ell_s]$ with x_v . As a reminder, the reasoning behind this identification is to interpret the graph as a subset of some \mathbb{R}^n , where the start and endpoints of an edge are actually the points where the vertices are located.

Let further $\mathcal{V} = \mathcal{G} \cup \mathcal{L}$ where \mathcal{G} denotes those nodes with a generator and \mathcal{L} denotes nodes with a load so that every node is either a generator or a load. Also for $v \in \mathcal{V}$ let $\mathcal{E}_v \subset \mathcal{E}$ be the set of edges connected to v . In addition we define the function

$$s(e, v) = \begin{cases} 1 & \text{if } v \text{ is the end of } e, \\ -1 & \text{if } v \text{ is the start of } e, \end{cases} \quad (3.9)$$

which distinguishes start and end nodes of an edge.

Then we formulate the linear inhomogeneous initial-boundary-value problem (3.10) on the graph \mathcal{G} ,

$$\partial_t u_e = -A_e \partial_x u_e - F_e u_e \quad \forall e \in \mathcal{E}, \quad (3.10)$$

with matrices A_e and F_e of the structure of (3.2) and under the conditions

$$u_e(0, x) = (u_0)_e(x) \quad \forall e \in \mathcal{E}, \quad (3.11)$$

$$u_e^1(t, x_v) = u_f^1(t, x_v) \quad \forall g \in \mathcal{L}, \forall e, f \in \mathcal{E}_g, \quad (3.12)$$

$$u_e^1(t, x_g) = v_g(t) \quad \forall g \in \mathcal{G}, \forall e \in \mathcal{E}_g, \quad (3.13)$$

$$\sum_{e \in \mathcal{E}_l} s(e, l) u_e^2(t, x_l) = i_l \quad \forall l \in \mathcal{L}, \quad (3.14)$$

where (3.11) is the initial condition, (3.12) is a continuity condition for the voltage in the loads (which is automatically fulfilled in the generators due to the next condition (3.13)), (3.13) is a coupling condition fixing the voltage at generators, and (3.14) is a coupling condition fixing the current at loads. In order to show well-posedness, we use the theory of operator semigroups, shortly introduced in Section 2.4, yet for a thorough introduction see [ENo1] and especially [KMNo3]. A different way to prove the following (or rather a more general case) is found in [Nic16]. The main difference is our condition of $R_e, G_e > 0$ in every transmission line, which allows us to consider also graphs that are not trees, because positive R, G result in stability of the problem itself, while in [Nic16] stability is the result of a control action, which is not present here. Nevertheless, the results on well-posedness can be found already in [Nic16], where the proof invokes the Lumer-Phillips theorem [ENo1, Theorem 3.15] directly, while we instead show that conditions needed for the application of [KMNo3, Proposition 3.9] are fulfilled.

We introduce the following setting: As underlying Banach (and actually Hilbert) space we take

$$\mathcal{X} = \prod_{e \in \mathcal{E}} L^2([0, \ell_e])^2,$$

with the scalar product

$$\langle u, v \rangle = \frac{1}{2} \sum_{e \in \mathcal{E}} \int_0^{\ell_e} (C_e u_e^1 v_e^1 + L_e u_e^2 v_e^2) dx. \quad (3.15)$$

As linear operator we take

$$\begin{aligned} \mathcal{A} : D(\mathcal{A}) &\rightarrow \mathcal{X} \\ u &\mapsto (e \ni x \mapsto -A_e \partial_x u_e(x) - F_e u_e) \end{aligned}$$

and as its domain the space

$$D(\mathcal{A}) = \prod_{e \in \mathcal{E}} \mathcal{H}^1([0, \ell_e])^2. \quad (3.16)$$

In addition we choose the boundary operator

$$\begin{aligned} \mathcal{L} : D(\mathcal{A}) &\rightarrow \mathbb{R}^{2|\mathcal{E}|} \\ u &\mapsto (B_V, B_C, B_I), \end{aligned}$$

with

$$\begin{aligned} B_V &= (u_e^1(t, x_g))_{g \in \mathcal{G}, e \in \mathcal{E}_g}, \\ B_C &= \left((u_e^1(x_l) - u_{e+1}^1(x_l))_{e=1}^{|\mathcal{E}_l|-1} \right)_{l \in \mathcal{L}}, \\ B_I &= \left(\sum_{e \in \mathcal{E}_l} s(e, l) u_e^2(t, x_l) \right)_{l \in \mathcal{L}}, \end{aligned}$$

which is well-defined, because any $u \in D(\mathcal{A})$ projected to a single edge $[0, \ell_e]$ of the graph fulfills $u \in \mathcal{H}^1([0, \ell_e], \mathbb{R})^2$, which has a well-defined trace operator for evaluation at the boundaries of $[0, \ell_e]$. Note that \mathcal{L} is continuous on $D(\mathcal{A})$ as a concatenation of the (continuous) trace operator and a finite-dimensional linear mapping.

To prove well-posedness, we first verify the prerequisites of [KMNo3, Assumption 3.1] in

Lemma 3.1. *The operator \mathcal{A} satisfies*

(G1) *The restriction of \mathcal{A} to $\ker \mathcal{L}$, namely*

$$\mathcal{A}_0 = \mathcal{A}|_{\ker \mathcal{L}} : D(\mathcal{A}_0) = \ker \mathcal{L} \rightarrow \mathcal{X}$$

is densely defined, closed and has non-empty resolvent set.

(G2) *\mathcal{L} is surjective.*

(G3) *The combined operator*

$$\begin{pmatrix} \mathcal{A} \\ \mathcal{L} \end{pmatrix} : D(\mathcal{A}) \rightarrow \mathcal{X} \times \mathbb{R}^{2|\mathcal{E}|}$$

is closed.

Proof. For this proof we consider the basis introduced in (3.3), where \mathcal{A} is of the form

$$\mathcal{A} = -\lambda \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \partial_x - \begin{pmatrix} a & b \\ b & a \end{pmatrix}.$$

For (G1) we first note that the space with zero boundary conditions, i.e.,

$$\prod_{e \in \mathcal{E}} \mathcal{H}_0^1([0, \ell_e])^2,$$

is a subset of $D(\mathcal{A}) \cap \ker \mathcal{L}$, which is well-known to be a dense subspace of $\prod_{e \in \mathcal{E}} L^2([0, \ell_e])^2$. Then we note that \mathcal{A} is just the sum of the spatial derivative and a bounded linear operator. As \mathcal{L} is continuous on \mathcal{H}^1 , its kernel is a closed subspace therein. Together we find that \mathcal{A}_0 is closed, if the derivative operator $\partial_x : \mathcal{H}^1 \rightarrow L^2$ is closed, which is well-known, see e.g. [Kre07, Example 4.13-4]. For the non-empty resolvent set we show that 0 is in the resolvent set of \mathcal{A}_0 , that is \mathcal{A}_0 is injective and its left inverse \mathcal{R} is densely-defined and bounded. The domain of \mathcal{R} is just the range of \mathcal{A}_0 , so we show this range to be dense in \mathcal{X} . Therefore we show that the space of test functions $\mathcal{T} = \{f \in \prod_{e \in \mathcal{E}} C^\infty([0, \ell_e])^2 \mid \text{supp}(f) \subset [0, \ell_e]\}$ fulfills $\mathcal{T} \subset \text{range}(\mathcal{A}_0)$ as it is well-known that \mathcal{T} is dense in L^2 .

Let $g \in \mathcal{T}$. There holds $\mathcal{T} \subset \mathcal{H}_0^1$ and as the zero boundary conditions essentially decouple different edges, we only consider one edge, given by $[0, \ell]$. The Fourier transform of g (which is well-defined as $g \in L^2$) is given by

$$g(x) = \int_{-\infty}^{\infty} \begin{pmatrix} \hat{g}_1(k) \\ \hat{g}_2(k) \end{pmatrix} e^{ikx} dk.$$

The operator \mathcal{A}_0 evaluated on g is given by

$$g(x) = - \int_{-\infty}^{\infty} \begin{pmatrix} a + \lambda ki & b \\ b & a - \lambda ki \end{pmatrix} \begin{pmatrix} \hat{g}_1(k) \\ \hat{g}_2(k) \end{pmatrix} e^{ikx} dk.$$

Let us call the matrix $D(k)$ and immediately note that it is invertible, as $a^2 > b^2$. In addition this matrix is diagonalizable whenever $b^2 \neq \lambda^2 k^2$. We find for its eigenvalues

$$\begin{aligned} \mu_1 &= a - \sqrt{b^2 - \lambda^2 k^2} \\ \mu_2 &= a + \sqrt{b^2 - \lambda^2 k^2}. \end{aligned}$$

Note, that these are proper complex numbers for k large enough. Nevertheless we can estimate their absolute values as:

$$|\mu_1|, |\mu_2| \geq |a| - |b| > 0.$$

As the eigenvalues of the inverse matrix $D(k)^{-1}$ are the inverses of μ_1, μ_2 , the norm of the inverse satisfies

$$\|D(k)^{-1}\| \leq \frac{1}{|a| - |b|}.$$

Hence the function $k \mapsto D(k)^{-1}$ is uniformly bounded. This also holds true for the remaining case $b^2 = \lambda^2 k^2$, because the operator norm is continuous. With this in mind we see that choosing

$$h(x) = \int_{-\infty}^{\infty} D(k)^{-1} \begin{pmatrix} \hat{g}_1(k) \\ \hat{g}_2(k) \end{pmatrix} e^{ikx} dk,$$

is well-defined and yields $\mathcal{A}_0 h = g$. There holds $h \in L^2([0, \ell], \mathbb{R}^2)$ and $h \in C^\infty([0, \ell], \mathbb{R}^2)$. It remains to show that h is again supported in $[0, \ell]$. This is the case because of the Paley–Wiener theorem, see [Rud87, Theorem 19.3], which relates L^2 -functions supported on an interval and their Fourier transforms. Together we have $h \in \mathcal{H}_0^1([0, \ell])^2 \subset D(\mathcal{A}_0)$. Finally using $\|g\|_2 = \|\hat{g}\|_2$ and the norm estimate of D^{-1} , it is easily computed that \mathcal{R} is bounded. The injectivity of \mathcal{A}_0 will be shown in Lemma 3.2.

For (G2) we note that functions, that on every edge have the form

$$u(x) = \begin{pmatrix} a + bx \\ c + dx \end{pmatrix},$$

are in $D(\mathcal{A})$ and can be used to construct any value for \mathcal{L} .

For (G3) we note again that \mathcal{L} is continuous on $D(\mathcal{A})$ and that \mathcal{A} is closed, hence their combination is as well. \square

Lemma 3.2. \mathcal{A}_0 is injective.

Proof. Here we consider again the basis, where the components of u are the voltage $u^1 = v$ and the current $u^2 = i$. Therefore consider an element $u_0 = (v, i)^T \in \ker \mathcal{A}_0$. Such u_0 is a valid initial condition for the initial-boundary-value problem (3.10) with homogeneous coupling conditions and a corresponding solution given by $u(t, x) = u_0(x)$ for all $t \geq 0$. This solution is at least once differentiable with respect to time and space. Therefore we can apply the reasoning of the following Proposition 3.4 to this special solution and find

$$\begin{aligned} 0 &= \sum_{e \in \mathcal{E}} \int_0^{\ell_e} C_e v_e \dot{v}_e + L_e i_e \dot{i}_e = \dot{\mathcal{V}}(\zeta(t)) \\ &\leq -2 \min_{e \in \mathcal{E}} \left(\min \left(\frac{R_e}{L_e}, \frac{G_e}{C_e} \right) \right) \mathcal{V}(u_0). \end{aligned}$$

As $\mathcal{V}(u_0) = 0$ if and only if $u_0 = 0$, we find that \mathcal{A}_0 is injective. \square

Lemma 3.3. Problem (3.10) is well-posed and therefore admits a solution $u : \mathcal{T} \rightarrow D(\mathcal{A})$ for $u_0 \in D(\mathcal{A})$ and suitably smooth v_g, i_l .

Proof. Following [KMNo3, Proposition 3.9], it remains to show that \mathcal{A}_0 generates a strongly continuous semigroup. For this we first omit the matrices F_e in \mathcal{A}_0 and note that the resulting operator is skew-symmetric with respect to the scalar product (3.15) of \mathcal{X} . (The boundary terms arising in integration by parts vanish, because $D(\mathcal{A}_0) \subset \ker \mathcal{L}$). Then we cite [ENo1, 3.24 Theorem. (Stone, 1932)] to see that \mathcal{A}_0 generates a unitary group. We now add F_e again, which is

a bounded linear operator and use [ENo1, Proposition 1.12], showing that the full operator \mathcal{A}_0 also generates a strongly continuous semigroup. The last step of non-vanishing v_g, i_l follows from application of the variation of constants formula [ENo1, VI, Corollary 7.8], which introduces some smoothness conditions on v_g, i_l . The exact nature of these condition is of no consequence to us but could be found by working through [KMNo3] to incorporate the boundary conditions as initial conditions together with a source term. \square

With Lemma 3.3 we have a \mathcal{H}^1 -solution to the initial-boundary-value problem (3.10). We now show its exponential stability using the Lyapunov function

$$\mathcal{V}(u(t)) = \frac{1}{2} \sum_{e \in \mathcal{E}} \int_0^{\ell_e} C_e (u_e^1)^2 + L_e (u_e^2)^2.$$

Proposition 3.4. *Two solutions u, w of Problem (3.10) with identical coupling conditions v_g and i_l but possibly different initial conditions $u_0, w_0 \in \prod_{e \in \mathcal{E}} \mathcal{H}^1([0, \ell_e])^2$ grow closer exponentially in the L^2 -norm.*

Proof. We take the difference of the two solutions $\zeta = u - w = (v, i)^T$ with initial condition $\zeta_0 = u_0 - w_0$ and examine the derivative of the Lyapunov function. As before the coupling condition for ζ is then homogeneous.

$$\begin{aligned} \dot{\mathcal{V}}(\zeta(t)) &= \sum_{e \in \mathcal{E}} \int_0^{\ell_e} C_e v_e \dot{v}_e + L_e i_e \dot{i}_e \\ &= \sum_{v \in \mathcal{V}} \sum_{e \in \mathcal{E}_v} v_e(t, x_v) s(e, v) i_e(t, x_v) - \sum_{e \in \mathcal{E}} \int_0^{\ell_e} G_e v_e^2 + R_e i_e^2 \\ &= \sum_{v \in \mathcal{V}} v_v(t, x_v) \sum_{e \in \mathcal{E}_v} s(e, v) i_e(t, x_v) - \sum_{e \in \mathcal{E}} \int_0^{\ell_e} G_e v_e^2 + R_e i_e^2 \\ &= - \sum_{e \in \mathcal{E}} \int_0^{\ell_e} G_e v_e^2 + R_e i_e^2 \\ &\leq -2 \min_{e \in \mathcal{E}} \left(\min \left(\frac{R_e}{L_e}, \frac{G_e}{C_e} \right) \right) \mathcal{V}(\zeta(t)), \end{aligned}$$

where the third equality stems from (3.12) and the fourth from (3.13) or (3.14) depending on the node being either a load or a generator. Note that because of the homogeneous boundary conditions, for generators v there holds $v_v(t, x_v) = 0$ and for loads there holds $\sum s(e, v) i_e(t, x_v) = 0$. Gronwall's inequality is then employed to get

$$\dot{\mathcal{V}}(\zeta(t)) \leq \mathcal{V}(\zeta(0)) \exp \left(-2 \min_{e \in \mathcal{E}} \left(\min \left\{ \frac{R_e}{L_e}, \frac{G_e}{C_e} \right\} \right) t \right). \quad (3.17)$$

Lastly noting that \mathcal{V} is equivalent to the standard L^2 -scalar product, yields exponential L^2 -stability. \square

In practical terms Proposition 3.4 means that for every set of boundary conditions one can choose an arbitrary initial condition u_0 , compute its corresponding solution u and all solutions for different initial conditions but the same boundary conditions converge exponentially to u . In the following section we will examine suitable boundary conditions and choose corresponding initial conditions to construct solutions to the power flow equations as solutions of the Telegrapher's equations.

3.2 FROM TELEGRAPHER'S EQUATION TO POWER FLOW EQUATIONS

3.2.1 Periodic solutions to the Telegrapher's equations

Now that we have established that only the coupling conditions matter for the long-term behavior of a solution, we examine periodic coupling conditions and their relation to the power flow equations. This means the functions v_g and i_l in (3.13) and (3.14) shall be of the form

$$\begin{aligned} v_g(t) &= \sum_{m \in \mathbb{Z}} \hat{V}_m^g e^{im\omega t}, \\ i_l(t) &= \sum_{m \in \mathbb{Z}} \hat{I}_m^l e^{im\omega t}, \end{aligned}$$

where quantities with a hat $\hat{\cdot}$ are defined at a node. Again, to have real voltage and current, there must hold $\hat{V}_m^g = (\hat{V}_{-m}^g)^*$ and likewise for the current. Therefore we once again only consider $m \geq 0$.

A solution to Problem (3.10) with these coupling conditions, that is smooth on every edge, is given by a Fourier series with coefficients of the form (3.8), where the coupling conditions translate to linear conditions on the complex voltage and current V_m and I_m for each transmission line. The correct coefficients in this expression can be calculated by solving the linear system stemming from the coupling conditions. As the system is linear, this can be done for each Fourier mode $m\omega$ separately.

The Fourier mode of the current at the ends of a line can be expressed as

$$\begin{pmatrix} I_m(0) \\ I_m(\ell) \end{pmatrix} = -\frac{Y_m^0}{\sinh(\gamma_m \ell)} \begin{pmatrix} -\cosh(\gamma_m \ell) & 1 \\ -1 & \cosh(\gamma_m \ell) \end{pmatrix} \begin{pmatrix} V_m^0 \\ V_m^\ell \end{pmatrix}.$$

We see that the incoming end behaves differently from the outgoing end, because of an overall minus sign in the second component. This is natural in the PDE setting as the sign of the current marks the direction of flow. By using $I_m^0 = I_m(0)$ and $I_m^\ell = -I_m(\ell)$, we can transform the system into

$$\begin{pmatrix} I_m^0 \\ I_m^\ell \end{pmatrix} = -\frac{Y_m^0}{\sinh(\gamma_m \ell)} \begin{pmatrix} -\cosh(\gamma_m \ell) & 1 \\ 1 & -\cosh(\gamma_m \ell) \end{pmatrix} \begin{pmatrix} V_m^0 \\ V_m^\ell \end{pmatrix}.$$

This makes the transmission line invariant under switching orientation and positive currents mean currents leaving the node into the

transmission line. A further benefit is that we can omit the function s from Equation (3.9). With this we come back to the whole network.

By the coupling conditions on the voltage, i.e. (3.12) and (3.13), voltage Fourier modes of all lines that meet at a node are equal. Therefore it is natural to define V_m^r as the m -th voltage Fourier mode at vertex r that then serves as the voltage Fourier mode of all lines at this vertex.

This is different for the current. For a transmission line between node r and node s we define the current leaving node r in direction s as

$$I_m^{rs} = -\frac{Y_m^{0,rs}}{\sinh(\gamma_m^{rs}\ell^{rs})} (-V_m^r \cosh(\gamma_m^{rs}\ell^{rs}) + V_m^s),$$

where Y, γ and ℓ now depend on the transmission line.

The net current I_m^r at node r is then just the sum of all currents leaving that node,

$$I_m^r = \sum_{s \in \mathcal{E}} -\frac{Y_m^{0,rs}}{\sinh(\gamma_m^{rs}\ell^{rs})} (-V_m^r \cosh(\gamma_m^{rs}\ell^{rs}) + V_m^s) \quad \forall r \in \mathcal{V}, \quad (3.18)$$

where we set $Y_m^{0,rs} = 0$ if node r and s are not connected and also $Y_m^{0,rr} = 0$ for all r . These equations are then combined for the whole network, i.e.,

$$I_m = Y_m V_m, \quad (3.19)$$

where $I_m, V_m \in \mathbb{R}^{|\mathcal{V}|}$ are the vectors of net current and voltage at each node and Y_m is the so-called admittance matrix. One can set up admittance matrices for different power networks, where not every edge is a transmission line. Yet in our case, Y_m is given as

$$\begin{aligned} (Y_m)_{ii} &= \sum_{j=1, j \neq i}^{|\mathcal{V}|} \frac{Y_m^{0,ij}}{\tanh(\gamma_m^{ij}\ell^{ij})}, \\ (Y_m)_{ij} &= (Y_m)_{ji} = -\frac{Y_m^{0,ij}}{\sinh(\gamma_m^{ij}\ell^{ij})}. \end{aligned} \quad (3.20)$$

It is an interesting and open question, whether admittance matrices are always invertible, for invertibility criteria see e.g. [TM22]. We now prove a criterion that asserts invertibility, whenever $|\gamma_m^{ij}\ell^{ij}|$ is great enough for all index pairs (i, j) . In order to do this we need two lemmas.

Lemma 3.5. *The argument $\arg(Y) = \theta$ of $Y = \frac{\sqrt{a+bi}}{\sqrt{c+di}} = \text{Re}^{i\theta}$ with $a, b, c, d > 0$, fulfills $|\arg(Y)| < \frac{\pi}{4}$.*

Proof. We find

$$\begin{aligned} 0 &< \arg(a+bi), \arg(c+di) < \frac{\pi}{2} \\ \Rightarrow 0 &< \arg(\sqrt{a+bi}), \arg(\sqrt{c+di}) < \frac{\pi}{4} \\ \Rightarrow -\frac{\pi}{4} &< \arg(Y) < \frac{\pi}{4}, \end{aligned}$$

because $\arg(Y) = \arg(\sqrt{a+bi}) - \arg(\sqrt{c+di})$. □

Lemma 3.6.

1. If $\Re(z) > \frac{1}{2} \sinh^{-1}(2 + \sqrt{3})$, there holds

$$|\arg(\tanh(z))| < \frac{\pi}{12}.$$

2. If $\Re(z) > \frac{1}{2} \cosh^{-1}(9)$, there holds

$$\frac{1}{2} |(\cosh(z))| > 1,$$

Proof. Let $z = x + iy$. For the first assertion we compute

$$\begin{aligned} x &> \frac{1}{2} \sinh^{-1}(2 + \sqrt{3}) \\ \Rightarrow \sinh(2x) &> 2 + \sqrt{3} \\ \Rightarrow 2 - \sqrt{3} &> \frac{1}{\sinh(2x)}, \end{aligned} \tag{3.21}$$

where we first used that \sinh is monotone increasing on the positive reals and then that $(2 + \sqrt{3})(2 - \sqrt{3}) = 1$. Now we use

$$\tanh(x + iy) = \frac{\sinh(2x) + i \sin(2y)}{\cosh(2x) + \cos(2y)},$$

to conclude

$$\begin{aligned} \tan\left(\frac{\pi}{12}\right) &= 2 - \sqrt{3} > \frac{1}{\sinh(2x)} > \frac{\sin(2y)}{\sinh(2x)} \\ &= \tan\left(\frac{\Im(\tanh(x + iy))}{\Re(\tanh(x + iy))}\right) \\ &= \tan\left(\arg(\tanh(x + iy))\right). \end{aligned} \tag{3.22}$$

As \tan is both monotone increasing and symmetrical with respect to the origin, the claim follows.

The second assertion follows similarly from

$$|\cosh(x + iy)|^2 = \frac{1}{2} (\cosh(2x) + \cos(2y)),$$

for $x, y \in \mathbb{R}$ and $\cos(2y) > -1$. \square

With these lemmas we can prove the following proposition.

Proposition 3.7. Consider an admittance matrix Y , whose coefficients are given by Equation (3.20). Let $\Re(\gamma_m^{ij} \ell^{ij}) > \frac{1}{2} \cosh^{-1}(9)$ for all index pairs. Then Y is strictly diagonally dominant and hence invertible.

Proof. First note that $\frac{1}{2} \cosh^{-1}(9) > \frac{1}{2} \sinh^{-1}(2 + \sqrt{3})$, therefore all conditions in Lemma 3.6 are fulfilled.

Let $\alpha_m^{ij} = \gamma_m^{ij} \ell^{ij}$. For an index pair (i, j) we find, according to Lemma 3.6,

$$\begin{aligned} &\frac{1}{2} |(\cosh(\alpha_m^{ij}))| - 1 > 0 \\ \Rightarrow &\left| \frac{Y_m^{0,ij}}{\sinh(\alpha_m^{ij})} \right| \left(\frac{1}{2} |(\cosh(\alpha_m^{ij}))| - 1 \right) > 0 \\ \Rightarrow &\frac{1}{2} \left| \frac{Y_m^{0,ij}}{\tanh(\alpha_m^{ij})} \right| - \left| \frac{Y_m^{0,ij}}{\sinh(\alpha_m^{ij})} \right| > 0. \end{aligned}$$

Now we sum over j for an arbitrary i and use

- $\psi = \arg\left(\left|\frac{Y_m^{0,ij}}{\tanh(\alpha_m^{ij})}\right|\right) < \frac{\pi}{4} + \frac{\pi}{12} = \frac{\pi}{6}$,
- $\cos(\psi) \geq \cos\left(\frac{\pi}{6}\right) = \frac{1}{2}$ for $0 \leq \psi \leq \frac{\pi}{6}$ and
- $\Re(z) = \cos(\arg(z)) * |z|$ for $z \in \mathbb{C}$,

to conclude that

$$\begin{aligned}
 0 &< \sum_{j=1}^{|V|} \frac{1}{2} \left| \frac{Y_m^{0,ij}}{\tanh(\alpha_m^{ij})} \right| - \left| \frac{Y_m^{0,ij}}{\sinh(\alpha_m^{ij})} \right| \\
 &\leq \sum_{j=1}^{|V|} \Re\left(\frac{Y_m^{0,ij}}{\tanh(\alpha_m^{ij})}\right) - \left| \frac{Y_m^{0,ij}}{\sinh(\alpha_m^{ij})} \right| \\
 &= \Re\left(\sum_{j=1}^{|V|} \frac{Y_m^{0,ij}}{\tanh(\alpha_m^{ij})}\right) - \sum_{j=1}^{|V|} \left| \frac{Y_m^{0,ij}}{\sinh(\alpha_m^{ij})} \right| \\
 &\leq \left| \sum_{j=1}^{|V|} \frac{Y_m^{0,ij}}{\tanh(\alpha_m^{ij})} \right| - \sum_{j=1}^{|V|} \left| \frac{Y_m^{0,ij}}{\sinh(\alpha_m^{ij})} \right|,
 \end{aligned}$$

which shows the strict diagonal dominance of Y . \square

With Proposition 3.7 we can guarantee invertibility of an admittance matrix. Unfortunately there is one admittance matrix Y_m for each $m \in \mathbb{N}$, defined by Equation (3.20). Luckily, for slightly stronger conditions we can prove the invertibility of Y_m for every $m \in \mathbb{N}$. To this end we again state a lemma.

Lemma 3.8. *For two complex numbers z_1, z_2 , each with positive real part, there holds $\Re(\sqrt{z_1 z_2}) \geq \sqrt{\Re(z_1)\Re(z_2)}$, that is, the real part of the geometric mean of two numbers with positive real part is at least as big as the geometric mean of the real parts.*

Proof. Consider $z = x + y i$ with $x > 0$. The real part of \sqrt{z} fulfills

$$\Re(\sqrt{z}) = \sqrt{\frac{x + \sqrt{x^2 + y^2}}{2}}, \quad (3.23)$$

as a short calculation shows. Now consider two numbers $z_1 = (a + b i)$, $z_2 = (c + d i)$, with $a, c > 0$.

$$\begin{aligned}
 \Re(\sqrt{z_1 z_2}) &= \sqrt{\frac{(ac - bd) + \sqrt{(ac - bd)^2 + (ad + bc)^2}}{2}} \\
 &= \sqrt{\frac{(ac - bd) + \sqrt{(ac + bd)^2 + (ad - bc)^2}}{2}} \\
 &\geq \sqrt{\frac{(ac - bd) + \sqrt{(ac + bd)^2}}{2}} \\
 &= \sqrt{\frac{(ac - bd) + (ac + bd)}{2}} \\
 &= \sqrt{ac} \\
 &= \sqrt{\Re(z_1)\Re(z_2)},
 \end{aligned} \quad (3.24)$$

where the sign switch under the square root is done by expanding the squares and using $2(ac)(bd) = 2(ad)(bc)$. \square

Proposition 3.9. *Let for every transmission line in the network hold that*

$$l^{ij} \sqrt{R^{ij} G^{ij}} > \frac{1}{2} \cosh^{-1}(9). \quad (3.25)$$

Then each admittance matrix Y_m arising from this network, given by Equation (3.20), is strictly diagonally dominant.

Proof. The conditions in Proposition 3.7 are lower bounds on the real part of $\gamma_m^{ij} l^{ij}$, which can be written as the geometric mean of $l^{ij} (R^{ij} + i m \omega L^{ij})$ and $l^{ij} (G^{ij} + i m \omega C^{ij})$. The claim then immediately follows from the preceding lemma. \square

Prescribing the left-hand sides I_m in Equation (3.19) for all $m \in \mathbb{Z}$ amounts to prescribing a problem of type (3.10) with only load nodes. For such a network a series $(V_m)_{m \in \mathbb{Z}}$ that fulfills Equation (3.19) for all Fourier modes m yields a solution to (3.10) given on every line by (3.8). This means, that there is a time periodic solution to (3.10), if all Y_m are invertible and all nodes are loads.

Of course, usually there are generator nodes. In that case, (3.19) can be reduced by inserting the values of \hat{V}_m^g and shifting them to the other side by altering I_m . Concretely, we define the vector \tilde{V}_m by

$$\tilde{V}_m^i = \begin{cases} V_m^i & \text{if the } i\text{-th node is a generator,} \\ 0 & \text{else,} \end{cases}$$

and

$$\bar{I}_m = I_m - Y_m \tilde{V}_m.$$

Then, we strike all rows (and in addition columns for Y_m) from \bar{I}_m and Y_m that correspond to generator nodes and arrive at a new equation

$$\tilde{I}_m = \tilde{Y}_m \tilde{V}_m.$$

For this equation to be solvable for all possible choices of generator nodes, we must demand that all principal minors¹ of Y_m are invertible. This is the case, when Y is strictly diagonally dominant, so for example under the conditions of Proposition 3.7. Therefore, if for every $m \in \mathbb{Z}$, Y_m , along with all its principal minors, is invertible, the coupling conditions,

$$\begin{aligned} \hat{V}_m^g &= V_m^g \quad \forall g \in \mathcal{G}, \\ \hat{I}_m^l &= \sum_{s \in \mathcal{L}} -\frac{Y_m^{0,ls}}{\sinh(\gamma_m^{ls} \ell)} \left(-V_m^l \cosh(\gamma_m^{ls} \ell) + V_m^s \right) \quad \forall l \in \mathcal{L}, \end{aligned} \quad (3.26)$$

yield a time-periodic solution to Problem (3.10), defined on every edge by Equations (3.6) through (3.8). Note that these conditions are again linear in the complex voltages.

¹ A principal minor of a square matrix M is every matrix that results from M by removing a number of rows and all columns with matching indices to the removed rows.

3.2.2 Power coupling conditions

In power networks we usually want to satisfy a power demand instead of a current demand.

Electric power at a node v is given by the product of voltage and the net current,

$$\begin{aligned} p_v(t) &= v_v(t) \cdot i_v(t) \\ &= \sum_{m,k \in \mathbb{Z}} V_m I_k e^{j(m+k)\omega t} \\ &= \sum_{m,k \in \mathbb{Z}} V_{m-k} I_k e^{jm\omega t} \\ &=: \sum_{m \in \mathbb{Z}} P_m e^{jm\omega t}. \end{aligned}$$

One way to replace the linear coupling conditions (3.26) would be to prescribe \hat{p}_m^l at every load node. In the case where v_g and p_l are Fourier polynomials, that is, there is $M \in \mathbb{N}_0$, such that $\hat{p}_m^l, V_m^g = 0$ for $|m| > M$, this yields $|\mathcal{V}|(2M+1)$ equations for just as many unknown voltage Fourier modes². So, depending on the concrete values of \hat{P}, \hat{V} , the system might be solvable. Unfortunately, except for small examples, it is hard to decide a priori whether a given set of power coupling conditions is feasible, see for example [Jer+20].

The mixing of different Fourier modes can make the resulting coupling conditions quite cumbersome. Luckily in applications a strong simplification is done: It is assumed that all inputs are sinusoidal, meaning that the only non-vanishing Fourier modes are those for $m \in \{\pm 1\}$.

Then, following any standard reference on power systems, for example [And15; GSC16], we use $v, i \in \mathbb{R}$ to express $V_{-1} = V_1^*$ and define $V := |V| e^{i\phi}$, similarly $I := |I| e^{i\psi}$, and find for the power

$$\begin{aligned} p(t) &= (Ve^{i\omega t} + V^*e^{-i\omega t}) (Ie^{i\omega t} + I^*e^{-i\omega t}) \\ &= VI^* + V^*I + VIe^{2i\omega t} + V^*I^*e^{-2i\omega t} \\ &= \Re(VI^*) + \Re(VIe^{2i\omega t}) \\ &= \Re(VI^*) (1 + \cos(2(\omega t + \phi))) + \Im(VI^*) \sin(2(\omega t + \phi)). \end{aligned}$$

This suggests to define the so-called *complex apparent power* $S = VI^*$, its real part, the *real power* $P = \Re(S)$ and its imaginary part, the reactive power $Q = \Im(S)$, so that at the node v it holds:

$$p_v(t) = P_v (1 + \cos(2(\omega t + \phi))) + Q_v \sin(2(\omega t + \phi)).$$

To define the power flow equations we use the admittance matrix Y from (3.18):

$$I^* = (YV)^*,$$

of which we write the real and imaginary part as

$$Y = G + iB.$$

² Up to now these are complex equations for complex Fourier modes but assuming $V_m = V_{-m}^*$ yields $|\mathcal{V}|(2M+1)$ real equations for as many real unknowns.

For the complex apparent power this means

$$S = \text{diag}(V)Y^*V^*.$$

For every node k we then find the nonlinear system of equations

$$\begin{aligned} P_k &= \sum_{i=1}^N |V_k||V_i| (G_{ki} \cos(\phi_k - \phi_i) + B_{ki} \sin(\phi_k - \phi_i)), \\ Q_k &= \sum_{i=1}^N |V_k||V_i| (G_{ki} \sin(\phi_k - \phi_i) - B_{ki} \cos(\phi_k - \phi_i)). \end{aligned} \quad (3.27)$$

These are combined to a system of equations for the whole network, called the *power flow equations*. Up to now we have 4 unknown quantities at every node. To find a unique solution, at every node two quantities are prescribed:

- All load nodes prescribe P and Q .
- All generators (with one exception) prescribe $|V|$ and P .
- The last generator, or even a number of generators, known as *slack nodes*, prescribe $|V|$ and ϕ .

These conditions lead to a system of nonlinear equations with two equations for each node to determine the two nodal variables $|V|$ and ϕ . For more information, we refer to [And15; FR16].

Any solution to the system of power flow equations yields a solution to the initial-boundary-value problem (3.10) with sinusoidal boundary conditions (with equal period at all nodes). Concretely a solution to the power flow system yields the complex voltages V at every node. With these one can compute $I = YV$ and use both V and I to prescribe sinusoidal boundary conditions at every node and compute the corresponding solution to (3.10). We prescribe the net current at load nodes and the voltage at generators (including the slack node). The justification for this is, that “voltage is pushed” (set by the voltage source) and “current is pulled” (drawn by the load). All in all we have shown how a solution of the power flow equations can be thought of as a solution to the Telegrapher’s equations with sinusoidal coupling conditions.

This solution necessarily has initial condition equal to Equation (3.6) evaluated at $t = 0$ but the results of Section 3.1 mean that any different initial condition will converge exponentially to the solution stemming from the power flow equations.

In the following sections we will examine the behavior of solutions numerically.

3.3 NUMERICAL SCHEME

We aim to mimic the behavior of the analytical solution of Problem (3.10) by a numerical approximation. We will first treat a single line, define the numerical scheme there and then prescribe numerical boundary conditions.

3.3.1 The scheme in the interior of a line

We compute a discrete approximation of a solution of (3.1) via a splitting scheme. The numerical approximation is based on Equation (3.4), where

$$\Lambda = \text{diag}(\lambda, -\lambda)$$

is diagonal. The balance law

$$\xi_t + \Lambda \xi_x = -B\xi$$

is split into an ODE part

$$\xi_t = -B\xi \quad (3.28)$$

and a PDE part

$$\xi_t + \Lambda \xi_x = 0. \quad (3.29)$$

We choose the discretization $t_n = n\Delta t$, $n = 0 \dots N$, $x_j = j\Delta x$, $j = 0 \dots J$ and

$$(\xi^\pm)_j^n = \xi^\pm(t_n, x_j).$$

Further, we define the lattice constant and the Courant number (see Definition 2.39) as

$$\Gamma = \frac{\Delta t}{\Delta x}, \quad \mathcal{C} = \Gamma|\lambda|.$$

As splitting we use Strang splitting (see Section 2.2.2 and [LeVo2, 17.4 Strang splitting]), meaning we make a half-step of the ODE, Equation (3.28), a full step of the PDE, Equation (3.29) and another half-step of the ODE.

Equation (3.28) is a linear ODE with constant coefficients, so we can solve it exactly up to machine precision with only constant computational costs by setting the ODE scheme to

$$\begin{aligned} \xi_j^* &= (\text{ODE}(\Delta t) \cdot \xi)_j = M \xi_j^n, \\ M &= \exp(-B\Delta t). \end{aligned} \quad (3.30)$$

Of course, this ODE only contains the source term of the Telegrapher's equations, which is why we employ the splitting technique. As Strang splitting itself introduces an error of second order, we cannot hope to get higher accuracy and therefore choose a scheme for the conservation law (3.29) which is at most second order, namely a flux-limited Lax-Wendroff scheme (see for example [LeVo2, 6.12 TVD limiters]), that for a single component is given for right and left moving waves as:

$$\begin{aligned} \hat{\xi}_j^+ &= \xi_j^+ - \Gamma|\alpha| (\xi_j^+ - \xi_{j-1}^+) \\ &\quad - \frac{1}{2}|\alpha| \Gamma(1 - |\alpha| \Gamma) \left(\phi(\theta_{j+\frac{1}{2}}^+)(\xi_{j+1}^+ - \xi_j^+) - \phi(\theta_{j-\frac{1}{2}}^+)(\xi_j^+ - \xi_{j-1}^+) \right), \\ \hat{\xi}_j^- &= \xi_j^- + \Gamma|\alpha| (\xi_{j+1}^- - \xi_j^-) \\ &\quad - \frac{1}{2}|\alpha| \Gamma(1 - |\alpha| \Gamma) \left(\phi(\theta_{j+\frac{1}{2}}^-)(\xi_{j+1}^- - \xi_j^-) - \phi(\theta_{j-\frac{1}{2}}^-)(\xi_j^- - \xi_{j-1}^-) \right), \end{aligned} \quad (3.31)$$

where a is the wave speed and we define as usual

$$\theta_{j+\frac{1}{2}}^+ = \begin{cases} \frac{\xi_j^+ - \xi_{j+1}^+}{\xi_{j+1}^+ - \xi_j^+} & \text{for } \xi_j \neq \xi_{j+1} \\ 0 & \text{for } \xi_j = \xi_{j+1} \end{cases}$$

and

$$\theta_{j+\frac{1}{2}}^- = \begin{cases} \frac{\xi_{j+2}^- - \xi_{j+1}^-}{\xi_{j+1}^- - \xi_j^-} & \text{for } \xi_{j+1}^- \neq \xi_j^- \\ 0 & \text{for } \xi_{j+1}^- = \xi_j^- \end{cases} \quad (3.32)$$

As limiter we choose the minmod-limiter,

$$\phi(x) = \begin{cases} 0 & \text{for } x < 0, \\ x & \text{for } 0 < x < 1, \\ 1 & \text{for } 1 < x, \end{cases} \quad (3.33)$$

which will become important for the numerical Lyapunov stability. Of course the choice of 0 in Equations (3.32) and (3.33) is arbitrary as the result of ϕ is anyways multiplied by the denominator of $\theta_{j+\frac{1}{2}}^-$ or $\theta_{j+\frac{1}{2}}^+$, making the product expression vanish, whenever the respective denominator is zero.

The CFL condition for this scheme is the same as that of the pure Lax-Wendroff scheme (Example 2.41), namely

$$\mathcal{C} \leq 1. \quad (3.34)$$

We write the scheme (although it is not linear) as

$$\hat{\xi}_j = (\text{PDE}(\Delta t)\xi)_j,$$

so that the full split scheme can be written as

$$\xi^{n+1} = \text{ODE}\left(\frac{\Delta t}{2}\right) \circ \text{PDE}(\Delta t) \circ \text{ODE}\left(\frac{\Delta t}{2}\right) \cdot \xi^n.$$

Having defined the scheme we come to the two main properties of it. We will now prove that the scheme is

- total-variation-diminishing and
- Lyapunov stable.

As numerical Lyapunov function we choose

$$\mathcal{V}(\xi) = \frac{1}{2} \sum_j \left((\xi_j^+)^2 + (\xi_j^-)^2 \right) \Delta x = \frac{1}{2} \sum_j (\xi_j)^T \xi_j \Delta x, \quad (3.35)$$

which is equivalent to the squared discrete L^2 -norm of the solution and hence corresponds directly to the analytical Lyapunov function.

As definition of total variation in two dimensions we choose the sum of the component-wise total variations:

$$\text{TV}(\xi) = \sum_j \left(\left| \xi_{j+1}^+ - \xi_j^+ \right| + \left| \xi_{j+1}^- - \xi_j^- \right| \right) = \sum_j \|\xi_{j+1} - \xi_j\|_1. \quad (3.36)$$

Now we are able to discuss the following results:

Proposition 3.10. *The ODE-scheme (3.30) is total-variation-diminishing and Lyapunov stable. The total variation after one step of the ODE scheme fulfills*

$$\text{TV}(\xi^{n+1}) \leq \exp(-b_{\min}\Delta t) \text{TV}(\xi^n) < \text{TV}(\xi^n)$$

where $b_{\min} = \min(b_1, b_2) = \min(\frac{R}{L}, \frac{G}{C})$ is the eigenvalue of smallest absolute value of the matrix B .

For the proof we need a short lemma.

Lemma 3.11. *There holds*

1. *The matrix $M = \exp(-B\Delta t)$ has the structure*

$$M = \begin{pmatrix} r & s \\ s & r \end{pmatrix}. \quad (3.37)$$

2. *Its eigenvalues are given by $m_1 = r + s$, $m_2 = r - s$ and fulfill $0 < m_1, m_2 < 1$.*

3. *There holds $r > 0$.*

Proof. Consider the base change matrix

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

Note that a matrix A fulfills

$$P^{-1}AP = A,$$

if and only if A has the form (3.37). Therefore there holds $P^{-1}BP = B$. But \exp commutes with change of basis and therefore

$$P^{-1}MP = \exp(-P^{-1}BP\Delta t) = \exp(-B\Delta t) = M,$$

proving 1. The eigenvalues of M can be computed from (3.37) and are $m_{1,2} = r \pm s$. So they are also given by $m_{1,2} = \exp(-b_{1,2}\Delta t)$ where $b_{1,2}$ are the eigenvalues of B . Assertion 2 follows because $b_{1,2} > 0$. Lastly $r = \frac{m_1 + m_2}{2} > 0$. \square

Next we come to the

Proof of Proposition 3.10. For the total variation after an ODE step we find

$$\begin{aligned}
\text{TV}(\text{ODE}(\Delta t)\xi) &= \sum_j \|M\xi_{j+1} - M\xi_j\|_1 \\
&\leq \|M\|_1 \sum_j \|\xi_{j+1} - \xi_j\|_1 \\
&= (|r| + |s|) \sum_j \|\xi_{j+1} - \xi_j\|_1 \\
&\leq \max(r + s, r - s) \sum_j \|\xi_{j+1} - \xi_j\|_1 \\
&= \max(m_1, m_2) \sum_j \|\xi_{j+1} - \xi_j\|_1 \\
&= \exp(-\min(b_1, b_2)\Delta t) \sum_j \|\xi_{j+1} - \xi_j\|_1 \\
&< \sum_j \|\xi_{j+1} - \xi_j\|_1 = \text{TV}(\xi),
\end{aligned}$$

due to Lemma 3.11, specifically, the second inequality is due to assertion 3 and the last one due to assertion 2.

For the Lyapunov function after an ODE step we get

$$\begin{aligned}
\mathcal{V}(\text{ODE } \xi) &= \frac{1}{2} \sum_j (M\xi_j)^T M\xi_j \Delta x \\
&= \frac{1}{2} \sum_j (\xi_j)^T M^T M \xi_j \Delta x \\
&\leq \frac{1}{2} \max(m_1^2, m_2^2) \sum_j (\xi_j)^T \xi_j \Delta x \\
&\leq \frac{1}{2} \sum_j (\xi_j)^T \xi_j \Delta x = \mathcal{V}(\xi),
\end{aligned}$$

again according to Lemma 3.11. \square

On the PDE side we find a similar result.

Proposition 3.12. *The PDE-scheme is total-variation-diminishing and Lyapunov stable.*

Proof. For the claim of TVD we note that the total variation (3.36) is given as the sum of the variations in both components. As the PDE-scheme does not mix the components, it is overall total variation diminishing, if it is TVD in each of its components. The TVD property of the minmod-limiter for the scalar case is a standard result found e.g. in [LeVo2, 6.12 TVD Limiters].

The Lyapunov function (3.35) is also given as a sum of the components, namely the sum of the L^2 -norms squared. Therefore we examine L^2 -stability of the PDE-scheme for the linear advection equation (which is one component of (3.29)), given by

$$u_t + au_x = 0. \quad (3.38)$$

with the wave speed a . The pair

$$E(u) = \frac{1}{2}u^2, \quad Q(u) = \frac{a}{2}u^2$$

is an entropy-entropy-flux pair for (3.38). According to [Tad87, Theorem 4.1] there is an entropy-conserving numerical flux for this entropy (see Definitions 2.35 and 2.43), which is defined by (see [FMT12, 2.1.2 Linear symmetrizable systems])

$$F_{j+\frac{1}{2}}^E = \frac{a}{2}(u_j + u_{j+1})$$

and numerical viscosity

$$Q_{j+\frac{1}{2}}^E = 0.$$

Note, that the corresponding scheme is the so-called naive scheme, which is unstable, although this poses no problem in the proof of [Tad87, Theorem 4.1].

Now, according to [Tad87, Theorem 5.2] a scheme for the advection equation with numerical viscosity $Q_{j+\frac{1}{2}}$ is entropy-stable with respect to E if

$$Q_{j+\frac{1}{2}} \geq Q_{j+\frac{1}{2}}^E = 0$$

for all j . The viscosity of the minmod scheme is given by

$$Q_{j+\frac{1}{2}}^{mm} = \mathcal{C} \left(1 - \left[(1 - \mathcal{C}) \phi(\theta_{j+\frac{1}{2}}^{\text{sign}(a)}) \right] \right) \geq 0,$$

because of the CFL condition (3.34) and $\phi \leq 1$. As the entropy E coincides with our Lyapunov function, this shows that the minmod scheme is Lyapunov stable on the advection equation and therefore our PDE scheme is Lyapunov stable for the PDE part (Equation (3.29)) of our splitting. \square

Before we come to boundary conditions we will remark on another PDE scheme that can be used, namely the pure (unlimited) Lax-Wendroff scheme introduced in Example 2.41. Unfortunately it is not TVD, yet it is Lyapunov stable, as its numerical viscosity fulfills

$$Q_{j+\frac{1}{2}}^{LW} = \mathcal{C}^2 \geq 0,$$

where again \mathcal{C} is the Courant number. Although it is not TVD, the total variation production is small. There holds

Proposition 3.13 (TV production of the LW scheme). *The total variation production of the Lax-Wendroff scheme fulfills*

$$\text{TV}(u^{n+1}) \leq (1 - \mathcal{C}^2 + \mathcal{C}) \text{TV}(u^n).$$

Proof. This can be easily seen by using the triangle inequality on $\text{TV}(u^{n+1})$, where u^{n+1} is replaced by the scheme $H(u^n)$, and then regrouping terms. \square

Obviously we cannot depend on the Lax-Wendroff scheme alone to be TVD. But the expression we just computed allows us to choose a Courant number that makes the total variation production of the PDE small enough to be canceled by the total variation reduction of the ODE scheme. We know that

$$\text{TV}(\xi^{n+1}) \leq (1 - \mathcal{C}^2 + \mathcal{C}) \exp(-b_{\min} \Delta t) \text{TV}(\xi^n).$$

Therefore the TVD property holds, if the scheme fulfills

$$(1 - \mathcal{C}^2 + \mathcal{C}) \exp(-b_{\min} \Delta t) \leq 1,$$

which is equivalent to

$$(1 - \mathcal{C}^2 + \mathcal{C}) \leq \exp(b_{\min} \Delta t). \quad (3.39)$$

This is true, if $\exp(b_{\min} \Delta t) \geq \max(1 - \mathcal{C}^2 + \mathcal{C}) = \frac{5}{4}$ but can also be satisfied, by choosing \mathcal{C} near 1. Because \mathcal{C} depends on Δt , this choice bounds Δt from below.

The advantage of using the ODE scheme to counteract the variation production of the PDE scheme is, that the combined scheme is actually of second order everywhere. If we demand the PDE scheme to be TVD on its own, we have to accept an order reduction at extremal points of the solution. Examples of this can be found in Section 3.4. Unfortunately the bounds we have just computed are much too conservative. This is so, because using the triangle equality in Proposition 3.13 and the minimum of the eigenvalues in the ODE scheme is overly pessimistic. And indeed it turns out that one can violate these bounds rather strongly without actually endangering the stability of the method. An example of this is given in Section 3.4.

3.3.2 Numerical coupling conditions

Now that we have established the stability of the scheme in the interior of the transmission line, we need to define the coupling conditions. Therefore we consider a node n with a set \mathcal{E}_n of transmission lines, that all start at it and reformulate the combined Telegrapher's equations on all these lines, as introduced in Section 2.1.6. We define the vector and the matrices

$$\begin{aligned} \xi &= \left(\xi_1^+ \quad \dots \quad \xi_N^+ \quad \xi_1^- \quad \dots \quad \xi_N^- \right)^T, \\ \Lambda &= \text{diag}(\lambda_1, \dots, \lambda_N, -\lambda_1, \dots, -\lambda_N), \\ B &= \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}, \\ \alpha &= \text{diag}(a_1, \dots, a_N), \\ \beta &= \text{diag}(b_1, \dots, b_N) \end{aligned}$$

and also

$$\xi^\pm = \left(\xi_1^\pm \quad \dots \quad \xi_N^\pm \right)^T, \quad \Lambda^\pm = \pm \text{diag}(\lambda_1, \dots, \lambda_N),$$

which leads to the PDE

$$\xi_t + \Lambda \xi_x + B\xi = \begin{pmatrix} \xi^+ \\ \xi^- \end{pmatrix}_t + \begin{pmatrix} \Lambda^+ & \\ & \Lambda^- \end{pmatrix} \begin{pmatrix} \xi^+ \\ \xi^- \end{pmatrix}_x + \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} \xi^+ \\ \xi^- \end{pmatrix} = 0.$$

For line $e \in \mathcal{E}_n$ we define $c_e = \sqrt{\frac{L_e}{C_e}}$. If the node is a load with prescribed current i_n , we have the following conditions:

$$\sum_{i=1}^N i_e = i_n \Leftrightarrow \sum_{e=1}^N \xi_e^+ = - \sum_{e=1}^N \xi_e^- + i_n,$$

$$v_e = v_f \quad \forall 1 \leq e, f \leq N \Leftrightarrow c_e(\xi_e^+ - \xi_e^-) = c_f(\xi_f^+ - \xi_f^-).$$

To reformulate them we define the matrices

$$\mathcal{M} = \begin{pmatrix} 1 & & \dots & 1 \\ c_1 & -c_2 & & \\ & \ddots & \ddots & \\ & & c_{N-1} & -c_N \end{pmatrix}, S = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}, \tilde{i}_n = \begin{pmatrix} i_n \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

which yields the conditions³ for the ghost cell values ξ_0 :

$$\begin{aligned} \mathcal{M}\xi_0^+ &= S\mathcal{M}\xi_0^- + \tilde{i}_n \\ \Leftrightarrow \xi_0^+ &= \mathcal{M}^{-1}S\mathcal{M}\xi_0^- + \mathcal{M}^{-1}\tilde{i}_n \\ \Leftrightarrow \xi_0^+ &= U\xi_0^- + \mathcal{M}^{-1}\tilde{i}_n. \end{aligned} \tag{3.40}$$

Note that $\|U\|_2 = 1$, so for $i_n = 0$ no energy is injected at the node. We still must choose ξ_0^- and do so by linear extrapolation, namely we choose $\xi_0^- = 2\xi_1^- - \xi_2^-$. In principle one should use stable extrapolation (see [BMZ15]) to guard against non-smooth solutions but this seems to be unnecessary with high enough R and G .

For the minmod scheme we need a second ghost cell ξ_{-1} . In principle it should be possible to use the time derivative of the coupling condition as in [BHH16] or [BK14], but in our setting it is much easier to use another linear extrapolation through the neighboring inner cell values and the newly-computed boundary value from (3.40).

Contrary, the generator coupling conditions are easier to determine as they essentially decouple different lines. When we prescribe the voltage at a node, this acts as a regular boundary condition for each attached line and can be treated as such. Alternatively we can write the generator coupling conditions similar to Equation (3.40),

$$\xi_0^+ = \xi_0^- + \text{diag}(c_1, \dots, c_N)^{-1} (1, \dots, 1)^T v_n. \tag{3.41}$$

Regarding the entropy production of coupling conditions (3.40) and (3.41) we only make the following proposition.

Proposition 3.14. *If R and G are sufficiently large, the whole scheme, including the boundary conditions, is entropy-stable.*

³ Note that M is invertible, as its determinant is given by $-(-1)^N (\prod_{i=1}^N c_i) (\sum_{i=1}^N \frac{1}{c_i})$ and $c_i > 0$.

Proof. The minmod scheme is piecewise linear on the space of all discretization points of all transmission lines, as can be seen in Equations (3.31), (3.32) and (3.33). The Lax-Wendroff scheme is linear. Also, there are only finitely many different pieces for the minmod limiter.

The coupling conditions are also linear on the same space. Together this means there is a collection of matrices $(W_k)_{k \in K}$ with finite index set K (enumerating all the different possible choices in Equation (3.33) for each discretization point ξ_i on all transmission lines), such that one step of the PDE scheme on the whole network can be expressed as $\text{PDE}(\xi) = W_k \xi$, where $k \in K$ is such that right branch in Equation (3.33) for the minmod limiter is chosen.

The ODE scheme on the other hand is also linear on the space of all ξ_i and can therefore be written as $\text{ODE}(\xi) = O\xi$, with a matrix O .

For the Lyapunov function after one step of both the PDE and the ODE scheme, we find

$$\begin{aligned} \mathcal{V}(\text{ODE} \circ \text{PDE} \xi) &= \xi^T W_k^T O^T O W_k \xi \\ &\leq \|O W_k\|^2 \xi^T \xi \\ &\leq \|O\|^2 \|W_k\|^2 \xi^T \xi \\ &\leq \|O\|^2 \max_{l \in K} \|W_l\|_l^2 \mathcal{V}(\xi), \end{aligned}$$

which does not increase $\mathcal{V}(\xi)$, if $\|O\| \max_{l \in K} \|W_l\|_l \leq 1$, which can be achieved by having sufficiently high R and G on all transmission lines as evidenced by Lemma 3.11. \square

We will rely on this property in the numerical examples in the next section and find the assumption to be justified for our purposes.

3.4 NUMERICAL EXAMPLES

For the computations we wrote a package⁴ for the Julia programming language⁵ (See [Bez+17]).

We will treat two examples, in which we compare exact periodic solutions, that correspond to power flow solutions and numerical solutions computed with the splitting scheme from Section 3.3. In addition we consider the behavior of the numerical Lyapunov function.

As we compare the splitting scheme to the power flow solution, the coupling conditions in the nodes are given by sinusoidal functions, namely by

$$\begin{aligned} v(t) &= \Re(V e^{i\omega t}), \\ i(t) &= \Re(I e^{i\omega t}). \end{aligned}$$

Therefore we provide only the quantity I and V for each node and of course ω for the whole network. The first example under consideration is a single transmission line between two nodes, one of which supplies voltage, while the other draws current.

⁴ The package is free software and can be used and modified by anyone. It can be found under https://bitbucket.org/efokken/telegraph_numerics

⁵ see <https://julialang.org/>

The parameters of this network are given in Table 3.1a and the parameters of the periodic boundary conditions in Table 3.1b. Note that this network satisfies the conditions of Proposition 3.9. We in-

Table 3.1: Parameters of the line, in addition $\omega = 4.0$.

(a) Line parameters.

R	4.0
L	6.0
G	2.0
C	1.0
ℓ	1.0

(b) Node parameters.

V_{start}	$5.0 + 3.0i$
I_{end}	$2.0 + 5.0i$

tegrate this problem from $t = 0$ to $t = 1.0$ with space discretization $\Delta x = 2^{-9}$. As Courant number we choose $\mathcal{C} = 0.8$. The numerical solution at $t = 1.0$ can be found in Figure 3.1a. To compute the error we first provide the analytical solution expressed in ξ^\pm . The analytical solution on every line is given by

$$\begin{aligned} v(x, t) &= \Re(V(x)e^{i\omega t}), \\ i(x, t) &= \Re(I(x)e^{i\omega t}), \end{aligned}$$

where V, I are taken from Equation (3.8) (with $m = \pm 1$). As voltages at the ends of the line we take the solution to the linear system (3.26). With these we compute the analytical expressions for ξ^\pm as

$$\xi^\pm(x, t) = \frac{1}{2} \left(i(x, t) \pm \sqrt{\frac{C}{L}} v(x, t) \right)$$

on every line. For an analytical solution ζ and a numerical solution ξ the maximal error of the numerical solution on a network over a time $T = N\Delta t$ is then given by

$$\Delta \xi = \max_{n=1:N} \max_{e \in \mathcal{E}} \left(\max_{j=1:J_e} \left(\left| \xi_{jn}^+ - \zeta_{jn}^+ \right|, \left| \xi_{jn}^- - \zeta_{jn}^- \right| \right) \right),$$

where \mathcal{E} is again the set of edges and J_e is the number of spatial steps on edge e and N is the total number of time steps. Here, $\zeta_{jn} = \zeta(j\Delta x, n\Delta t)$ is just the analytical solution evaluated at cell centers.

The time-dependent error over the whole line at time $t = 1$ is shown in Figure 3.1b for illustrating its usual appearance. The spikes in the error chart come from the minmod-scheme and lie exactly at the extrema of the solution. We remark that the boundary conditions do not seem to introduce any further error.

A second example is given by the network shown in Figure 3.2 on page 64, where the node in the middle draws current and the outer nodes supply voltage. The three transmission lines start at the node in the middle. Line parameters and the coupling conditions are given in Table 3.2 on page 64 and these parameters again fulfill the conditions of Proposition 3.9. This network is also integrated with a

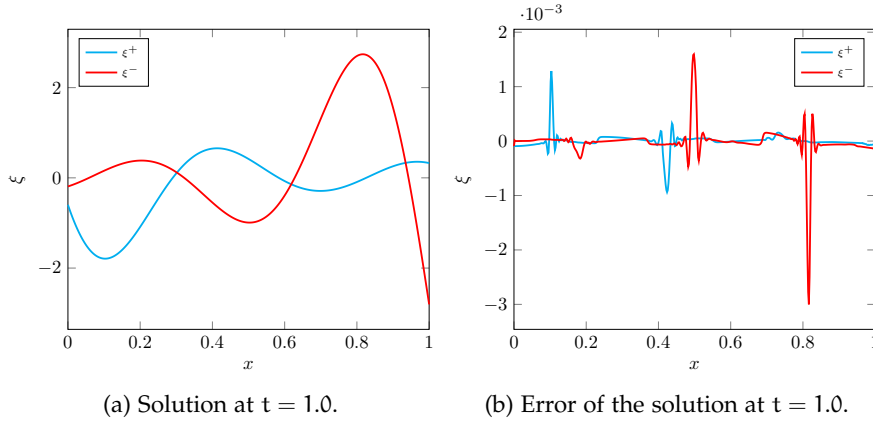


Figure 3.1: The single line example.

spatial step size of $\Delta x = 2^{-9}$ and again a Courant number of $\mathcal{C} = 0.8$. For this network we also examine the order of convergence. Therefore, we solve the problem with step sizes $\Delta x_i = 2^{-i}$ for $1 \leq i \leq 10$ and compute the corresponding maximal errors $\Delta \xi_i$ as well as the i -th estimate of the convergence order by

$$c_i = \frac{\frac{\Delta \xi_i}{\Delta x_{i-1}}}{\frac{\Delta \xi_{i-1}}{\Delta x_{i-1}}}.$$

In Table 3.3 on page 65 we see that the order of convergence approaches 2 with the pure Lax-Wendroff scheme and stays below 2 for the minmod-scheme as can be expected due to the order reduction at extrema. The fact that the pure Lax-Wendroff method approaches convergence order 2 may be surprising as it is well-known that it introduces oscillations at extrema due to its total variation production. Yet, as noted in and below Proposition 3.13, this effect is counteracted by the ODE part of the scheme. For the left-hand side of inequality (3.39) we find

$$1 - \mathcal{C}^2 + \mathcal{C} = 1.16.$$

For its right-hand side we compute

$$\Delta t = \mathcal{C} \frac{\Delta x}{\lambda} = \Delta t = \mathcal{C} \Delta x * \sqrt{LC} = 0.8 * 2^{-9} * 3 = 0.0046875$$

and

$$\exp(b_{\min} \Delta t) = \exp\left(\frac{1}{9} \cdot 0.0046875\right) \approx 1.0005 > 1.16.$$

for the finest discretization. So inequality (3.39) is far from fulfilled, yet the convergence order is 2. This demonstrates what we said earlier, namely that the inequality is much too conservative.

For this second example, the plots of the line solutions and errors are similar to those of the single line and can be found in Figure 3.3a and Figure 3.3b. Different parameters do not change the pictures much, which is why we do not show further images.

Lastly we compare the numerical (3.35) and analytical Lyapunov functions. The analytical Lyapunov function is given by

$$\mathcal{V}(t) = \mathcal{V}(0) \exp \left\{ -2 \min_{e \in \mathcal{E}} \left(\min \left(\frac{R_e}{L_e}, \frac{G_e}{C_e} \right) \right) t \right\}.$$

Table 3.2: Parameters of the network, in addition $\omega = 4.0$.

(a) Line parameters.				(b) Node parameters.		
line	$1 \rightarrow 2$	$1 \rightarrow 3$	$1 \rightarrow 4$	node	type	coupling condition
R	2.0	3.0	1.0	N1	current	$10.0 + 3.0i$
L	6.0	6.0	9.0	N2	voltage	$4.0 + 4.0i$
G	2.0	1.0	2.0	N3	voltage	$2.0 + 5.0i$
C	1.0	1.0	1.0	N4	voltage	$3.0 + 6.0i$
ℓ	2.0	2.0	2.0			

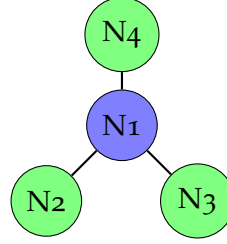


Figure 3.2: A network with three generators and one load.

For the comparison we use the network of the second example and the same initial data, but choose $\tilde{v}_l = 0$ and $v_g = 0$ for load and generator nodes respectively. With the numerical scheme we compute a solution and evaluate its Lyapunov function. The analytical Lyapunov function in this case is given by $\mathcal{V}(0) \exp(-\frac{2}{3}t)$. In Figure 3.4 we show both the estimate and the actually computed Lyapunov function.

Lastly we want to illustrate the severe restriction, the Courant number places on utilizing an explicit scheme for realistic power networks. Therefore we take some sample line parameters from [Brao4, Table 3 on page 26], namely parameters for an aerial transmission line for a voltage of 380 kV:

$$\begin{aligned}
 R &= 0.028 \text{ m}\Omega \text{ km}^{-1}, \\
 L &= 0.8 \text{ mH km}^{-1}, \\
 G &= 15 \text{ nS km}^{-1}, \\
 C &= 14 \text{ nF km}^{-1}
 \end{aligned} \tag{3.42}$$

and compute the wave speed given in (3.5) from them,

$$\lambda = \sqrt{LC}^{-1} \approx 298\,807 \text{ km s}^{-1} \approx 3 \times 10^8 \text{ m s}^{-1},$$

which is nearly the speed of light. The wave speed λ for a buried cable still fulfills $\lambda > 1 \times 10^8 \text{ m s}^{-1}$. Therefore for a rather coarse spatial step size of $\Delta x = 1 \text{ km}$ we get a restriction on the time step of

$$\Delta t \leq \frac{\Delta x}{\lambda} \leq 10^{-5} \text{ s},$$

so that 10^5 time steps per second have to be taken. For any realistic power network, computing so many time steps is at least difficult to implement and probably either too slow or too expensive to be practical. On the other hand, the mean lifetime of the exponential

Table 3.3: Order estimates of split schemes with different PDE schemes.

i	Lax Wendroff	Minmod
2	0.574	0.584
3	1.137	0.846
4	1.754	1.304
5	1.986	1.236
6	1.914	1.145
7	1.955	1.212
8	1.976	1.245
9	1.988	1.293
10	1.994	1.369

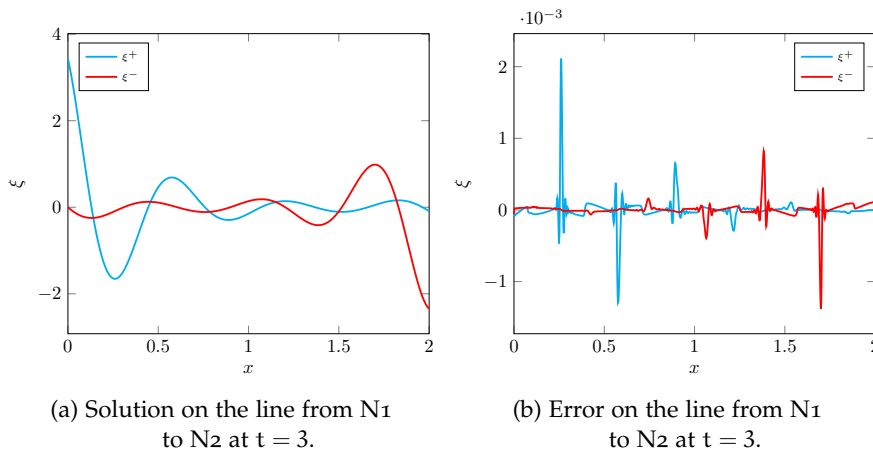


Figure 3.3: The network example.

decay of the Lyapunov function computed from (3.42) is about 15 s and this is a worst-case estimate as illustrated by Figure 3.4. Therefore for the simulation of normal network states using the power flow equations, as is done in energy system simulation, seems a valid choice.

3.5 SUMMARY AND OUTLOOK

In this chapter we have shown how the power flow equations used in engineering applications can be thought of as special solutions to the Telegrapher's equations on the transmission lines between the nodes of the power network. We also provided a second-order numerical scheme to compute these PDE solutions. The scheme — with possible exclusion of the coupling points — is furthermore Lyapunov stable, which mimics the analytical solution.

The numerical method is extendable to higher order as there are splitting schemes at least up to order six⁶ and there are entropy-stable ENO schemes of arbitrary order [FMT12] so that the Telegrapher's equations could be solved with much higher accuracy. Again however,

⁶ see <https://www.asc.tuwien.ac.at/~winfried/splitting/index.php>

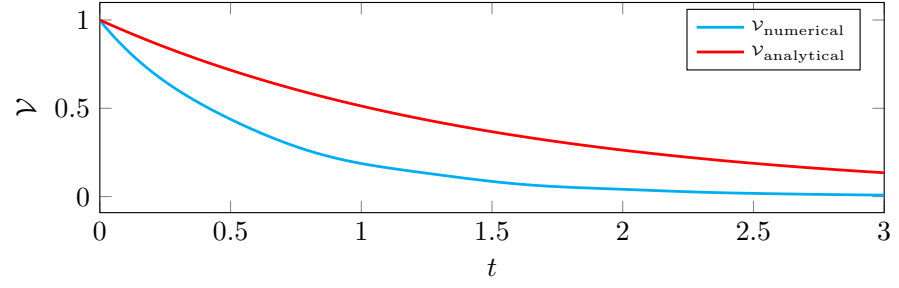


Figure 3.4: Comparison of the analytical estimate and the computed numerical Lyapunov function, both normed to 1.

comparing the computational costs, actually computing PDE solutions for realistic physical parameters can be only done for small instances and with high implementation effort.

Therefore we will use the power flow equations for all further power network modeling in this work. Although they are non-linear, their algebraic nature makes them much cheaper to handle than PDEs.

THE ISENTROPIC EULER EQUATIONS AND GAS-POWER-COUPLING

Having established the power flow equations as an appropriate model for power networks we now shift our attention to gas networks. In this chapter we will investigate the isentropic Euler equations as a gas model. These form a well-known non-linear hyperbolic balance law used to describe gas flow in pipelines. The purpose of this chapter is to find a suitable model for coupling the gas network to the power network. In order to do this we will first deal with the gas network itself, proving well-posedness of the Riemann problem as defined in Definition 2.8 of the isentropic Euler equations for certain pressure laws. Then we will extend the analysis to coupling conditions at vertices in the gas network and use this to prove the well-posedness of the gas-power coupling.

An alternative would be to go the route of the preceding chapter and replace the physically accurate model of the isentropic Euler equations with equations describing a steady state. Yet, the gas network is not quite as fast-paced as the power network, as the signal speed of roughly $c \approx 3.3 \times 10^2 \text{ m s}^{-1}$ suggests when compared to the signal speed of the power system, which fulfills $c \geq 1 \times 10^8 \text{ m s}^{-1}$. Therefore it may be desirable to resolve more than just a steady state. To this end, the implicit box scheme from Section 2.2.3 will prove valuable, as it allows for coarse and fine time discretizations, depending on whether more accuracy or less numerical cost is desired.

Coupling of the networks happens at nodes common to both, where a gas-fired generator is placed such that it can convert gas to electricity, see [HMS19; Zlo+16]. The coupling is modeled by an equality constraint relating generated power to consumed gas flow.

In Sections 4.1 and 4.2, we derive conditions on the gas pressure and power demand, respectively, that guarantee that the coupling conditions lead to well-posedness of the gas-power coupling. For the numerical study in Section 4.3, we give some validation results for the proposed discretizations and the coupling. Also, we present the influence of different pressure laws, while the actual coupling of gas and power networks will take place in Chapter 5.

4.1 GAS NETWORKS

Gas networks have been investigated very intensively during the last decade, see for example [BGH11; BHKo6b; KLB10]. Coupling conditions at nodes have been established to ensure well-posedness of the network solution [CGo8] and a rigorous numerical treatment [Egg18; GZ19; KLB10; ZA96]. As underlying structure for gas networks we use — as introduced in Section 2.1.6 — a directed graph with nodes

\mathcal{V}_{Gas} and edges \mathcal{E}_{Gas} . On this graph we define a graph problem, as in Definition 2.31.

The balance law in each pipe $e \in \mathcal{E}_{\text{Gas}}$ describes the gas flow. As balance law we choose the isentropic Euler equations in one space dimension:

$$\begin{pmatrix} \rho \\ q \end{pmatrix}_t + \begin{pmatrix} q \\ p(\rho) + \frac{q^2}{\rho} \end{pmatrix}_x = \begin{pmatrix} 0 \\ S(\rho, q) \end{pmatrix}, \quad (4.1)$$

where again $t \in \mathbb{R}_0^+$ is the time, $x \in [0, l_e]$ is the position along the pipeline of length l_e , the first state variable ρ is the density of the gas, the second state variable q is the momentum of the gas, p is a pressure function and S is a source terms that models friction at the pipeline walls. The pressure law must be a function of the form

$$p \in C^1(\mathbb{R}^+, \mathbb{R}^+), \quad p'(\rho) > 0 \text{ for all } \rho \in \mathbb{R}^+,$$

to ensure strict hyperbolicity. Note that our convention is $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$, meaning that we exclude the vacuum. A discussion of the vacuum will follow in Section 4.1.2. Usual examples are the isothermal pressure law

$$p(\rho) = c^2 \rho,$$

where c is the speed of sound in the gas, or, more generally, the γ -law

$$p(\rho) = \kappa \rho^\gamma, \quad (4.2)$$

for suitable constants κ and γ , which we examine in Proposition 4.2. In fact, we show that the class of possible pressure laws can be enlarged, leading to non-standard pressure functions. For all pressure laws, the *speed of sound* $c(\rho)$ inside the gas is given by

$$c(\rho) = \sqrt{p'(\rho)},$$

while the flow speed of individual gas particles is given by

$$v(\rho) = \frac{q}{\rho}.$$

It is straightforward to compute the eigenvalues λ and the eigenvectors r for the system (4.1),

$$\begin{aligned} \lambda_1(\rho, q) &= \frac{q}{\rho} - \sqrt{p'(\rho)} = v(\rho, q) - c(\rho), \\ \lambda_2(\rho, q) &= \frac{q}{\rho} + \sqrt{p'(\rho)} = v(\rho, q) + c(\rho), \\ r_1(\rho, q) &= \begin{pmatrix} -1 \\ -\lambda_1(\rho, q) \end{pmatrix}, \\ r_2(\rho, q) &= \begin{pmatrix} 1 \\ \lambda_2(\rho, q) \end{pmatrix}. \end{aligned} \quad (4.3)$$

Using Definition 2.10, it is easy to see that these eigenvalues are genuinely non-linear if and only if there holds $2p'(\rho) + p''(\rho) \neq 0$ for all $\rho \in \mathbb{R}^+$. As is customary we concentrate on the case

$$2p'(\rho) + p''(\rho) > 0. \quad (4.4)$$

At the network boundaries, typically the pressure p_{in} or the flow q_{out} is given. Further, the following coupling conditions, which were described in [BHKo6b; Hero7], are used at all inner nodes of \mathcal{V}_{Gas} in order to implicitly prescribe boundary conditions for the isentropic Euler equations.

- equality of pressure: the pressure values at the ends of all arcs connected to the same node must be equal, that is, there is a coupling pressure p_{coupling} such that

$$p_e = p_{\text{coupling}} \quad (4.5)$$

holds at the end of all arcs e connected to the junction. Although this condition yields entropy-violating solutions at certain junctions, see [HHW20; Rei15] for details, it is well-suited for realistic pipeline scenarios, as will be seen in Chapter 5. The great advantage of it is, that equality of pressure also yields equality of density, which simplifies the analysis of Riemann problems at junctions,

$$\rho_e = \rho_{\text{coupling}},$$

because p as a strictly increasing function is one-to-one.

- conservation of mass: the sum of all incoming fluxes must equal the sum of all outgoing fluxes (including eventual boundary terms), i.e.,

$$\sum_{\text{incoming pipes}} q_{\text{pipe}} = \sum_{\text{outgoing pipes}} q_{\text{pipe}}, \quad (4.6)$$

which encodes conservation of mass at a junction.

Here, using the eigen fields from Equation (4.3), it is straightforward to show that the condition (2.18) is fulfilled, as long as all states at the junction fulfill

$$\left| \frac{q}{\rho} \right| < c(\rho) = \sqrt{p'(\rho)}. \quad (4.7)$$

This condition is called the *sub-sonic* condition and we will assume it to hold true for the rest of the thesis. Not though, that other coupling conditions are possible, see for example [GHM17]. As the name suggests, it means, that the flow speed $v(\rho) = \frac{q}{\rho}$ of the gas is less than the speed of sound. This condition also fixes the eigenvalue structure to $\lambda_1 < 0 < \lambda_2$.

Concretely for n attached pipes, all outgoing, the determinant D of Equation (2.18) is given by

$$D = \sum_{i=1}^n \lambda_2(\rho_i, q_i),$$

which is positive for sub-sonic states ρ_i, q_i in all attached pipes. For a single node with a finite number of adjacent edges extending to infinity and with initial data existence of weak entropic solutions in the space of functions with bounded variation has been shown e.g.

in [CGo8]. In [HHW20] existence of weak integrable solutions on a graph has been established. Similar results are also available for other choices of coupling conditions and can be found for example in [Bre+14].

Note that these coupling conditions are virtually the same as those of the power network, if one interprets voltage as the pressure of electricity and current as the flow.

4.1.1 Well-posedness of the Riemann problem for the isentropic Euler equation

We now investigate conditions for the pressure function (4.2) that guarantee well-posedness of the Riemann problem for the isentropic Euler equations without a source term, see [Bre+14] for an introduction. This is an important ingredient in applying the front-tracking technique of Section 2.1.4, see e.g. [CHSo8; Gug+12]. To guarantee strict hyperbolicity (see Definition 2.4), we start with the assumption $p' > 0$. We also assume that the initial data $U_l = (\rho_l, q_l)$ and $U_r = (\rho_r, q_r)$ of the Riemann problem is sub-sonic (see Equation (4.7)).

A solution to the Riemann problem with left state U_l and right state U_r can be constructed with an intersection point of the first forward Lax curve through U_l and the second backward Lax curve through U_r as per Theorem 2.22.

The Lax curves as defined in Definition 2.13 are curves in the state space \mathcal{S} . But for the isentropic Euler equations it is possible to parameterize them by the density. In that case, the first component of the Lax curves is just the parameter itself and therefore we consider only the projection to the second component, that is, instead of

$$L : \rho \mapsto \begin{pmatrix} \rho \\ q(\rho) \end{pmatrix},$$

we consider $L : \rho \mapsto q(\rho)$. In this sense, the Lax curve of all states $U = (\rho, q)$ reachable via 1-rarefaction or 1-shock (that is, waves corresponding to λ_1 above, see Definition 2.18) from a left state $U_l = (\rho_l, q_l)$ is given by (taken from [CGo8])

$$\begin{aligned} L_1(\rho; \rho_l, q_l) &= L_1^+(\rho; \rho_l, q_l) \\ &= \begin{cases} \rho \left(\frac{q_l}{\rho_l} + \int_{\rho_l}^{\rho} \frac{c(s)}{s} ds \right) & \text{for } \rho \leq \rho_l \text{ (rarefaction),} \\ \rho \frac{q_l}{\rho_l} - \sqrt{f(\rho, \rho_l)} & \text{for } \rho_l \leq \rho \text{ (shock),} \end{cases} \end{aligned}$$

where f is defined by

$$\begin{aligned} f(\rho, \rho_l) &= a(\rho, \rho_l) \Delta p(\rho, \rho_l), \\ a(\rho, \rho_l) &= \frac{\rho}{\rho_l} (\rho - \rho_l), \\ \Delta p(\rho, \rho_l) &= p(\rho) - p(\rho_l), \end{aligned} \tag{4.8}$$

such that $U = (\rho, L_1(\rho))$ holds.

Analogously, the Lax curve of all states $U = (\rho, q)$ reachable via 2-rarefaction or 2-shock from a right state $U_r = (\rho_r, q_r)$ is given by

$$\begin{aligned} L_r(\rho; \rho_r, q_r) &= L_2^-(\rho; \rho_r, q_r) \\ &= \begin{cases} \rho \left(\frac{q_r}{\rho_r} - \int_{\rho_r}^{\rho} \frac{c(s)}{s} ds \right) & \text{for } \rho \leq \rho_r \text{ (rarefaction),} \\ \rho \frac{q_r}{\rho_r} + \sqrt{f(\rho, \rho_r)} & \text{for } \rho_r \leq \rho \text{ (shock).} \end{cases} \end{aligned}$$

Whenever we write $L_l(\rho)$ and $L_r(\rho)$, it will mean $L_l(\rho, \rho_l, q_l)$ and $L_r(\rho, \rho_r, q_r)$, respectively. Beware that later, in Section 4.1.3, different second and third arguments will appear and we will write them out again.

A solution to the Riemann problem can then be constructed from a point ρ , where

$$L_l(\rho) - L_r(\rho) = 0, \quad (4.9)$$

by setting the solution as a 1-wave (shock or rarefaction, depending on whether $\rho > \rho_l$ or $\rho < \rho_l$) between the states U_l and $U = (\rho, L_{r/l}(\rho))$ and a 2-wave between U and U_r . The ordering of the eigenvalues and the Lax entropy condition (Definition 2.17) guarantee that these waves do not meet. An example of Lax curves and their intersection is shown in Figure 4.1. Note that

$$-L_r(\rho; \rho_r, q_r) = L_l(\rho; \rho_r, -q_r). \quad (4.10)$$

So L_l and $-L_r$ share most properties. Therefore Equation (4.9) is best

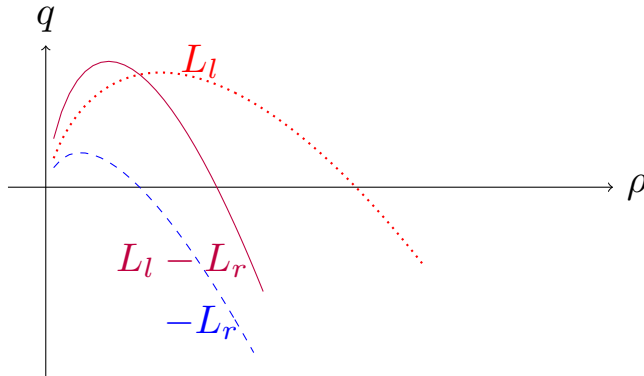


Figure 4.1: Typical Lax curves with pressure law $p(\rho) = c^2\rho$. The zero of the purple curve is the desired solution.

understood not as a difference of two functions but as a sum of two very similar functions. Similar to [HMS19], we show certain properties of the Lax curves that are used to define the new pressure laws. There will appear three propositions, which are numbered by A, B and C. These propositions present lists of conditions that are closely related. To make things tractable, each list bears the same numbering but has the letter of the corresponding proposition in front. For example, the conditions A1, B1 and C1 all govern concavity of the Lax curves. Before we state the proposition, recall from Proposition 2.19 that Lax curves are twice continuously differentiable.

Proposition A (Proposition 4.1). *Let $L \in C^2(\mathbb{R}^+, \mathbb{R}^+)$ and let it fulfill Conditions A1 through A3 (where only one of A3(a) and A3(b) must hold):*

- (A1) $L'' \leq 0$,
- (A2) $\exists \rho > 0 : L(\rho) < 0$,
- (A3) (a) $0 < \lim_{\rho \rightarrow 0} L(\rho) < \infty$ or
(b) $\lim_{\rho \rightarrow 0} L(\rho) = 0$ and $\lim_{\rho \rightarrow 0} L'(\rho) > 0$.

Then there is a unique $\rho > 0$ with $L(\rho) = 0$ such that for all $\hat{\rho} > \rho$ there holds $L(\hat{\rho}) < 0$ and $L(\rho) \rightarrow -\infty$ for $\rho \rightarrow \infty$.

Proof. Condition A3 ensures positive values near 0, Condition A2 ensures negative values for some $\rho > 0$, together yielding a zero in between and condition A1 makes L concave, guaranteeing uniqueness of the zero and implying then that $L(\rho) \rightarrow -\infty$ for $\rho \rightarrow \infty$. \square

Note, that if two functions L_a, L_b satisfy the prerequisites of Proposition A, so does $L_a + L_b$. Therefore we search for conditions on pressure functions that make L_l and $-L_r$ satisfy the conditions of Proposition A, as then their sum in Equation (4.9) will, too, and therefore this sum will have a unique zero.

The main result of this section is the following. It generalizes the usual notion of γ -laws to allow for $\gamma < 1$. Further, as we will show below (Proposition C), positive linear combinations of such valid pressure laws again lead to valid pressure laws.

Proposition 4.2 (generalized γ -laws). *Let the derivative of the pressure be given by*

$$p'(\rho) = \alpha \rho^\delta.$$

This translates to pressure functions given by

$$p(\rho) = \frac{\alpha}{\gamma} \rho^\gamma + \text{const},$$

for $\delta \neq -1$ and

$$p(\rho) = \alpha \log(\rho) + \text{const},$$

for $\delta = -1$. For these pressure functions there holds:

- *The Lax curves through any sub-sonic initial states U_l, U_r have a unique intersection point at some $\rho > 0$ if and only if $-2 < \delta \leq 2$ and $\alpha > 0$.*
- *For every δ with $|\delta| > 2$ there are sub-sonic states U_l, U_r such that the Lax curves have no intersection at all.*

Note that the case $\delta = -2$ or $\gamma = -1$ violates genuine non-linearity (Inequality (4.4)) and is hence not admissible.

Expressed in the usual form of γ -laws, Proposition 4.2 means $p(\rho) = \kappa \rho^\gamma$ is a valid pressure function if and only if $0 < \gamma \leq 3$ and $\kappa > 0$ or $-1 < \gamma < 0$ and $\kappa < 0$. The proof of Proposition 4.2 will be given at the end of this section. For this proof we will reformulate Conditions A1 through A3 in Proposition A in terms of the pressure function.

Because of Equation (4.10), both Lax curves behave essentially identically and therefore we prove our findings for L_l only, as the proofs for $-L_r$ are the same. Along the way the derivatives of the Lax curves will be important and so we provide them here,

$$L'_l(\rho) = \begin{cases} \frac{q_l}{\rho_l} + \int_{\rho}^{\rho_l} \frac{c(s)}{s} ds - c(\rho) & \text{for } \rho \leq \rho_l, \\ \frac{q_l}{\rho_l} - \frac{f'}{2\sqrt{f}} & \text{for } \rho_l \leq \rho, \end{cases}$$

$$L''_l(\rho) = \begin{cases} -\left(\frac{c(\rho)}{\rho} + c'(\rho)\right) & \text{for } \rho \leq \rho_l, \\ -\frac{2f''f - (f')^2}{4\sqrt{f^3}} & \text{for } \rho_l \leq \rho, \end{cases}$$

and

$$L'_r(\rho) = \begin{cases} \frac{q_r}{\rho_r} - \int_{\rho}^{\rho_r} \frac{c(s)}{s} ds + c(\rho) & \text{for } \rho \leq \rho_r, \\ \frac{q_r}{\rho_r} + \frac{f'}{2\sqrt{f}} & \text{for } \rho_r \leq \rho, \end{cases}$$

$$L''_r(\rho) = \begin{cases} \frac{c(\rho)}{\rho} + c'(\rho) & \text{for } \rho \leq \rho_r, \\ \frac{2f''f - (f')^2}{4\sqrt{f^3}} & \text{for } \rho_r \leq \rho. \end{cases}$$

Next, we will give conditions under which Proposition A is applicable to L_l . Before we do so, we provide three lemmas. Lemma 4.5 will be used afterwards in the proof of Proposition 4.6, while Lemmas [ref*wellposedness:lemma:1](#) and 4.4 are needed for the proof of Lemma 4.5 itself. Note here again that our convention is $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$. If we included 0, the lemmas would hold trivially by continuity.

Lemma 4.3. *Let $g \in C^1(\mathbb{R}^+, \mathbb{R}^+)$ be a non-negative function, $g \geq 0$, and let G be given by $G(\rho) = \int_{\rho}^{\rho_l} g(s) ds$. Then there holds*

1. *If $\rho^2 g(\rho) \xrightarrow{\rho \rightarrow 0} 0$, then $\rho G(\rho) \xrightarrow{\rho \rightarrow 0} 0$.*
2. *If $\rho G(\rho) \xrightarrow{\rho \rightarrow 0} 0$, then $\liminf_{\rho \rightarrow 0} \rho^2 g(\rho) = 0$.*

Proof.

1. By assumption $\rho^2 g(\rho) \xrightarrow{\rho \rightarrow 0} 0$. For $m \in \mathbb{N}$ choose $\rho_m > 0$ such that $g(\rho) \leq \frac{1}{m\rho^2}$ for $\rho < \rho_m$. Now choose $\rho_{m,0} < \rho_m$ so small that

$$\rho_{m,0} \left(\int_{\rho_m}^{\rho_l} g(s) ds - \frac{1}{m\rho_m} \right) \leq \frac{1}{m}.$$

Then, for $\rho < \rho_{m,0}$, there holds

$$\begin{aligned} \rho \int_{\rho}^{\rho_l} g(s) ds &\leq \rho \int_{\rho_m}^{\rho_l} g(s) ds + \frac{1}{m} \rho \int_{\rho}^{\rho_m} \frac{1}{s^2} ds \\ &= \rho \left(\int_{\rho_m}^{\rho_l} g(s) ds - \frac{1}{m\rho_m} \right) + \frac{1}{m} \\ &\leq \frac{2}{m}. \end{aligned}$$

Therefore $\lim_{\rho \rightarrow 0} \rho \int_{\rho}^{\rho_l} g(s) ds \leq \frac{2}{m}$ for all $m \in \mathbb{N}$. As $g \geq 0$, we also have $0 \leq \lim_{\rho \rightarrow 0} \rho \int_{\rho}^{\rho_l} g(s) ds$. Summarizing, we get $\lim_{\rho \rightarrow 0} \rho G(\rho) = \lim_{\rho \rightarrow 0} \rho \int_{\rho}^{\rho_l} g(s) ds = 0$.

2. We prove by contradiction: Assume there are $\rho_0 > 0$ and $\alpha > 0$ such that $\rho^2 g(\rho) \geq \alpha$ for $\rho < \rho_0$. Then $g(\rho) \geq \frac{\alpha}{\rho^2}$ for such ρ . Therefore

$$\begin{aligned} \rho \int_{\rho}^{\rho_1} g(s) ds &\geq \rho \int_{\rho_0}^{\rho_1} g(s) ds + \rho \alpha \int_{\rho}^{\rho_0} \frac{1}{s^2} ds \rightarrow 0 + \alpha \rho \left[-\frac{1}{s} \right]_{\rho}^{\rho_0} \\ &\rightarrow \alpha, \end{aligned}$$

which contradicts $\rho G(\rho) \xrightarrow{\rho \rightarrow 0} 0$. \square

Lemma 4.4. Let $g \in C^1(\mathbb{R}^+, \mathbb{R}^+)$, $g \geq 0$ and $\liminf_{\rho \rightarrow 0} \rho^2 g(\rho) = 0$. Let also $(\rho^2 g(\rho))' \geq 0$. Then, $\limsup_{\rho \rightarrow 0} \rho^2 g(\rho) = \liminf_{\rho \rightarrow 0} \rho^2 g(\rho) = 0$.

Proof. We prove by contradiction. Let $\limsup_{\rho \rightarrow 0} \rho^2 g(\rho) > \liminf_{\rho \rightarrow 0} \rho^2 g(\rho)$. Then it is easily seen that $\liminf_{\rho \rightarrow 0} (\rho^2 g(\rho))' = -\infty < 0$, resulting in a contradiction. \square

The following lemma contains the result we need later on.

Lemma 4.5. Let $g \in C^1(\mathbb{R}^+, \mathbb{R}^+)$ be a non-negative function, $g \geq 0$, such that also $(\rho^2 g(\rho))' \geq 0$ for all $\rho > 0$ and let G be given by $G(\rho) = \int_{\rho}^{\rho_1} g(s) ds$. Then, there holds $\rho^2 g(\rho) \xrightarrow{\rho \rightarrow 0} 0$ if and only if $\rho G(\rho) \xrightarrow{\rho \rightarrow 0} 0$.

Proof. Applying Lemmas 4.3 and 4.4 yields the result. \square

The role of g in Lemma 4.5 will be played by $p'(\rho)$ and $\frac{c(\rho)}{\rho}$ in the following. Let us now put the focus on the Lax curve L_1 again. The following Proposition B will turn Conditions A1 through A3 in Proposition A into conditions on the pressure function.

Proposition B (Proposition 4.6). Let $p \in C^2(\mathbb{R}^+)$ with $p' > 0$. Then Conditions A1, A2, A3 hold for L_1 for all ρ_l, q_l with $\left| \frac{q_l}{\rho_l} \right| < \sqrt{p'(\rho_l)}$ if and only if Conditions B1 through B3 hold (again with only one of B3(a) and B3(b) fulfilled).

(B1) These inequalities hold:

$$2p'(\rho) + \rho p''(\rho) \geq 0 \quad \forall \rho > 0, \quad (4.11)$$

$$\Delta p^2 + \alpha^2 \left(2\Delta p p'' - (p')^2 \right) + \frac{1}{2} (\alpha^2)' (\Delta p^2)' \geq 0 \quad \forall \rho_l > 0, \rho > \rho_l, \quad (4.12)$$

where the arguments ρ and ρ_l of α and Δp have been omitted for readability, see (4.8).

(B2) Let $p_{\infty} = \lim_{\rho \rightarrow \infty} p(\rho) \in \mathbb{R} \cup \{\infty\}$. For all $\rho > 0$ there holds

$$p_{\infty} - \rho p'(\rho) - p(\rho) \geq 0.$$

(B3) (a) There is $p_0 > 0$ such that $p(\rho) = -\frac{p_0}{\rho} + o_{\rho \rightarrow 0} \left(\frac{1}{\rho} \right)$ or

(b) $p(\rho) \in o_{\rho \rightarrow 0} \left(\frac{1}{\rho} \right)$ and $\lim_{\rho \rightarrow 0} \int_{\rho}^{\rho_l} \frac{c(s)}{s} ds - c(\rho) - c(\rho_l) \geq 0$ for all $\rho_l > 0$.

Proof.

$A_1 \Leftrightarrow B_1$: The equivalence $A_1 \Leftrightarrow B_1$ is immediate from the definitions, Inequality (4.11) is for the rarefaction part, inequality (4.12) for the shock part. Note that

$$2ff'' - (f')^2 = \Delta p^2 + a^2 \left(2\Delta p p'' - (p')^2 \right) + \frac{1}{2}(a^2)'(\Delta p^2)'.$$

$A_1, A_3(a) \Rightarrow B_1, B_3(a)$: Assume Conditions A_1 and $A_3(a)$. Let $F(\rho) = \int_{\rho}^{\rho_1} \frac{c(s)}{s} ds$. Then

$$0 < l = \lim_{\rho \rightarrow 0} L_1(\rho) = \lim_{\rho \rightarrow 0} \rho \int_{\rho}^{\rho_1} \frac{c(s)}{s} ds = \lim_{\rho \rightarrow 0} \rho F(\rho).$$

Therefore $F(\rho) = \frac{l}{\rho} + o_{\rho \rightarrow 0} \left(\frac{1}{\rho} \right)$. Note now that

$$\left(\rho^2 \frac{c(\rho)}{\rho} \right)' = \frac{1}{2\sqrt{p'(\rho)}} (2p'(\rho) + \rho p''(\rho)) \geq 0,$$

due to Condition A_1 . Lemma 4.5 therefore shows that

$$\frac{c(\rho)}{\rho} = -F'(\rho) = \frac{l}{\rho^2} + o_{\rho \rightarrow 0} \left(\frac{1}{\rho^2} \right),$$

and hence

$$p'(\rho) = c(\rho)^2 = \frac{l^2}{\rho^2} + o_{\rho \rightarrow 0} \left(\frac{1}{\rho^2} \right),$$

which again with Lemma 4.5 yields

$$p(\rho) = -\frac{l^2}{\rho} + o_{\rho \rightarrow 0} \left(\frac{1}{\rho} \right).$$

$A_1, A_3(a) \Leftarrow B_1, B_3(a)$: Assume now Conditions B_1 and $B_3(a)$. We now note that

$$\left(\rho^2 p'(\rho) \right)' = \rho (2p'(\rho) + \rho p''(\rho)) \geq 0,$$

because of B_1 and use the Lemma to arrive at $A_3(a)$.

$A_1, A_3(b) \Leftarrow B_1, B_3(b)$: We assume Condition $B_3(b)$. The last proof (here we need again Conditions A_1 and B_1 respectively) also shows that

$$\lim_{\rho \rightarrow 0} L_1(\rho) = 0 \Leftrightarrow p \in o_{\rho \rightarrow 0} \left(\frac{1}{\rho} \right).$$

For the derivative we find

$$\begin{aligned} \lim_{\rho \rightarrow 0} L_1'(\rho) &= \lim_{\rho \rightarrow 0} \frac{q_l}{\rho_l} + \int_{\rho}^{\rho_1} \frac{c(s)}{s} ds - c(\rho) \\ &> \lim_{\rho \rightarrow 0} \int_{\rho}^{\rho_1} \frac{c(s)}{s} ds - c(\rho) - c(\rho_l), \end{aligned}$$

due to the sub-sonic condition. This is non-negative due to Condition $B_3(b)$.

A1, A3(b) \Rightarrow B1,B3(b): Assume that B3(b) does not hold, that is

$$\lim_{\rho \rightarrow 0} \int_{\rho}^{\rho_1} \frac{c(s)}{s} ds - c(\rho) - c(\rho_1) \leq -\delta < 0.$$

Choosing ρ_1, q_1 such that $\frac{q_1}{\rho_1} = -c(\rho_1) + \frac{1}{2}\delta$ yields

$$\lim_{\rho \rightarrow 0} L_1'(\rho) \leq -\frac{1}{2}\delta < 0.$$

A2 \Leftarrow B2: First of all we note that the limit exists, as $p' > 0$. For $\rho > \rho_1$ we write out

$$\begin{aligned} L_1(\rho) &= \rho \left(\frac{q_1}{\rho_1} - \sqrt{\frac{f(\rho, \rho_1)}{\rho^2}} \right) \\ &= \rho \left(\frac{q_1}{\rho_1} - \sqrt{\frac{1}{\rho_1} \left(1 - \frac{\rho_1}{\rho} \right) (p(\rho) - p(\rho_1))} \right). \end{aligned} \quad (4.13)$$

This is less than zero if and only if the bracket is less than zero. If now $p_\infty = \infty$, this is fulfilled for some $\rho > \rho_1$. If $p_\infty < \infty$, we have the limit

$$\lim_{\rho \rightarrow \infty} \sqrt{\frac{1}{\rho_1} \left(1 - \frac{\rho_1}{\rho} \right) (p(\rho) - p(\rho_1))} = \sqrt{\frac{p_\infty - p(\rho_1)}{\rho_1}}.$$

Let now δ and ϵ_ρ be defined by

$$\begin{aligned} \frac{q_1}{\rho_1} + \delta &= c(\rho_1), \\ \sqrt{\frac{1}{\rho_1} \left(1 - \frac{\rho_1}{\rho} \right) (p(\rho) - p(\rho_1))} &= \sqrt{\frac{p_\infty - p(\rho_1)}{\rho_1}} - \epsilon_\rho. \end{aligned}$$

Then $\delta > 0$ as the state is sub-sonic and $\epsilon_\rho \rightarrow 0$ for $\rho \rightarrow \infty$. Therefore let ρ be so great that $\delta - \epsilon_\rho > 0$. Then, due to B2, we find

$$\begin{aligned} \frac{q_1}{\rho_1} &< \frac{q_1}{\rho_1} + \delta - \epsilon_\rho = c(\rho_1) - \epsilon_\rho \leq \sqrt{\frac{p_\infty - p(\rho_1)}{\rho_1}} - \epsilon_\rho \\ &= \sqrt{\frac{1}{\rho_1} \left(1 - \frac{\rho_1}{\rho} \right) (p(\rho) - p(\rho_1))}, \end{aligned}$$

which, according to (4.13), implies A2.

A2 \Rightarrow B2: In case $p_\infty = \infty$, condition B2 is fulfilled. Let now $p_\infty < \infty$ and let there be for all sub-sonic (ρ_1, q_1) a $\rho_- > 0$ such that

$$\frac{q_1}{\rho_1} < \sqrt{\frac{1}{\rho_1} \left(1 - \frac{\rho_1}{\rho_-} \right) (p(\rho_-) - p(\rho_1))}, \quad (4.14)$$

that is, let A2 be true. Note that the right-hand-side is increasing in ρ_- because its derivative is positive. Therefore taking suprema of (4.14) yields

$$c(\rho_1) \leq \sqrt{\frac{p_\infty - p(\rho_1)}{\rho_1}} \Rightarrow p'(\rho_1) \leq \frac{p_\infty - p(\rho_1)}{\rho_1}. \quad \square$$

With Proposition B we have a list of Conditions B₁, B₂, B₃ which is equivalent to Conditions A₁, A₂, A₃ from Proposition A but now only involves the pressure function. This is fortunate as we want to prove Proposition 4.2, which only contains pressure functions. In principle, we could just plug the pressure functions into our Conditions B and check, what generalized γ -laws are allowed. But there is more insight to be gained by simplifying the conditions further. We will do so in Propositions 4.7, 4.8 and 4.9. Yet, this comes at a price, the next list of conditions in Proposition C below will only be sufficient conditions for Propositions A and B to hold.

Proposition 4.7. *Let $p \in C^3(\mathbb{R}^+)$. Condition B₁ holds if*

$$\begin{aligned} 2p'(\rho) + \rho p''(\rho) &\geq 0, \\ 6p'(\rho) + 6\rho p''(\rho) + \rho^2 p'''(\rho) &\geq 0 \end{aligned}$$

hold for all $\rho > 0$.

Proof. Here we use f from Equation (4.8) and denote the derivative with respect to ρ by a prime. We note that $(2ff'' - (f')^2)(\rho_l, \rho_l) = 0$ and that

$$(2ff'' - (f')^2)' = 2ff''''.$$

As $f \geq 0$, we see that $2ff'' - (f')^2 \geq 0$ for all $\rho \geq \rho_l$ if $f'''(\rho, \rho_l) \geq 0$ for all $\rho \geq \rho_l$. This was also proved in [GHM17]. This is the case if and only if $\rho_l f'''(\rho, \rho_l) \geq 0$. For this we find

$$\begin{aligned} \rho_l f'''(\rho, \rho_l) &= 6p'(\rho) + 3(2\rho - \rho_l)p''(\rho) + \rho(\rho - \rho_l)p'''(\rho) \\ &= [6p'(\rho) + 6\rho p''(\rho) + \rho^2 p'''(\rho)] - \rho_l [3p''(\rho) + \rho p'''(\rho)], \end{aligned}$$

which is an affine function in ρ_l . Hence for given $\rho > 0$ this function takes its minimum in one of the edges of the simplex $\{\rho_l \mid 0 \leq \rho_l \leq \rho\}$. The values on these are given by

$$\begin{aligned} &6p'(\rho) + 6\rho p''(\rho) + \rho^2 p'''(\rho) \text{ and} \\ &3(2p'(\rho) + \rho p''(\rho)). \end{aligned} \quad \square$$

The next condition we will replace is .

Proposition 4.8. *Condition B3(b) is fulfilled if either of these conditions is satisfied:*

- C₃(b) (i) $\lim_{\rho \rightarrow 0} c(\rho) = 0$ and $2p'(\rho) - \rho p''(\rho) \geq 0$ for all $\rho > 0$.
(ii) $0 < \lim_{\rho \rightarrow 0} c(\rho) < \infty$.
(iii) There is a $\eta \in (0, 1)$ such that $\lim_{\rho \rightarrow 0} \rho^\eta c(\rho) \in \mathbb{R}$, so that in particular the limit is finite.

Proof.

(i) Here we have

$$\begin{aligned}
\lim_{\rho \rightarrow 0} \int_{\rho}^{\rho_l} \frac{c(s)}{s} ds - c(\rho) - c(\rho_l) &= \lim_{\rho \rightarrow 0} \int_{\rho}^{\rho_l} \left(\frac{c(s)}{s} - c'(s) \right) ds - 2c(\rho) \\
&= \lim_{\rho \rightarrow 0} \int_{\rho}^{\rho_l} \left(\frac{c(s)}{s} - c'(s) \right) ds \\
&= \lim_{\rho \rightarrow 0} \int_{\rho}^{\rho_l} \frac{1}{2s\sqrt{p'(s)}} (2p'(s) - sp''(s)) ds \\
&\geq 0.
\end{aligned}$$

(ii) In this case the integral is unbounded near zero, but $c(\rho) + c(\rho_l)$ is finite.

(iii) In this case there is $\alpha > 0$ such that $\rho^\eta c(\rho) = \alpha(1 + r(\rho))$ with $r(\rho) \rightarrow 0$. Then, for every $m \in \mathbb{N}$ there is $\rho_m > 0$ such that for all $\rho < \rho_m$ there holds $c(\rho) \geq \alpha \left(1 - \frac{1}{m}\right) \rho^{-\eta}$. Choose now m so great that

$$\left(1 - \frac{1}{m}\right) \frac{1}{\eta} > 1,$$

and then $\rho_{m,0}$ so small that for all $\rho < \rho_{m,0}$ there holds

$$\left(1 - \frac{1}{m}\right) \frac{1}{\eta} - r(\rho) = 1 + \theta,$$

with $\theta > 0$. Define also

$$C_m = \int_{\rho_m}^{\rho_l} \frac{c(s)}{s} ds.$$

Then we find for $\rho < \rho_{m,0}$

$$\begin{aligned}
&\int_{\rho}^{\rho_l} \frac{c(s)}{s} ds - c(\rho) - c(\rho_l) \\
&= \int_{\rho}^{\rho_m} \frac{c(s)}{s} ds + \int_{\rho_m}^{\rho_l} \frac{c(s)}{s} ds - c(\rho) - c(\rho_l) \\
&\geq \alpha \left(1 - \frac{1}{m}\right) \int_{\rho}^{\rho_m} s^{-\eta-1} ds + C_m - c(\rho) - c(\rho_l) \\
&= \alpha \left(\left(1 - \frac{1}{m}\right) \frac{1}{\eta} - 1 - r(\rho) \right) \rho^{-\eta} \\
&\quad - \alpha \left(1 - \frac{1}{m}\right) \frac{1}{\eta} \rho_m^{-\eta} + C_m - c(\rho_l) \\
&> \alpha \theta \rho^{-\eta} - \alpha \left(1 - \frac{1}{m}\right) \frac{1}{\eta} \rho_m^{-\eta} + C_m - c(\rho_l) \\
&\xrightarrow{\rho \rightarrow 0} +\infty > 0.
\end{aligned}$$

□

Proposition 4.9. *All conditions only depend on differences in pressures, that is, if $C \in \mathbb{R}$ is a constant then if a pressure function p satisfies our*

conditions, so does $p + C$. This means that if $p_\infty < \infty$ in Condition B2, we can instead choose $p_\infty = 0$, yielding the clearer condition

$$p'(\rho) \leq -\frac{p(\rho)}{\rho}.$$

We now have all tools at hand but now we want to apply them to the Riemann problem of the isentropic Euler equations. Here we have to adhere to Inequality (4.4) again to guarantee genuine non-linearity. Therefore the following Condition C1 contains a strict inequality.

Proposition C (Proposition 4.10). *The Riemann problem for arbitrary sub-sonic left- and right-hand states is well-posed if the pressure function p satisfies the following conditions (in addition to $p' > 0$):*

(C1) *both inequalities of Proposition 4.7:*

$$\begin{aligned} 2p'(\rho) + \rho p''(\rho) &> 0, \\ 6p'(\rho) + 6\rho p''(\rho) + \rho^2 p'''(\rho) &\geq 0, \end{aligned}$$

(C2) *one of the conditions according to Condition B2:*

- (a) $p \rightarrow \infty$ for $\rho \rightarrow \infty$ or
- (b) $p_\infty = \lim_{\rho \rightarrow \infty} p(\rho) < \infty$ and
 $p'(\rho) \leq \frac{p_\infty - p(\rho)}{\rho}$ for all $\rho > 0$,

(C3) *according to Proposition 4.8:*

- (a) *There is $p_0 > 0$ such that $p(\rho) = -\frac{p_0}{\rho} + o\left(\frac{1}{\rho}\right)$ for $\rho \rightarrow 0$ or*
- (b) *one of the following conditions is fulfilled:*
 - (i) $\lim_{\rho \rightarrow 0} c(\rho) = 0$ and $2p'(\rho) - \rho p''(\rho) \geq 0$ for all $\rho > 0$,
 - (ii) $0 < \lim_{\rho \rightarrow 0} c(\rho) < \infty$,
 - (iii) *there is a $\eta \in (0, 1)$ such that $\lim_{\rho \rightarrow 0} \rho^\eta c(\rho)$ exists and $0 < \lim_{\rho \rightarrow 0} \rho^\eta c(\rho) < \infty$.*

These conditions are positively linear in p and hence all pressure functions satisfying them define a convex cone.

Note that the convex cone property means that also integrals over valid pressure functions are valid pressure functions, as the integral as an operator is linear and monotone.

With Proposition C we have a useful list of conditions which can be checked easily for any given candidate for a pressure function. The proposition makes it easy to prove our central result Proposition 4.2 and we will do so now.

Proof of Proposition 4.2.

- For the first part of the claim we note that for $-2 < \delta \leq 2$ and $\alpha > 0$ the generalized γ -law fulfills the conditions in Proposition C. Note that at $\delta = 0$ the conditions in C3 switch and at $\delta = -1$ the conditions in C2 switch.

- For the second part of the claim we observe: For $\delta > 2$, choose $-c(\rho_l) < \frac{q_l}{\rho_l} < -\frac{2}{\delta}c(\rho_l)$, which is obviously sub-sonic. Then L_l is strictly negative for $\rho > 0$, as is easily computed. A similar range for $\frac{q_r}{\rho_r}$ shows the same for $-L_r$. For $\delta < -2$, choose $\sqrt{\frac{-1}{\delta+1}}c(\rho_l) < \frac{q_l}{\rho_l} < c(\rho_l)$ and L_l is strictly positive for $\rho > 0$ and similarly for $-L_r$. \square

Example 4.11. As just proved, the functions

$$\begin{aligned} p(\rho) &= \rho^3, \\ p(\rho) &= \rho, \\ p(\rho) &= -\frac{1}{1-\epsilon} \rho^{-1+\epsilon} \quad (\text{for } 0 < \epsilon < 1), \\ p(\rho) &= \ln(\rho) \end{aligned}$$

and especially

$$p(\rho) = \frac{a\rho}{b+c\rho}, \quad \text{with } a, b, c > 0 \quad (4.15)$$

are all valid pressure functions. A rather exotic pressure law

$$p(\rho) = \frac{\rho^3 - \rho}{\ln(\rho)} = \int_1^3 \rho^\gamma d\gamma$$

is also valid.

The pressure law (4.15) will reappear in Chapter 5, where we will use it to simulate realistic gas networks.

4.1.2 A note on the vacuum

Up to now we have ignored the possibility of a vacuum state, that is, of a state with $\rho = 0$, and just excluded it from consideration by demanding $\rho > 0$. Further examination is in order.

With regards to a vacuum state, all examined pressure laws fall in one of two categories. On the one hand, for pressure laws that fulfill the bound

$$\int_0^\rho \frac{c(s)}{s} ds < \infty, \quad (4.16)$$

a discussion of vacuum states can be found in [LS80]. In the language of Proposition 4.2, these are exactly the well-known γ -laws with $1 < \gamma < 3$.

On the other hand, for all new pressure functions of Proposition 4.2, as well as the isothermal γ -law, i.e. $\gamma = 1$, the bound (4.16) is violated and —as shown in [LS80]— the vacuum cannot be connected to a non-vacuum state.

Therefore, for all those of the examined pressure laws that diverge at $\rho = 0$, no vacuum can appear as a solution of any Riemann problem, if the initial conditions don't contain a vacuum.

4.1.3 Extension to junctions

We have shown well-posedness of Riemann problems with sub-sonic initial conditions under certain conditions on the pressure function.

The next step is the analysis of the coupling conditions (4.5) and (4.6). We will carry it out by considering the generalized Riemann problem at the junction in accordance with [BHKo6a; BHKo6b; CGo8].

Consider a junction v with incoming pipes indexed by $s \in \mathcal{E}_v^s$ and outgoing pipes indexed by $f \in \mathcal{E}_v^f$ as in Section 2.1.6. At the junction-facing end of each pipe there are initial states $U_i = (\rho_i, q_i)$ for each $i \in \mathcal{E}_v = \mathcal{E}_v^s \cup \mathcal{E}_v^f$. To make things tractable, we restrict the solution of the junction Riemann problem to be of this form: In each pipe $i \in \mathcal{E}_v$, there appears exactly one new state V_i next to the junction such that the V_i satisfy the junction conditions and are connected to their respective U_i by an admissible shock or a rarefaction wave. On incoming pipes these must be 1-waves, on outgoing pipes these must be 2-waves. A sketch of this is shown in Figure 4.2.

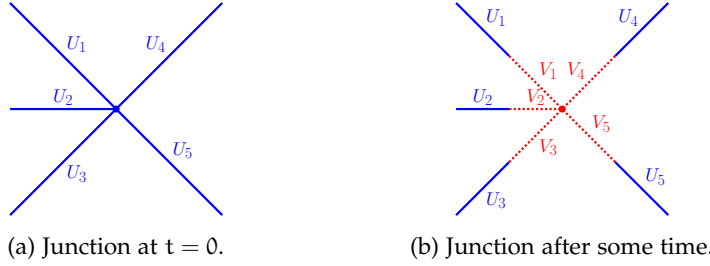


Figure 4.2: A junction with initial states. $i = 1, 2, 3$ are incoming pipes, $i = 4, 5$ are outgoing.

This is not a great restriction as this is the only solution structure found in pipelines with low Mach number, yet for different solutions, see [GHM17]. For these new states to appear at all, the wave speed between V_i and U_i must be negative on incoming pipes and positive on outgoing ones. To keep things simple, we examine a single incoming pipe i with initial condition U_i and new state $V_i = (\rho, L_1(\rho; U_i))$.

Proposition 4.12. *Let Inequality (4.4) and Proposition B and hence Proposition A be fulfilled for the pressure law. Let $\rho_{i,min}$ be such that $L'_1(\rho_{i,min}; U_i) = 0$, if it exists, otherwise let $\rho_{i,min} = 0$. The wave speed of the wave between U_i and V_i is negative if and only if $\rho > \rho_{i,min}$.*

Proof. By definition there is always a $\rho_{i,min}$. We first show that it is unique. It is easily seen that for $\rho \leq \rho_i$ there holds $\lambda_1(\rho, L_1(\rho; U_i)) = L'_1(\rho; U_i)$. Therefore there holds $L'_1(\rho_i, U_i) = \lambda_1(\rho_i, q_i) < 0$. With the concavity of L_1 we see that there is no $\rho_{i,min} > \rho_i$. Because of Inequality (4.4), L_1 is strictly concave for $\rho < \rho_i$ and this implies that $\rho_{i,min}$ is unique. Because L_1 is concave, L'_1 is decreasing, so $\lambda_1(\rho, L_1(\rho; U_i)) < 0$ if and only if $\rho > \rho_{i,min}$.

We now come to the wave speeds.

\Leftarrow : Let $\rho > \rho_{i,\min}$. Then $\lambda_1(V_i) < 0$. If V_i is in the rarefaction part of $L_1(\rho; U_i)$, the wave speed is negative. If V_i is in the shock part of $L_1(\rho; U_i)$, we compute:

$$\begin{aligned} s'(\rho) &= \left(\frac{L_1(\rho; U_i) - L(\rho_i; U_i)}{\rho - \rho_i} \right)' \\ &= \frac{L'_1(\rho; U_i)(\rho - \rho_i) - (L_1(\rho; U_i) - L_1(\rho_i; U_i))}{(\rho - \rho_i)^2} \leq 0, \end{aligned}$$

since L_1 is concave. So

$$s(\rho) \leq s(\rho_i) = \lambda_1(\rho_i) < 0,$$

because U_i is sub-sonic.

\Rightarrow : Let $\rho \leq \rho_{i,\min}$. Because U_i is sub-sonic, we have $\rho \leq \rho_{i,\min} < \rho_i$ and are dealing with a rarefaction wave. But then the wave speed is just given by $\lambda_1(\rho, L_1(\rho; U_i))$, which is non-negative in this case as shown in the beginning of the proof. \square

For outgoing pipes one defines $\rho_{i,\min} \geq 0$, such that $L'_r(\rho_{i,\min}; U_i) = 0$, if possible, or $\rho_{i,\min} = 0$ otherwise, and obtains in the same way the following proposition.

Proposition 4.13. *Let Proposition B be fulfilled for the pressure. On an outgoing pipe the wave speed between U_i and $V_i = (\rho, L_r(\rho; U_i))$ is positive if and only if $\rho > \rho_{i,\min}$.*

Coming back to junctions, we define the *minimal junction density* as

$$\rho_{\min} = \max_{i \in \mathcal{E}_v^s \cup \mathcal{E}_v^f} \rho_{i,\min}.$$

As just proved, for a solution ρ of the equation

$$\sum_{i \in \mathcal{E}_v^s} L_1(\rho; U_i) - \sum_{j \in \mathcal{E}_v^f} L_r(\rho; U_j) = 0 \quad (4.17)$$

the states $(V_i = (\rho, L_1(\rho; U_i)))_{i \in \mathcal{E}_v^s}$ fulfill $\lambda_1(V_i) < 0$ and the states $(V_j = (\rho, L_r(\rho; U_j)))_{j \in \mathcal{E}_v^f}$ fulfill $\lambda_2(V_j) > 0$ if and only if there holds $\rho > \rho_{\min}$.

Therefore a solution $(V_i)_{i \in \mathcal{E}_v}$ to the junction Riemann problem is admissible if and only if the density ρ at the junction fulfills $\rho > \rho_{\min}$.

Note that a usual Riemann problem with sub-sonic initial conditions can be treated as a junction with one incoming and one outgoing pipe. In this case, only one new state V is created and the admissibility criterion guarantees that $\lambda_1(V) < 0 < \lambda_2(V)$. For the sake of completeness, we classify the solutions to the usual Riemann problem by wave types in Table 4.1.

4.1.4 Additional constraints for consistency

Up to now, we have given conditions under which the Riemann problem at the junction is well-posed for sub-sonic initial conditions. But

Table 4.1: (r)arefaction waves and (s)hocks for different values of the density ρ .

(a) wave types for $\rho_l \leq \rho_r$.		
value of ρ	left wave	right wave
$\rho \leq \rho_{\min}$	invalid	
$\rho_{\min} < \rho \leq \rho_l$	r	r
$\rho_l \leq \rho \leq \rho_r$	s	r
$\rho_r \leq \rho$	s	s
(b) wave types for $\rho_r \leq \rho_l$.		
value of ρ	left wave	right wave
$\rho \leq \rho_{\min}$	invalid	
$\rho_{\min} < \rho \leq \rho_r$	r	r
$\rho_r \leq \rho \leq \rho_l$	r	s
$\rho_l \leq \rho$	s	s

we have no guarantee that the appearing states $(V_k)_{k \in \mathcal{E}_v}$ are themselves sub-sonic, as Propositions 4.12 and 4.13 only restrict the sign of the wave speed between U_k and V_k . For example, it is still possible that on an incoming pipe i we have $\lambda_1(V_i) < \lambda_2(V_i) < 0$. While this may still be a valid junction solution, this is unsatisfactory for simulation purposes as this introduces super-sonic states. Therefore we introduce another restriction on the solution of the coupling condition, namely that there must hold

$$\begin{aligned} \lambda_2(V_i) &> 0, i \in \mathcal{E}_v^s \\ \lambda_1(V_i) &< 0, i \in \mathcal{E}_v^f, \end{aligned} \quad (4.18)$$

because only then can we guarantee that the states V_i are sub-sonic. The following example shows that without this further condition, Equation (4.18) may actually be violated.

Example 4.14. We consider a junction with three incoming and no outgoing pipes. As pressure law we use the isothermal pressure law, $p(\rho) = c^2 \rho$, and take $c = 1$. Then choosing the following values for $U_i = (\rho_i, q_i)$, $i = 1, 2, 3$,

ρ_1	ρ_2	ρ_3	q_1	q_2	q_3
3×10^{-3}	1.5×10^{-3}	3×10^{-4}	$0.3\rho_1$	$0.3\rho_2$	$0.3\rho_3$

we find as the solution of Equation (4.17) $\rho \approx 1.25 \times 10^{-3}$. For the eigenvalues we get

$$\begin{aligned} \lambda_1(\rho, L_1(\rho, U_1)) &< 0 < \lambda_2(\rho, L_1(\rho, U_1)) \\ \lambda_1(\rho, L_1(\rho, U_2)) &< 0 < \lambda_2(\rho, L_1(\rho, U_2)), \end{aligned}$$

but

$$\lambda_1(\rho, L_1(\rho, U_3)) < \lambda_2(\rho, L_1(\rho, U_3)) < 0,$$

while U_1, U_2, U_3 are all sub-sonic. This means that the state V_3 cannot stay at the junction to fulfill the coupling condition.

We can define a range of densities ρ at the junction, inside of which corresponding states V_k are guaranteed to be sub-sonic. Consider again an incoming pipe i with initial condition U_i . We define

$$\rho_{i,\max} = \min(\{\rho > 0 \mid \lambda_2(\rho, L_l(\rho; U_i)) < 0\} \cup \{\infty\})$$

for incoming pipes and

$$\rho_{i,\max} = \min(\{\rho > 0 \mid \lambda_1(\rho, L_r(\rho; U_i)) > 0\} \cup \{\infty\})$$

for outgoing pipes and further the *maximal junction density*

$$\rho_{\max} = \min_{i \in \mathcal{E}_v^s \cup \mathcal{E}_v^f} \rho_{i,\max}. \quad (4.19)$$

Proposition 4.15. *A collection $(V_k)_{k \in \mathcal{E}_v}$ of states at a junction with sub-sonic initial conditions U_k in each pipe, that fulfills the coupling condition (4.17) is a valid solution of the junction problem and all V_k are sub-sonic, if the minimal junction density is strictly less than the maximal junction density and if the common density ρ of the V_k lies in between these,*

$$\rho_{\min} < \rho < \rho_{\max}.$$

Although there is no knowledge about $\lambda_2(V_i)$ on incoming pipes (and $\lambda_1(V_j)$ on outgoing ones), as it is not conveniently given by the derivative of a Lax curve, we still have $\lambda_2(V_i) = \lambda_1(V_i) + 2c(\rho) > \lambda_1(V_i)$, and especially

$$\lambda_2(\rho, L_l(\rho; U_i)) > \lambda_1(\rho, L_l(\rho; U_i)) \geq 0 \quad \text{for } \rho \leq \rho_{i,\min}.$$

Therefore we have at least $\rho_{i,\max} > \rho_{i,\min}$ in every pipe. Unfortunately there is no guarantee that $\rho_{\max} > \rho_{\min}$.

4.2 COUPLING TO AN EXTERNAL GAS SINK

In this section, we focus on the coupling of the gas network to an external gas consumer. The role of this consumer will be taken by gas-fired power plants in the next chapter. Following the ideas in [HMS19; Zlo+16], this external coupling is done by extending the already existing coupling, namely the conservation of fluxes in Equation (4.6), by an additional flux demand.

Well-posedness of the coupling

We focus on a junction with one incoming pipeline, one outgoing pipeline and an outlet that draws a set amount of flow ε . The coupling condition is taken from [HMS19] and reads

$$\begin{aligned} p_{\text{in}} &= p_{\text{out}}, \\ q_{\text{in}} &= q_{\text{out}} + \varepsilon, \end{aligned} \quad (4.20)$$

where the outflow ε is non-negative. To find what ε are allowed, we examine another generalized Riemann problem at the junction. We

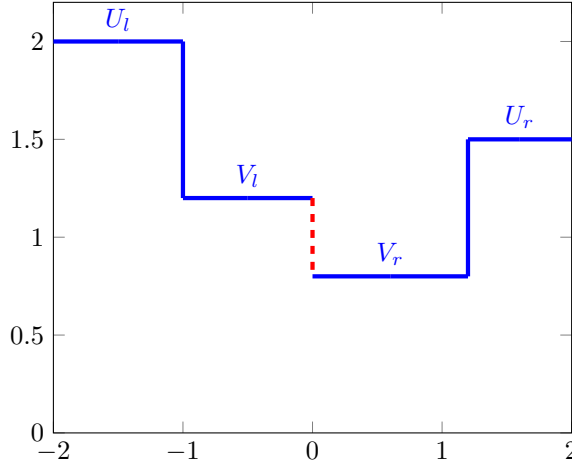


Figure 4.3: Solution after a short time, the junction is located at $x = 0$.

start with two constant states on the incoming and outgoing pipes and solve the Riemann problem where the flow q jumps by an amount ε at the junction. We again demand the waves to leave the junction, giving rise to a picture like that in Figure 4.3, where next to the junction in red two new states V_l, V_r appear that fulfill the coupling conditions (4.20). The analysis of this setting is similar to a three-way junction but with the flow on one outgoing pipe fixed to ε . We must now solve the equation

$$L_l(\rho) - L_r(\rho) = \varepsilon, \quad (4.21)$$

which is similar to Equation (4.9) of a usual (two-way) Riemann problem but with a non-zero right-hand side. For now we can forego the additional consistency constraints of Section 4.1.4: One half of the eigenvalues is on the right side of zero due to the ρ_{\min} -criterion, which we need here again to make the waves have the right direction:

$$\lambda_1(V_l) \leq 0 \quad \text{and} \quad \lambda_2(V_r) \geq 0.$$

Instead of invoking ρ_{\max} , we note that $\varepsilon \geq 0$ and compute for the remaining two eigenvalues

$$\lambda_2(V_l) = \frac{L_l(\rho)}{\rho} + c(\rho) = \frac{L_r(\rho) + \varepsilon}{\rho} + c(\rho) = \lambda_2(V_r) + \frac{\varepsilon}{\rho} \geq \lambda_2(V_r) \geq 0,$$

and

$$\begin{aligned} \lambda_1(V_r) &= \frac{L_r(\rho)}{\rho} - c(\rho) = \frac{L_l(\rho) - \varepsilon}{\rho} - c(\rho) \\ &= \lambda_1(V_l) - \frac{\varepsilon}{\rho} \leq \lambda_1(V_l) \leq 0. \end{aligned}$$

In our setting this is sufficient. When we also consider power-to-gas plants in Chapter 5, we will need to also take negative ε into account. In this case, we have to adhere to the ρ_{\max} -criterion (4.19) instead.

Solution structure for different outflows ε

For the usual Riemann problem we had a unique non-zero solution due to our findings in Section 4.1.1. For $\varepsilon > 0$ we now have two

solutions to $L_l(\rho) - L_r(\rho) = \varepsilon$, one of which is not admissible as it lies to the left of the maximum and hence has non-negative derivative (which means it would be super-sonic). As $L_l - L_r$ is decreasing in the admissible regime, greater ε result in smaller ρ . One of the two solution structures is given in Table 4.2. The other, for $\rho_r \leq \rho_l$ is similar.

Table 4.2: (r)arefaction waves and (s)hocks in the solutions for $\rho_l \leq \rho_r$.

value of ε	left wave	right wave
$\varepsilon \geq (L_l - L_r)(\rho_{\min})$	invalid	
$(L_l - L_r)(\rho_{\min}) > \varepsilon \geq (L_l - L_r)(\rho_l)$	r	r
$(L_l - L_r)(\rho_l) \geq \varepsilon \geq (L_l - L_r)(\rho_r)$	s	r
$(L_l - L_r)(\rho_r) \geq \varepsilon$	s	s

This concludes the theoretical study of the gas-power coupling (4.20). We now validate our findings with some simulation examples.

4.3 NUMERICAL RESULTS

Within the following numerical examples, we consider two different discretization schemes. The first is the third-order CWENO3 scheme (see Section 2.2.4 and [Kol14]) with a local Lax-Friedrichs numerical flux. This scheme is used as a reference for validation, but is not used further on, because the (usual) CFL condition makes it numerically very expensive.

The coupling conditions of the CWENO3 scheme are handled by the technique described in [NKS18], mentioned also in Section 2.2.4, which retains the third-order accuracy of the CWENO3 method.

The second scheme is the implicit box scheme of Section 2.2.3. In order to fulfill its inverse CFL condition (Proposition 2.45), we need the gas flow to be not only sub-sonic (Equation (4.7)), but for the eigenvalues to be bounded away from zero. Luckily, for gas flow in pipeline networks, the application we are interested in, an estimate of $v(\rho) < 0.3c(\rho)$ or even less is fulfilled due to economical reasons, as a high velocity can only be the result of a high pressure drop, which is inefficient, as in that case, energy is lost in the expansion of the gas, instead of its movement.

The inverse CFL condition is beneficial for problems with large characteristic speeds whose solution is quasi-stationary. Along with the low flow speed, this is usually the case for daily operation tasks in gas networks and therefore motivates the choice of this scheme for our scenarios.

The first test example in Section 4.3.1 is supposed to demonstrate the different cases revealed in the analysis above (Section 4.2). Further, since the applied implicit box scheme does not explicitly make use of any Riemann solver, this scenario is also considered as a numerical validation of its applicability, where the CWENO3 scheme with Riemann solver at the junction serves as reference.

For the second example (Section 4.3.2) we demonstrate the differences resulting from various pressure functions, which are all covered by our theoretical results from Section 4.1.1.

4.3.1 Validation

We consider the isentropic Euler equations without source term, pressure law $p(\rho) = \kappa \rho^\gamma$ and parameters $\kappa = 1.0$, $\gamma = 1.4$, and a Riemann problem with left state $U_l = \begin{pmatrix} 4.0 & 1.0 \end{pmatrix}^T$ and right state $U_r = \begin{pmatrix} 3.0 & -1.0 \end{pmatrix}^T$ for the setting described in Section 4.2. Accordingly, we assume a gas demand ε at the coupling point of the two states (whose position is chosen at $x = 0$ for a better readability of the corresponding graphs). Then, from Table 4.2, we get the following solution structure:

- s-s solution for $\varepsilon \leq 0.57877$,
- r-s solution for $0.57877 \leq \varepsilon \leq 3.0594$,
- r-r solution for $3.0594 \leq \varepsilon$.

Further, one can easily compute $\rho_{1,\min} \approx 1.8819$, $\rho_{2,\min} \approx 1.5041$, and therewith $\rho_{\min} \approx 1.8819$ and the maximum gas demand $\varepsilon \approx 4.3892$. We will consider the numerical simulation of the described setting until time $t = 0.1$ for $\varepsilon \in \{0.25, 1.75, 3.25\}$ and the following discretization parameters:

- CWENO3: $\Delta t = 5 \cdot 10^{-5}$, $\Delta x = 5 \cdot 10^{-4}$,
- IBOX: $\Delta t = 5 \cdot 10^{-4}$, $\Delta x = 5 \cdot 10^{-5}$.

The different choices result from the (usual) CFL condition (similar to Proposition 2.40) the explicit CWENO3 scheme has to obey, in contrast to the inverse CFL condition of the IBOX scheme (see Proposition 2.45). Figures 4.4 to 4.6 show the computed densities at the final time. Both schemes show the correct solution structure (shock/rarefaction waves), where CWENO3 expectably achieves the sharper resolution.

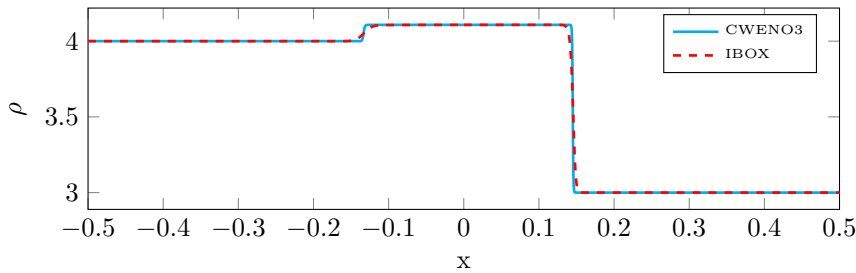


Figure 4.4: Density profile at $t = 0.1$ for $\varepsilon = 0.25$. (s-s solution)

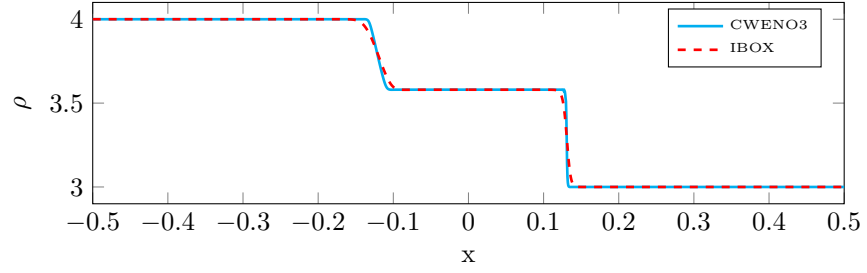


Figure 4.5: Density profile at $t = 0.1$ for $\varepsilon = 1.75$. (r-s solution)

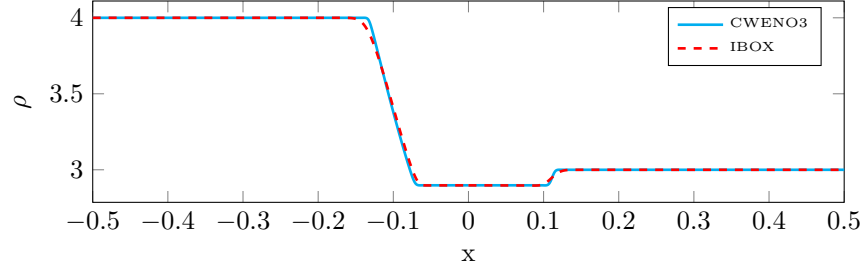


Figure 4.6: Density profile at $t = 0.1$ for $\varepsilon = 3.25$. (r-r solution)

4.3.2 Different pressure laws

In our second test case, we apply various pressure laws, which are all covered by our theoretical study, and are interested in the different dynamics one may observe even on a single pipeline. Therefore, we consider a single pipe with length $l = 0.1$ and the following pressure laws:

- $p(\rho) = \frac{1}{\gamma} \rho^\gamma$ with $\gamma = 1.4$ (“ γ -law”),
- $p(\rho) = \frac{-1}{0.99} \rho^{-0.99}$ (“inverse”, corresponding to $\gamma = -0.99$),
- $p(\rho) = \ln(\rho)$ (“logarithmic”),
- $p(\rho) = \frac{1}{10} \sum_{i=1}^{10} \frac{\rho^{1+i/5}}{1+i/5}$ (“sum of γ -laws”).

Apart from the usual γ -law, we also consider the “inverse” law as it lies close to the boundary of the additional pressure laws. The “logarithmic” law marks the transition from positive to negative γ and the sum is an example of the last property shown in Proposition C of sums of pressure functions again being pressure functions. It is peculiar that the “inverse” and “logarithmic” pressure functions are can be negative for positive densities. This may be counter-intuitive, yet, as only differences in pressure are relevant, the sign of the pressure function is of little consequence. Note that all considered pressure functions are scaled in such a way that $p'(\rho = 1) = 1$ and there is no source term. Initially, we have $\rho = 1$ and $q = 0$ in the whole pipe. Further, we fix $\rho = 1$ on the left-hand boundary, whereas q at the right-hand boundary linearly increases from 0 to 0.2 until time $t = 0.1$ and stays constant afterwards until the final time $t =$

0.5. We approximate the solution to this problem by CWENO3 with discretization parameters $\Delta t = 5 \cdot 10^{-4}$ and $\Delta x = 10^{-3}$. The variety of the resulting dynamics is demonstrated in Figures 4.7 and 4.8, which show the density in the pipeline at times $t = 0.25$ and $t = 0.5$, respectively.

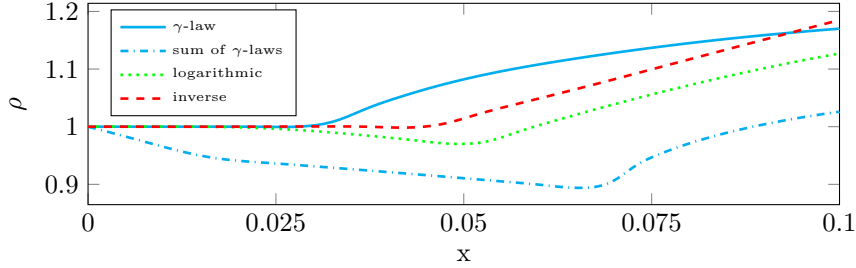


Figure 4.7: Simulation result at time $t = 0.25$ for different pressure laws.

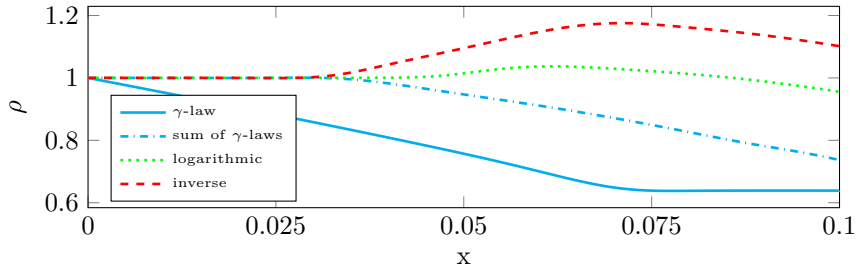


Figure 4.8: Simulation result at time $t = 0.5$ for different pressure laws.

4.4 SUMMARY

In this chapter we have presented a coupled model for gas and power allowing for a mathematically well-defined transition from gas to power. Along the way we have provided conditions on the pressure function of the isentropic Euler equations for the well-posedness of the Riemann problem. This led to some non-standard pressure functions. Various simulation results show the properties of the presented approach. In this context, we also validated that the implicit box scheme of [KLB10] is able to resolve the correct solution/wave structure of standard Riemann problems. This is of particular interest since the IBOX scheme allows for practically relevant discretization parameters and does not need any (even approximate) Riemann solver at junctions like the high resolution CWENO3 scheme we used for the validation.

Now that the theoretical ground work is done, we can move on to the simulation of realistic gas networks

APPLICATION TO COUPLED GAS-POWER NETWORKS

In the last two chapters we have established the power flow equations as a suitable model for power grids and the isentropic Euler equations with pressure coupling conditions for gas networks. In addition we have established the well-posedness of the gas-power coupling on the gas side in Section 4.2.

We will use all of these findings in this chapter to propose a combined gas-power model (Section 5.1) for a gas network coupled to a power grid via gas-fired power plants, which consume gas to generate power and also power-to-gas plants, which transform power into gas. This model will then be put to the test by combining well-known benchmark data for both gas networks (namely from GasLib [Sch+17]) and power grids (IEEE power networks from Matpower [ZMT11]) and simulating operation of the combined gas-power network in the rest of the chapter.

In doing so, we will also discuss vertex coupling conditions for the gas network (Section 5.3.1), that represent a physically more accurate model of gas flow, as already alluded to below Equation (4.5). Yet, ultimately, we will come to the conclusion that the additional effort, stemming mostly from lack of suitable network data, is not worth the minuscule changes in simulation results for the task of simulating realistic gas network settings.

Lastly we consider a slightly revised version of the combined network just mentioned to examine a scenario of volatile and uncertain power demand in Section 5.4. Here we will introduce volatility via the Ornstein–Uhlenbeck process into the power network and examine repercussions of this on the gas network coupled to the power network.

It is worth mentioning that for electrical grids and the corresponding optimization problems, there exists a large number of established benchmarks. This includes IEEE test cases, CIGRE benchmarks [Stro6] and, in case of the German system, the scigrid model [Mat+17]. Likewise for gas networks, the GasLib suite offers test cases for simulation and optimization purposes, respectively.

However, when it comes to coupled gas-power systems, there do not exist widely accepted benchmarks. One of the few exceptions appears to be the case study presented in [Zlo+17], which comprises the IEEE RTS96 One Area 24 node electrical grid and a 24 pipeline gas network. Hence, the present chapter takes steps towards a more realistic simulation benchmark for multi-energy systems. Specifically, we couple a model of the Greek gas network — the GasLib-134 model [Sch+17] which includes 86 pipelines — with the IEEE 300-bus system under AC conditions. We formulate a combined simulation framework, for which we observe a significant influence of the power consumption

on the gas pressure at the gas-fired plant nodes while the gas consumed varies over time. In addition we provide simulation data of the adapted version of the combined network, for which we investigate the influence of uncertainty. The data provided in two data repositories can be used to benchmark simulation software for gas-power networks.

5.1 THE COMBINED MODEL

For the combined model we start from a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, that contains two subgraphs $\mathcal{G}_{\text{Power}} = (\mathcal{V}_{\text{Power}}, \mathcal{E}_{\text{Power}})$ and $\mathcal{G}_{\text{Gas}} = (\mathcal{V}_{\text{Gas}}, \mathcal{E}_{\text{Gas}})$, such that $\mathcal{V} = \mathcal{V}_{\text{Power}} \cup \mathcal{V}_{\text{Gas}}$ and $\mathcal{E} = \mathcal{E}_{\text{Power}} \cup \mathcal{E}_{\text{Gas}}$. For the gas part, we choose the model from Section 4.1, for the power part we choose the power flow equations (3.27).

Up to now we have two separated models, the underlying graph is disconnected and the two connected components are the gas network and the power network respectively and these different parts behave quite differently. In the power network the edges just carry two parameters and their topological information, i.e. their starting and ending node and the nodes carry most of the physical information, namely active and reactive power, while the situation in the gas network is reversed. Here the arcs carry a balance law describing gas dynamics while the nodes only carry coupling information. Yet in all parts of the network, only the nodes have boundary conditions¹.

In order to connect gas and power networks, we introduce additional edges \mathcal{E}_{PG} , which model the conversion of fuel ε , taken from the gas network via Equation (4.20) into electrical power. Therefore a gas-fired power plant $e \in \mathcal{E}_{\text{PG}}$ must connect a node $v \in \mathcal{V}_{\text{Gas}}$ to a node $w \in \mathcal{V}_{\text{Power}}$, such that gas can be turned into power.

The exact dependence of the amount of burned fuel on the generated power is encoded in a *heat rate* formula. For example [Zl0+16] chooses a heat rate of the form

$$\varepsilon(P) = a_0 + a_1 P + a_2 P^2, \quad (5.1)$$

where P is the real power available at the node p and a_0, a_1 and a_2 are constants. For our first example in Section 5.2, we will do the same. Later on, in Section 5.3, we will use a linear relation as a simplification, because realistic values for the coefficients are hard to come by.

An illustration of a combined network can be found in Figure 5.1.

Before we show simulation results, we describe the model for all network components in detail, this will in part be a reprise of the last two chapters, but will introduce some subtleties that arise for realistic networks.

¹ In the power network these are rightfully called *node specifications* and we only call them boundary conditions to unify the wording between gas and power networks.

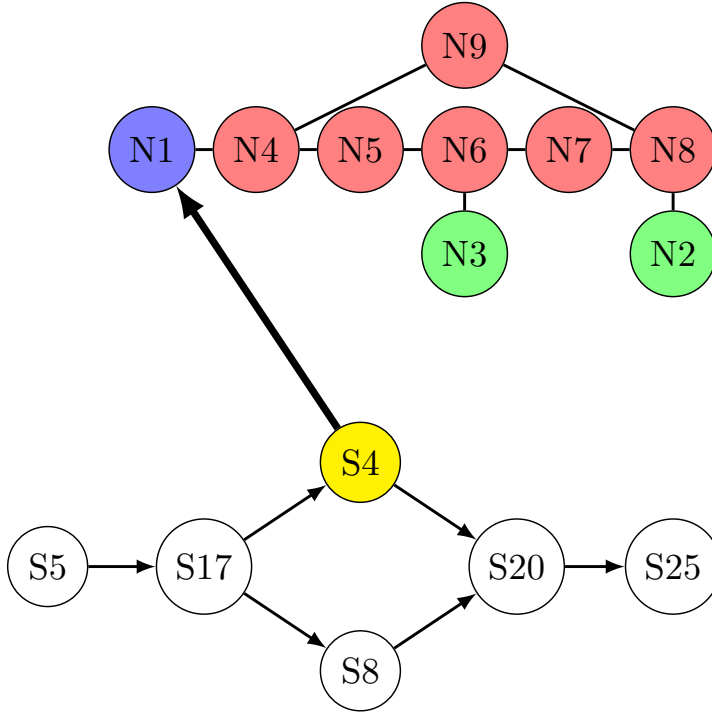


Figure 5.1: A small combined gas-power network. The upper right part is the power network “case9” from Matpower[ZMT11] with blue slack node, green power plants (generator nodes) and red consumers (load nodes). The lower part is a part of the GasLib-40 network ([Sch+17]). The bold arc is a gas-power connection.

5.2 A SIMPLE NETWORK

5.2.1 Model

For our first simulation scenario we introduce the bare minimum of models to get a meaningful coupled gas-power network. Let us start with the power model.

Modeling electric power flow on $\mathcal{G}_{\text{Power}}$

On $\mathcal{G}_{\text{Power}}$ we use the power flow equations (3.27). As stated, this model is suitable to describe the behavior of power networks operating at sinusoidal alternating current (AC). The power network consists of nodes and transmission lines. The nodes fall into two categories, generators and loads, where a generator models a power plant and a load models a consumer (or a collection thereof).

At a node $k \in \mathcal{V}_{\text{Power}}$ there are active (or real) power $P_k(t)$, reactive power $Q_k(t)$, the voltage magnitude $|V_k|(t)$ and the phase angle $\phi_k(t)$. These are now time dependent, because we are interested in scenarios where the power demand changes over time.

Further, we model the admittance of nodes and transmission lines, denoted by Y , which is split into real and imaginary part $Y = G + iB$. The admittance is the inverse of the impedance which in turn is a

complex extension of ohmic resistance in the power network. These quantities are assumed constant.

The admittance of a transmission line $e \in \mathcal{E}_{\text{Power}}$, connecting nodes $j, k \in \mathcal{V}_{\text{Power}}$ is denoted by $Y_{jk} = G_{jk} + iB_{jk}$, which we set to zero, if no arc connects j and k . The admittance of a node $k \in \mathcal{V}_{\text{Power}}$ is denoted by $Y_{kk} = G_{kk} + iB_{kk}$.

At every node at every point in time we prescribe two out of the four quantities $P_k, Q_k, |V|_k, \phi_k$, depending on the node:

- Load nodes k specify values for P_k, Q_k .
- Generators k specify values for P_k, V_k .
- Slack nodes (which are also generators) k specify V_k, ϕ_k .

These settings are usually called *bus specifications*, as the nodes in a power network are called *buses*. All in all we have $2|\mathcal{V}_{\text{Power}}|$ free variables at every time point, that can be determined as a solution to the power flow equations (3.27), which we restate here for the readers convenience:

$$\begin{aligned} P_k &= \sum_{j \in \mathcal{V}_{\text{Power}}} |V_k| |V_j| (G_{kj} \cos(\phi_k - \phi_j) + B_{kj} \sin(\phi_k - \phi_j)), \\ Q_k &= \sum_{j \in \mathcal{V}_{\text{Power}}} |V_k| |V_j| (G_{kj} \sin(\phi_k - \phi_j) - B_{kj} \cos(\phi_k - \phi_j)), \end{aligned} \quad (5.2)$$

where $P, Q, |V|$ and ϕ are time-dependent and these equations must hold at each point in time.

Note, that a necessary condition for uniqueness of the solution is to have at least one slack node k_0 , prescribing $|V_{k_0}|$ and ϕ_{k_0} , because otherwise for any solution $(P_k(t), Q_k(t), |V_k(t)|, \phi_k(t))_{k \in \mathcal{V}_{\text{Power}}}$ of the power flow equations, a different solution is obtained by shifting all phase angles ϕ_k by the same amount $r \in \mathbb{R}$. This is possible, as without a slack node, the equations only depend on differences in the phase angles. Often only a single slack node is used, although also multiple slack nodes can be used [Chao8] and we will do so in the second of our numerical studies in Section 5.3.4.

Instead of the described AC power flow equations, it is possible to use so-called direct current (DC) power flow equations, which are a linear approximation, see [GS01] for an overview. This approach simplifies the numerical treatment greatly at the cost of some accuracy. Other linearizations are subject of active research, see e.g. [LPL19].

Modeling gas flow on \mathcal{G}_{Gas}

We model the following quantities of the gas flow, namely, the pressure $p = p(x, t)$ as well as the flux $q = q(x, t)$. The units of those quantities are (bar) and ($\text{m}^3 \text{s}^{-1}$). Note that we use the volumetric flow as opposed to mass flow, as is customary in real-world gas networks. The pressure is given by a function of the gas density $\rho = \rho(x, t)$. An overview as well as recent results on gas flow can be found for example in [BGH11; BHK06a; BHK06b; Bre+14; CG08; Rei14; Rei15]

Pipelines

The direction of the edges in \mathcal{E}_{Gas} determines the positive direction of the flow. Pipelines $e \in \mathcal{E}_{\text{Gas}}$ in the gas network are modeled as an interval $[0, \ell_e]$ with further structure. In these the gas flow is modeled with the isentropic Euler equations from Section 4.1 as e.g. proposed in [BHK06b].

Here a subtlety arises. The density ρ and flux q , that are governed by the isentropic Euler equations as introduced in the last chapter have implicit units, namely $\rho = \rho_l$ is a line density with unit (kg m^{-1}) , while $q = q_m$ is a mass flow with unit (kg s^{-1}) .

For real-world scenarios we need to work with a three-dimensional density with unit (kg m^{-3}) and volumetric flow with unit $(\text{m}^3 \text{s}^{-1})$. These variables are given in terms of line density and mass flow as

$$\rho = \frac{\rho_l}{A}, \quad q = \frac{q_m}{\rho_0}, \quad (5.3)$$

where A is the cross section of the pipe in question and ρ_0 is the density of the gas at standard conditions. This is relevant for the coupling of pipes with possibly different cross sections. With these, the isentropic Euler equations (4.1) on every pipe $e \in \mathcal{E}_{\text{Gas}}$ acquire additional coefficients:

$$\begin{pmatrix} \rho_e \\ q_e \end{pmatrix}_t + \begin{pmatrix} \frac{\rho_0}{A_e} q_e \\ \frac{A_e}{\rho_0} p(\rho_e) + \frac{\rho_0}{A_e} \frac{q_e^2}{\rho_e} \end{pmatrix}_x = \begin{pmatrix} 0 \\ S(\rho_e, q_e) \end{pmatrix}. \quad (5.4)$$

Here p is the pressure function and S is the source term modeling wall friction in the pipes, both of which we prescribe presently. The source term S is given by

$$S(\rho, q) = \frac{\rho_0 \lambda(q_e) |q_e|}{2 A_e d_e \rho_e} (-q_e),$$

where d_e is the diameter of the pipe. The friction factor $\lambda(q_e)$ is the flux-dependent Darcy friction factor, as for example detailed in [Bro]. The friction is governed by the so-called Reynolds number,

$$\text{Re}(q_e) = \frac{d_e}{A_e \eta} \rho_0 |q_e|,$$

where η is the dynamic viscosity of the gas. For $\text{Re}(q_e) < 2000$ the friction is dominated by laminar flow and according to [Men15] we may assume

$$\lambda(q_e) = \frac{64}{\text{Re}(q_e)}. \quad (5.5)$$

For $\text{Re}(q_e) > 4000$ the friction is dominated by turbulent flow and must be determined by the Prandtl-Colebrook formula, see again [Men15]:

$$\frac{1}{\sqrt{\lambda}} = -2 \log_{10} \left(\frac{2.51}{\text{Re}(q_e) \sqrt{\lambda}} + \frac{k_e}{3.71 d_e} \right),$$

where now k_e is the roughness of the pipe. Instead of solving this implicit formula, we rely on the Swamee-Jain approximation [SJ76], namely,

$$\lambda(q_e) = \frac{1}{4} \frac{1}{\log_{10} \left(\frac{k_e}{3.71 d_e} + \frac{5.74}{\text{Re}(q_e)^{0.9}} \right)^2}. \quad (5.6)$$

For the intermediate regime $2000 \leq \text{Re}(q_e) \leq 4000$, the two expressions (5.5) and (5.6) are interpolated by the unique cubic polynomial, that makes the function λ differentiable at $\text{Re}(q_e) = 2000$ and $\text{Re}(q_e) = 4000$.

Now we come to the pressure function p . We use the isothermal pressure function with compressibility factor, as was also used for example in [Kol11],

$$p(\rho) = \frac{c_{\text{vac}}^2 \rho}{1 - \alpha c_{\text{vac}}^2 \rho}, \quad (5.7)$$

where α is an empirical correction coefficient and c_{vac} is the limit of the speed of sound in the vacuum limit, that is, for $\rho \rightarrow 0$. We also define the compressibility factor $z(p)$:

$$c_{\text{vac}} = \sqrt{\frac{p_0}{z_0} \frac{T}{T_0} \frac{1}{\rho_0}},$$

$$z(p) = 1 + \alpha p.$$

The numerical values for parameters $\rho_0, p_0, z_0, T_0, T, \alpha$ are listed in Table 5.1.

Table 5.1: Gas net constants.

ρ_0 [kg m ⁻³]	p_0 [bar]	z_0	T_0 [K]	T [K]	α [bar ⁻¹]
0.785	1.01325	1.005	273.15	283.15	-0.00224

In this table we see that $\alpha < 0$. This means, that the pressure function is of the form

$$p(\rho) = \frac{a\rho}{b + c\rho}, \text{ with } a, b, c > 0,$$

which, according to Proposition 4.10 yields a well-posed Riemann problem for sub-sonic initial states, namely the pressure function fulfills Conditions C1, C2(b) and C3(b)(ii).

Note that the pressure function can be inverted, yielding

$$\frac{p}{c_{\text{vac}}^2 z(p)} = \rho. \quad (5.8)$$

Nodes of the gas network \mathcal{V}_{Gas}

The previous set of differential equations on the edges has to be accompanied by boundary conditions, as pipelines have a finite length $\ell_e < \infty$. Therefore the nodes $v \in \mathcal{V}_{\text{Gas}}$ between gas pipelines will prescribe coupling conditions.

For our first examples we will simply use the coupling conditions from Section 4.1, which we describe here in detail. Consider a vertex

$v \in \mathcal{V}_{\text{Gas}}$ with attached pipelines $e \in \mathcal{E}_v = \mathcal{E}_v^s \cup \mathcal{E}_v^f$ and the function (already introduced in Equation (3.9))

$$s : \mathcal{E}_v \rightarrow \{\pm 1\},$$

$$s(e) = \begin{cases} 1 & e \in \mathcal{E}_v^s \text{ (e starts at } v), \\ -1 & e \in \mathcal{E}_v^f \text{ (e ends at } v). \end{cases}$$

Also let $p_e(t), q_e(t)$ for $e \in \mathcal{E}_v$ denote the boundary values at time t of edge e in node v , that is, $p_e(t) = p_e(x = 0, t)$, if $s(e) = 1$ and $p_e(t) = p_e(x = \ell_e, t)$, if $s(e) = -1$ and similarly for q_e . Then the coupling and boundary conditions at node v read

$$\begin{aligned} p_e(t) &= p_f(t) \text{ for all } e, f \in \mathcal{E}_v, \\ q_v(t) &= \sum_{e \in \mathcal{E}_v} s(e) q_e(t) \text{ for all } e \in \mathcal{E}_v, \end{aligned} \quad (5.9)$$

where $q_v : \mathbb{R}_0^+ \rightarrow \mathbb{R}$ is a possibly time dependent deterministic external demand or supply function, which is prescribed at boundary nodes of the network to model external gas inflow or outflow. Boundary nodes, where gas is supplied to the network, that is, $q_v \geq 0$, are called *sources*, while boundary nodes that demand gas, $q_v \leq 0$, are called *sinks*. For inner nodes, we have $q_v = 0$. These are in total $|\mathcal{E}_v|$ equations at a node v at each point in time.

5.2.2 Discretizations

Having defined our model we now need to discretize it in order to search solutions numerically. This search will be carried out by Newton's method at each time step. Along the lines of Section 2.2 we discretize the simulation time horizon $[0, T]$ as $(t^n = n\Delta t)_{n=0, \dots, N}$ with $\Delta t = \frac{T}{N}$ for some $N \in \mathbb{N}$. Every gas pipeline $e \in \mathcal{E}_{\text{Gas}}$ of length ℓ_e is discretized into $(x_j^e = j\Delta x_e)_{j=0, \dots, J_e}$ with $\Delta x = \frac{\ell_e}{J_e}$.

Power flow discretization

For the power network, the discretization is straightforward: We simply evaluate the power flow equations (5.2) at each time point $n\Delta t$.

Gas pipeline discretization

For the pipeline discretization we follow Section 2.2 and replace the continuous values of pressure (or density, see Equation (5.8)) and flow with values at each discretization point:

$$\begin{aligned} (p_e)_j^n &= p(j\Delta x_e, n\Delta t) \\ (\rho_e)_j^n &= \rho(j\Delta x_e, n\Delta t) \\ (q_e)_j^n &= q(j\Delta x_e, n\Delta t) \end{aligned}$$

where $x = x_j := j\Delta x_e$, $0 \leq j \leq J_e$. The isentropic Euler equations are discretized with the implicit Box scheme [IBOX], Definition 2.44, which is due to Kolb et. al. [KLB10].

For a general hyperbolic balance law

$$u_t + f(u)_x = S(u)$$

with space discretization $(x_j = j\Delta x)_{0 \leq j \leq J}$ as above we have for the time step between t and $t^* = t + \Delta t$

$$\begin{aligned} \frac{u_j^* + u_{j-1}^*}{2} = & \frac{u_j + u_{j-1}}{2} \\ & - \frac{\Delta t}{\Delta x} \left(f(u_j^*) - f(u_{j-1}^*) \right) \\ & + \frac{\Delta t}{2} \left(S(u_j^*) + S(u_{j-1}^*) \right), \end{aligned} \quad (5.10)$$

where $u_j = u(x_j, t)$ and $u_j^* = u(x_j, t^*)$. In our case, u_j has two components, density and flux, and hence we get $2J_e$ equations on a pipeline for $2J_e + 2$ variables. Therefore for each pipeline we need an additional 2 equations for the possibility of a unique solution. In our case, as one eigenvalue of the isentropic Euler equations is negative and the other positive, we need one boundary condition at the start and one at the end of the pipeline, as per Proposition 2.34.

Note that no diagonalization is needed before a time step and Equation (5.10) can be used directly. But the inverse CFL condition (Proposition 2.45) must be fulfilled,

$$\Delta t \geq \frac{\Delta x}{2\lambda_{\min}}.$$

Here, λ_{\min} denotes the minimum of the set

$$\{|\lambda| \mid \lambda \text{ is an eigenvalue of } f'(u), u \in B\},$$

where B must contain all states that appear in the physical system in question. For our purposes of pipelines it is reasonable to use

$$B = \left\{ (\rho, q) \in \mathcal{S} \mid \left| \frac{v(\rho, q)}{c(\rho)} \right| < 0.2 \right\},$$

which means that states are far in the interior of the sub-sonic domain. Apparently the scheme breaks down for transonic flow, where $v(\rho, q)$ approaches $c(\rho)$.

We refer to [Kol11, Prop 4.2, following remark] for a proof of well-posedness of the scheme in the scalar case. A proof for the systems case is unfortunately still missing. The numerical study of the last chapter, Section 4.3.1, directly evaluates the performance of the scheme for the isentropic Euler equations.

Note that the inverse CFL condition is well-suited for the simulation of gas networks, as large time steps are desirable for numerical feasibility when simulating over large time horizons.

Gas node and gas-power discretization

As was the case in the power network, the node equations (5.9) can be evaluated at each time step $n\Delta t$.

Also for the gas-power conversion plant equations (5.1) no further challenges arise and they are also evaluated at every time step.

5.2.3 Network properties and numerical results

With the model and discretizations we have just stated, we are ready to examine the first example of a combined gas-power network.

We consider the network(s) depicted in Figure 5.1 on page 93, containing a power grid from the example “case9” of the Matlab programming suite Matpower [ZMT11] and a small part of the GasLib-40 network [Sch+17], extended with a gas-to-power generator between S4 and N1, providing the necessary power at the latter node.

We consider an increasing power demand within the power grid, that leads to an increasing fuel demand of a gas-fired power plant and further to a significant pressure drop in the gas network.

In the power network a per-unit system is used, such that power and voltage are measured in multiples of base values. The base power and voltage are 100 MW and 345 kV, respectively. For example, a power of 200 MW is given in the per-unit system by 2 p.u..

Some parameters of the gas network are equal among all pipes in this small network, namely

- diameter: $d = 0.6$ m,
- roughness: $k = 0.05$ mm,
- dynamical viscosity: $\eta = 10^{-5}$ kg/meter/s.

All other parameters of the gas network are gathered in Table 5.2, the parameters of the power network in Table 5.3. Within this scenario,

Table 5.2: Parameters of the gas network.

Pipe	From	To	Length [km]
P10	S4	S20	20.322
P20	S5	S17	20.635
P21	S17	S4	10.586
P22	S17	S8	10.452
P24	S8	S20	19.303
P25	S20	S25	66.037

we use the discretization step sizes of ($\Delta x = 1$ km, $\Delta t = 900$ s) and apply the same time step sizes for the power flow equations. The entire discretized system is solved time step by time step with Newton’s method.

Initially, the gas network is in a stationary state: The pressure at S5 is fixed at 60 bar, the outflow at S25 is $q = 100 \text{ m}^3 \text{ s}^{-1}$, and there is an additional gas consumption at S4 resulting from the gas-to-power transformation ($a_0 = 2$, $a_1 = 5$, $a_2 = 10$) due to the power demand at the slack bus N1. The initial (stationary) state of the power grid is determined by boundary conditions given in Table 5.4.

In the course of the simulation, the power and reactive power demand at N5 are linearly increased between $t = 1$ hour and $t = 1.5$ hours from 0.9 p.u. to 1.8 p.u. (reactive power from 0.3 p.u. to 0.6 p.u.),

Table 5.3: Parameters of the power grid in per-unit quantities.

(a) Nodes			(b) Transmission lines				
Node	G	B	Edge	From	To	G	B
N1	0.0000	-17.3611	TL14	N1	N4	0.0000	17.3611
N2	0.0000	-16.0000	TL45	N4	N5	-1.9422	10.5107
N3	0.0000	-17.0648	TL56	N5	N6	-1.2820	5.5882
N4	3.3074	-39.3089	TL36	N3	N6	0.0000	17.0648
N5	3.2242	-15.8409	TL67	N6	N7	-1.1551	9.7843
N6	2.4371	-32.1539	TL78	N7	N8	-1.6171	13.6980
N7	2.7722	-23.3032	TL82	N8	N2	0.0000	16.0000
N8	2.8047	-35.4456	TL89	N8	N9	-1.1876	5.9751
N9	2.5528	-17.3382	TL94	N9	N4	-1.3652	11.6041

Table 5.4: Initial boundary conditions of the power grid [p.u.].

Node	P	Q	V	ϕ
N1	-	-	1	0
N2	1.63	-	1	-
N3	0.85	-	1	-
N4	0	0	-	-
N5	-0.90	-0.30	-	-
N6	0	0	-	-
N7	-1	-0.35	-	-
N8	0	0	-	-
N9	-1.25	-0.50	-	-

see Figure 5.2a. Accordingly, the power demand at the slack bus N1 increases, see Figure 5.2b, and therewith the gas consumption at S4, which also results in an increase of the inflow at S5 (see Figure 5.3). Due to the increased flow values, the pressure in the gas network decreases, see Figure 5.4 for the pressure at the nodes S20 and S25.

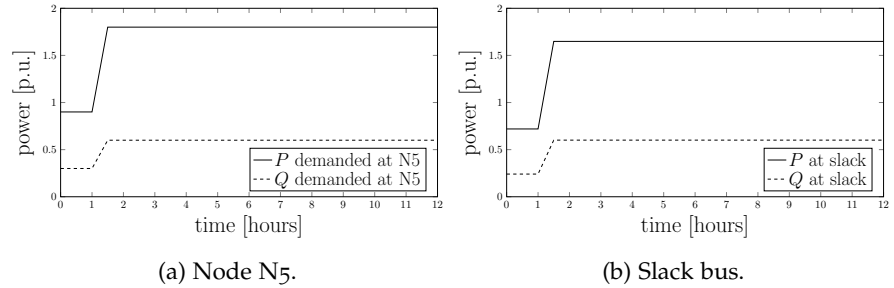


Figure 5.2: Power and reactive power at two example nodes.

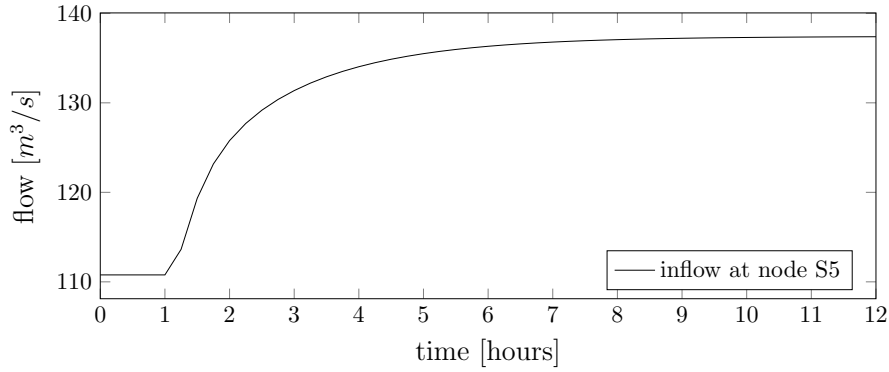


Figure 5.3: Inflow at node S5.

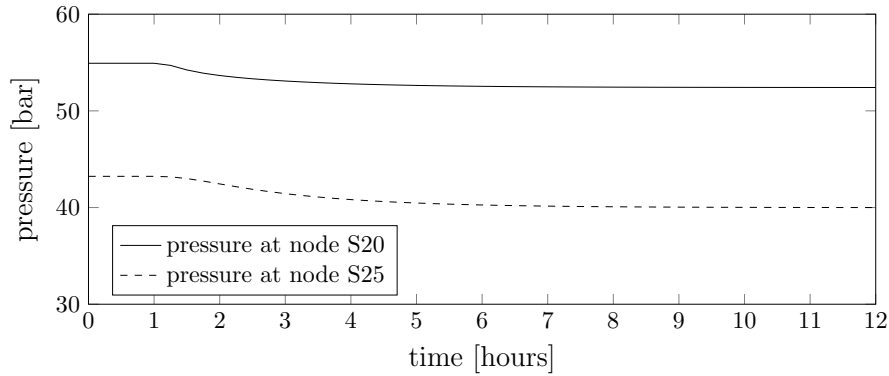


Figure 5.4: Pressure at nodes S20 and S25.

5.3 A REALISTIC NETWORK

5.3.1 A refined model for realistic networks

Having checked the validity of our approach on a small network, a “toy model”, it is time to examine our approach on realistic data. Therefore we employ larger networks and more gas-fired power plants. In addition gas-fired power plants are now accompanied by power-to-gas plants, that transform excess electrical power into gas for storage in the pipeline network. An overview on power-to-gas possibilities can be found in [Sch+15].

In order to model the physics of junctions more accurately, we will also examine coupling conditions that have been proposed in [Rei14; Rei15].

To simulate the operation of the network we have in mind, namely the GasLib-134 network, which is — as the name suggests — also found in [Sch+17], we must introduce some further models for network components that appear in this network. These will be applied in addition to those of Section 5.2.1.

Short pipes, control valves and compressors

In addition to pipelines there are three additional edge types in GasLib-134. As they behave similarly, we discuss them together. Let us start with short pipes.

In pipelines we employ a physical model, in our case the isentropic Euler equations, for the gas transport. In contrast, a short pipe has no physical properties and its model simply encodes that the state at its incoming end must be equal to the state at its outgoing end. That is, for a short pipe with incoming and outgoing pressures $p_{\text{in}}(t)$, $p_{\text{out}}(t)$ and fluxes $q_{\text{in}}(t)$, $q_{\text{out}}(t)$ there holds

$$\begin{aligned} p_{\text{out}}(t) &= p_{\text{in}}(t), \\ q_{\text{out}}(t) &= q_{\text{in}}(t). \end{aligned} \tag{5.11}$$

Short pipes can be used to separate boundary conditions from coupling conditions on the computational level by inserting an artificial short pipe in between a node with multiple attached pipes and an outer node of the network, see also [Kol11].

Next we come to control valves, which were also described in [ES05; Kol11]. These are similar to short pipes but can be controlled to reduce the pressure. Therefore they carry an additional control variable $u(t)$, that is externally prescribed. The relevant equations are

$$\begin{aligned} p_{\text{out}}(t) &= p_{\text{in}}(t) - u(t), \\ q_{\text{out}}(t) &= q_{\text{in}}(t). \end{aligned}$$

Note that in order to model the technical component correctly, the control must satisfy $u(t) \geq 0$ at all times.

Lastly there are compressor stations. These can be thought of as inverse control valves, in that they raise the pressure of the gas flowing through them. As this requires energy, this pressure increase incurs a cost. If the compressor to be modeled burns gas in order to produce the required energy, one needs to reduce the flux by the amount of burned fuel. Examples of this approach can be found in [ES05; Her07; Kol11]. On the other hand there are compressors that are powered from an external source and therefore a higher pressure increase should be more costly but not change the gas flow. We will follow this approach and arrive at the compressor equations

$$\begin{aligned} p_{\text{out}}(t) &= p_{\text{in}}(t) + u(t), \\ q_{\text{out}}(t) &= q_{\text{in}}(t), \end{aligned} \tag{5.12}$$

where $u(t)$ is again a control variable that must also be non-negative, $u(t) \geq 0$, to capture the behavior of real compressors. Later, in Chapter 6, we will introduce a cost associated with compressor operation and compute optimal controls.

Physical coupling conditions

We have already introduced coupling conditions in Equation (5.9), that enforce the continuity of the pressure over a node. In addition, the source nodes and the sink nodes carry the boundary conditions describing inflow and outflow, respectively. For the next numerical study we will compare these pressure boundary conditions, indicated by the subscript “p”, to so called *Bernoulli invariant* coupling conditions, indicated by the subscript “b” and introduced in [Rei14; Rei15].

There are also further physical coupling conditions found in [HHW20], which we do not investigate, although it is plausible that our findings, especially the numerical study in Section 5.3.4, may be applicable to these as well.

We first reformulate the coupling conditions (5.9) with a new coupling function H , to unify the coupling description:

$$\begin{aligned} H(\rho_e, q_e) &= H(\rho_f, q_f) \text{ for all } e, f \in \mathcal{E}_v, \\ q_v(t) &= \sum_{e \in E} s(e) q_e(t) \text{ for all } e \in \mathcal{E}_v. \end{aligned}$$

This results in the pressure coupling conditions, if we set

$$H_p(\rho, q) = p(\rho), \quad (5.13)$$

while the new Bernoulli coupling condition is defined by

$$H_b(\rho, q) = \frac{1}{2} \left(\frac{\rho_0 q}{\rho A} \right)^2 + \int_{\rho_0}^{\rho} \frac{p'(\hat{\rho})}{\hat{\rho}} d\hat{\rho}. \quad (5.14)$$

Note that we use the space-dependent density, not the line density, as introduced in Equation (5.3). This is important, because the line density is usually discontinuous over a node, when traveling from one pipe through the node to another pipe. On the other hand, for the pressure coupling function H_p , the three-dimensional space-dependent density is continuous over the node.

The Bernoulli coupling condition has the advantage that it assures that no energy is produced in a junction. This is not guaranteed by the pressure coupling condition. For a derivation and an example of this, see [Rei14]. Because $\rho \mapsto p(\rho)$ is one-to-one, (5.14) can be written as

$$H_b(p, q) = \frac{1}{2} \left(\frac{\rho_0 q}{\rho(p) A} \right)^2 + \int_{p_0}^p \frac{1}{\rho(\hat{p})} d\hat{p}. \quad (5.15)$$

Note that $v = \frac{\rho_0 q}{\rho A}$ is simply the flow velocity of the gas. If we were to omit the first part of H_b , this would be equivalent to the usual condition of pressure equality.

If we rewrite the coupling condition with only pipes that start at the node, using Lemma 2.32, we arrive at

$$\Psi(u(l, t)) = \begin{pmatrix} H(\rho_1, q_1) - H(\rho_2, q_2) \\ \vdots \\ H(\rho_{n-1}, q_{n-1}) - H(\rho_n, q_n) \\ \sum_{i=1}^n q_i \end{pmatrix}.$$

Here n enumerates the pipes connected in the node under consideration and the dependence of all ρ_i, q_i on $x = l$ and t has been omitted for readability. Let us define for notational convenience

$$\begin{aligned} H_i &= H(\rho_i, q_i), \\ \lambda_i &= \lambda_2(\rho_i, q_i), \\ c_i &= \frac{\partial H_i}{\partial \rho_i} + \frac{\partial H_i}{\partial q_i} \lambda_i, \end{aligned}$$

where λ_2 denotes the second (and therefore positive) eigenvalue of the isentropic Euler equations, see Equation (4.3). With these, the matrix in Condition (2.18), needed for Proposition 2.34, is given by

$$M = \begin{pmatrix} c_1 & -c_2 & & & \\ & c_2 & -c_3 & & \\ & & \ddots & \ddots & \\ \lambda_1 & & \dots & & \lambda_n \end{pmatrix},$$

and its determinant is given by

$$\det(M) = \sum_{1 \leq i \leq n} \lambda_i \prod_{1 \leq k \leq n, k \neq i} c_i,$$

as induction over n and Laplace expansion of the first column shows. This is non-vanishing if $\lambda_i > 0$ and $c_i > 0$. The first of which is just the sub-sonic condition, the second is true, because

$$\begin{aligned} \frac{\partial H}{\partial \rho} &> 0, \\ \frac{\partial H}{\partial q} &> 0. \end{aligned}$$

This is immediate for H_p , because $p' > 0$ and for H_b it can be computed from Equation (5.14), using again the sub-sonic condition.

Coming back the Equation (5.15), the integral can be solved when $\rho(p)$ is inserted from Equation (5.8). This yields

$$H_b(p, q) = \frac{1}{2} \left(\frac{\rho_0 q}{\rho(p)A} \right)^2 + c_{\text{vac}}^2 \left[\ln \left(\frac{p}{p_0} \right) + \alpha(p - p_0) \right].$$

Actually H_b behaves differently from H_p only because of the first part, which involves the gas velocity

$$v(\rho, q) = \frac{\rho_0 q}{\rho(p)A}.$$

Note that this velocity in realistic pipelines is usually rather small compared to the speed of sound in the gas $c(\rho) = \sqrt{p'(\rho)}$ and this suggests, that the impact of the first part may be small compared to that of the second part. We will investigate this in our numerical studies in Section 5.3.4.

Finally, let us remark on a technical property of the Bernoulli coupling. Although H_b represents the more accurate physical model, it brings about implementation issues: At a node where in addition to pipelines a short pipe or any other connection type, as detailed in Equations (5.11) through (5.12), is attached, the term H_b cannot be easily evaluated, as this requires to know the cross section of the component, a quantity often not available for these components. This is of course an issue that stems mostly from the used data, in our case the GasLib library. If it turns out that the physical coupling is desirable, this problem could be addressed by supplying this physical data.

Gas-Power-Conversion revisited

In our second example we simplify the dependence of burned fuel on generated power, and instead of Equation (5.1), we choose a linear relation, which encodes the generation of power with a constant efficiency E_{GtP} , where the subscript indicates “gas-to-power”.

In contrast to our first example, whose model only dealt with gas consumption, we would now like to incorporate the opposite direction as well, namely the generation of gas from a power supply. Therefore the edges \mathcal{E}_{PG} shall now model gas-fired power plants with an additional power-to-gas plant, where surplus electric power is converted to natural gas, e.g. by electrolysis and methanation.. An overview on power-to-gas capabilities can be found in [Sch+15], while a description of a state-of-the art plant is found in [TA14]. The conversion of power into gas is also modeled as a process with constant efficiency, called E_{PtG} , where this time the subscript indicates “power-to-gas”.

There are three reasons, why we put power-to-gas plants next to gas-fired power plants in our model: First, existing gas-fired power plants evidently have access to both the power and the gas network. Second, from a numerical perspective it elegantly deals with the possibility of power demand at a gas-fired power plant becoming negative as the model in that case can simply switch to the gas generation. Third, as one aim of power-to-gas plants is the reduction of carbon emissions, it is promising to capture the carbon dioxide from the burning of fuel during power generation for use during the gas generation. In that way and provided enough average electrical power, the emission of carbon can be avoided altogether.

The model equations for these combined plants relate the outgoing flux q in the gas node to the power demand at the power node and are given by

$$q = E(\text{sign}(P))P, \quad (5.16)$$

where P is the power demand (positive) or supply (negative) of the connected power node, q is the outflow of the sink, and

$$\begin{aligned} E(1) &= E_{\text{GtP}} \\ E(-1) &= E_{\text{PtG}} \end{aligned}$$

determines the efficiency of the conversion processes dependent on whether power is demanded or supplied. This piecewise linear model serves as an approximation of the heat rate of a power plant, respectively the efficiency of a power-to-gas plant. To overcome the non-differentiability of (5.16) at $P = 0$, we employ an interpolating function S with small parameter $\epsilon > 0$ (which we choose in Section 5.3.3),

$$S(x, a, b, \epsilon) = x \left(\frac{1}{2}(a + b) - \frac{3}{4}(b - a)\frac{x}{\epsilon} + \frac{(b - a)}{4} \left(\frac{x}{\epsilon} \right)^3 \right).$$

It is the unique polynomial (in x) of degree 4 satisfying

$$\begin{aligned} S(0, a, b, \epsilon) &= 0, \\ S(\epsilon, a, b, \epsilon) &= a \cdot \epsilon, \\ S(-\epsilon, a, b, \epsilon) &= b \cdot (-\epsilon), \\ \frac{\partial S}{\partial x}(\epsilon, a, b, \epsilon) &= a, \\ \frac{\partial S}{\partial x}(-\epsilon, a, b, \epsilon) &= b. \end{aligned}$$

Using S we replace the conversion (5.16) by

$$q = \begin{cases} E_{PtG} \cdot P & \text{for } P < -\epsilon, \\ S(P, E_{GtP}, E_{PtG}, \epsilon) & \text{for } -\epsilon < P < \epsilon, \\ E_{GtP} \cdot P & \text{for } \epsilon < P, \end{cases} \quad (5.17)$$

which makes $P \mapsto q(P) \in C^1(\mathbb{R})$ and $q(0) = 0$, so that no gas is taken from or injected into the gas network if electrical power is neither drawn nor supplied. In this way, we don't have to account for switching times beforehand. Instead the switching happens automatically by virtue of this conversion function.

5.3.2 Discretization

The new models introduced in Section 5.3.1 can be discretized directly by evaluating them at each time step.

5.3.3 Network properties

We now describe the network of the first real example. It consists of a gas network that is modeled after a real Greek gas network with some distortions to protect trade secrets of the gas network operator and a benchmark power network. The data we describe here can be found in a git repository² for future use.

Gas network

We use the GasLib-134 model [Sch+17] with inactive compressor and inactive valve. This is a network with 90 sink nodes, 3 source nodes and 86 inner nodes and an image of it can be found in Figure 5.5. As connections, there are 86 pipes, 45 short pipes, one compressor and one valve. The pipes have a total length of approximately 1500 km. As valve and compressor are inactive, they just let gas flow through them. We let them act like short pipes with the following exception: Although GasLib-134 doesn't provide it, we attach a cross section to these components so they can partake in the Bernoulli coupling (5.14). The compressor begins at the end of a single pipe and the valve ends at a single pipe. Therefore we endow them with the cross section of their respective attached pipes. We do so in order to have the coupling reach through the entire network. Otherwise, there would be three

² <https://bitbucket.org/efokken/gas-power-benchmark>

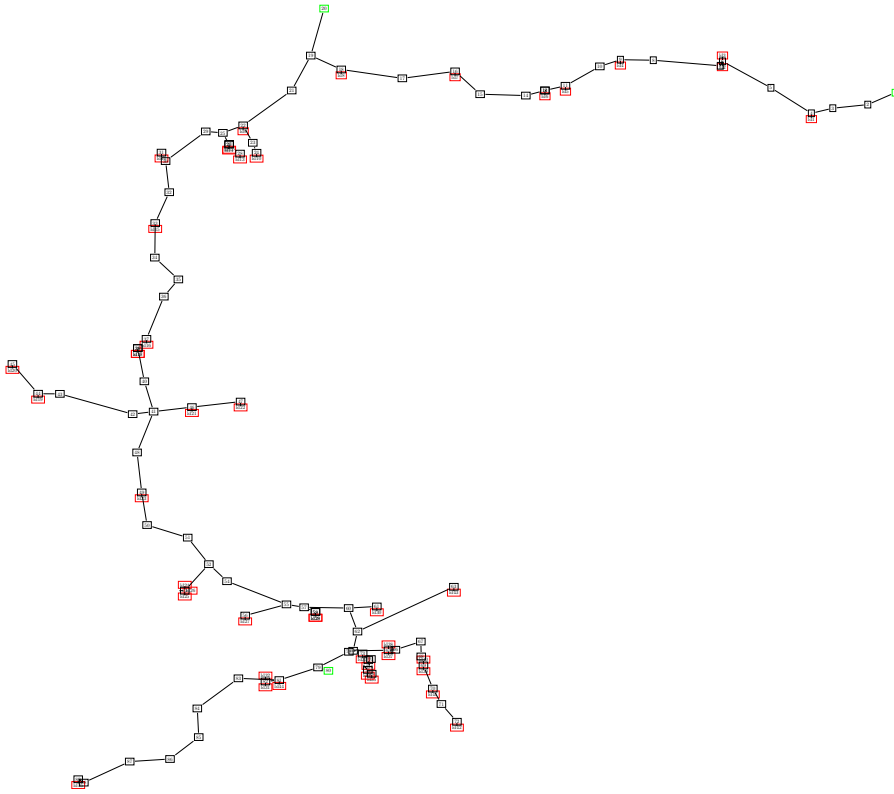


Figure 5.5: The gas network with green sources, red sinks and black junctions.

distinct parts, one before the compressor, one after the valve and one in between, that are not coupled through the Bernoulli coupling. The inflow of gas into the three source nodes and outflow at (non-gas plant) sink nodes of the network is chosen constant. A list of both can be found in Table 5.5.

Power network

For the power model we adapt the IEEE 300-bus test case that is part of the Matpower software [ZMT11]. A possible depiction of this network is shown in Figure 5.6. Originally, this system has a total of $n_{\text{bus}} = 300$ nodes (1 slack bus, 68 PV-nodes, 231 PQ-nodes), and $n_{\text{line}} = 411$ lines. We modify the grid such that the original slack bus is now a PV-node, and the nodes listed in Table 5.6 are all slack nodes. These are linked to sinks of the gas network. At these nodes, gas and electricity can be converted into each other. As GasLib-134 is inspired by the Greek gas network we aim to have a plausible number of gas-fired power plants for Greece. Unfortunately the author was unable to find reliable data on the total number of gas-fired power plants in Greece, although a choice of ten gas-fired power plants seems to be reasonable.³ The IDs of the connected sinks in the gas network are given in Table 5.6. Therefore we have a total of 10 slack nodes, 59 PV-nodes and 231 PQ-nodes. The nominal total active power generation of the grid is about

³ See e.g. https://de.wikipedia.org/wiki/Liste_von_Kraftwerken_in_Griechenland and https://en.wikipedia.org/wiki/List_of_power_stations_in_Greece

Table 5.5: Volumetric inflow and outflow at source and sink nodes of the gas network.

		(b) Volumetric outflow at sinks other than conversion plants.	
		Gas network ID	Outflow [$\text{m}^3 \text{s}^{-1}$]
(a) Volumetric inflow for gas network.		node_ld1	0.000000
		node_ld3	0.000000
		node_ld4	0.121019
		node_ld5	0.000000
		node_ld7	1.490446
		node_ld8	2.089172
		node_ld9	0.000000
		node_ld11	5.490446
		node_ld14	0.452229
		node_ld15	0.280255
		node_ld16	0.076433
		node_ld17	4.617834
		node_ld18	4.617834
		node_ld19	0.802548
		node_ld20	0.445860
		node_ld21	0.286624
		node_ld22	7.592357
		node_ld23	0.082803
		node_ld25	0.802548
		node_ld26	0.000000
		node_ld27	0.012739
		node_ld28	0.000000
		node_ld30	1.426752
		node_ld32	0.000000
		node_ld33	1.101911
		node_ld34	0.000000
		node_ld35	0.000000
		node_ld37	7.732484
		node_ld38	0.000000
		node_ld39	0.000000
		node_ld40	7.732484
		node_ld41	1.528662
		node_ld43	0.000000
		node_ld44	0.000000
		node_ld45	0.000000
Node ID	Inflow [$\text{m}^3 \text{s}^{-1}$]		
node_1	58.993631		
node_20	190.815287		
node_80	61.866242		

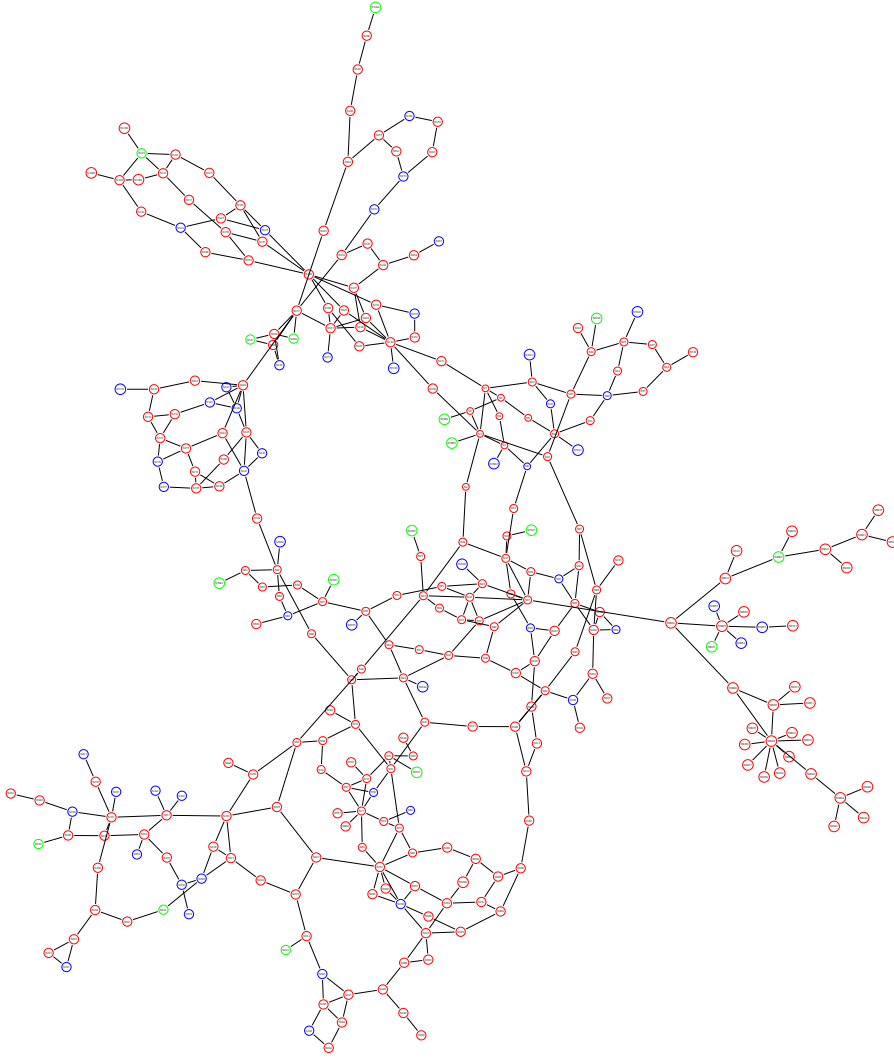


Figure 5.6: Power network with green gas-fired power plants, blue non-gas power plants and red loads.

Table 5.6: Connection of gas nodes and power nodes for conversion.

Power grid ID	Gas network ID
213	node_ld31
221	node_ld24
230	node_ld13
7001	node_ld36
7017	node_ld2
7024	node_ld12
7039	node_ld42
7057	node_ld6
7061	node_ld29
7071	node_ld10

24,000 MW. For the slack nodes and the PV-nodes, we use the bus specifications from the original case file “case300” in the Matpower suite. For PQ-nodes we use time-dependent bus specifications given by

$$\begin{aligned} P(t) &= P_{300} \left(0.9 + 0.4 \sin \left(\frac{2\pi t}{24h} \right) \right), \\ Q(t) &= Q_{300} \left(0.9 + 0.4 \sin \left(\frac{2\pi t}{24h} \right) \right), \end{aligned} \quad (5.18)$$

where P_{300} and Q_{300} are the active and reactive power demand from the original case file.

Parameters of Gas-Power conversion

We need to specify values for the conversion factors E_{GtP} and E_{PtG} in (5.16). For the operation as a gas-fired power plant we choose an efficiency of $\eta_{GtP} = 0.4$ with respect to the lower heating value of the gas. This is a realistic value, given that there are gas-fired power plants with efficiencies of up to 60 % [Age19]. The lower heating value L of natural gas is usually in the range of $36 \text{ MJ kg}^{-1} \leq L \leq 50 \text{ MJ kg}^{-1}$, depending on the gas composition [CL16]. We choose $L = 40 \text{ MJ kg}^{-1}$. The parameter E_{GtP} is then obtained from

$$E_{GtP} = \frac{1}{\rho_0 L \eta_{GtP}} \approx 0.0796 \text{ m}^3 \text{ MJ}^{-1}.$$

For PtG conversion we choose an efficiency of $\eta_{PtG} = 0.8$, this time with respect to the upper heating value according to [TA14]. The upper heating value U of natural gas is given by $U = 1.11L$. Therefore we obtain

$$E_{PtG} = \frac{\eta_{PtG}}{\rho_0 U} \approx 0.0229 \text{ m}^3 \text{ MJ}^{-1}.$$

Lastly for the parameter ϵ , which smooths Equation (5.16), we use $\epsilon = 0.01$, which was found by trial and error, chosen to yield reliable convergence of the Newton method while being as small as possible to keep the conversion factor close to the piecewise linear model. Note that the choice of ϵ must depend on the time step size, where smaller time steps allow for smaller ϵ and therefore more sudden switching behavior.

5.3.4 Numerical results

Using the network data from Section 5.3.3 we simulate the combined network over a time horizon of 24 hours. Therefore we use a time step size of $\Delta t = 30 \text{ min}$ and a pipeline-dependent spatial step size that is chosen near $\Delta x = 10 \text{ km}$, such that for every pipeline e there holds that $\frac{\ell_e}{\Delta x_e}$ is an integer.

Comparison of coupling conditions

In order to quantify the difference of the two coupling conditions H_p and H_b in Equations (5.13) and (5.14), we simulate once (simulation

Table 5.7: Absolute and relative differences for the two coupling constants.

(a) Absolute and relative difference of pressure.

$\max p_p - p_b $	$\max \frac{ p_p - p_b }{p_p}$
0.1828 bar	0.0041

(b) Absolute and relative difference of flow for different ranges near 0.

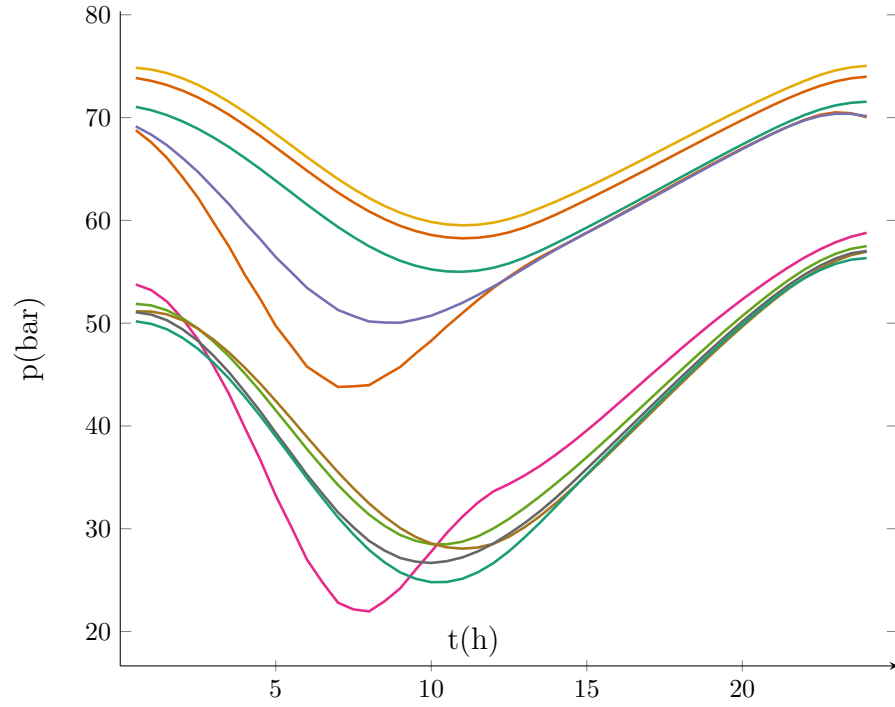
range [$\text{m}^3 \text{s}^{-1}$]	$\max q_p - q_b $	$\max \frac{ q_p - q_b }{ q_p }$
$10^{-3} < q_p < 10^{-2}$	$0.0004 \text{ m}^3 \text{s}^{-1}$	0.3270
$10^{-2} < q_p < 10^{-1}$	$0.0180 \text{ m}^3 \text{s}^{-1}$	0.32670
$10^{-1} < q_p < 10^0$	$0.0225 \text{ m}^3 \text{s}^{-1}$	0.1105
$10^0 < q_p < 10^1$	$0.0227 \text{ m}^3 \text{s}^{-1}$	0.0207
$10^1 < q_p $	$0.0570 \text{ m}^3 \text{s}^{-1}$	0.0029

“p”) with the pressure coupling constant H_p , and once (simulation “b”) with the Bernoulli coupling constant H_b . We compare the values of the pressures and the volumetric flows on the whole gas network at every time step. Let $p_p(x, t)$ be the pressure obtained from simulation “p” at time t and at some position in the network (x ranges over all pipes and all pipe lengths). Further, let $p_b(x, t)$ be the analogue for simulation “b” and let $q_p(x, t)$ and $q_b(x, t)$ be the corresponding values for the volumetric flow. Table 5.7 shows our findings with regard to the different coupling constants. For the relative differences in the flow we used different ranges of q , because although the relative error grows when approaching $q = 0$, the absolute values are very small and hence probably of little significance. In contrast to [Rei14] we find little difference for the two coupling conditions. The key difference of our model, compared to that of [Rei14], is the absence of a friction term in the latter, which allows errors to accumulate. In our case, artificial energy produced at the nodes is consumed by friction and cannot cause much error. In light of the small size of the error introduced by using the physically unsound pressure coupling constant, practitioners should trade-off carefully the need for more accuracy against the practical hurdles mentioned at the end of Section 5.3.1.

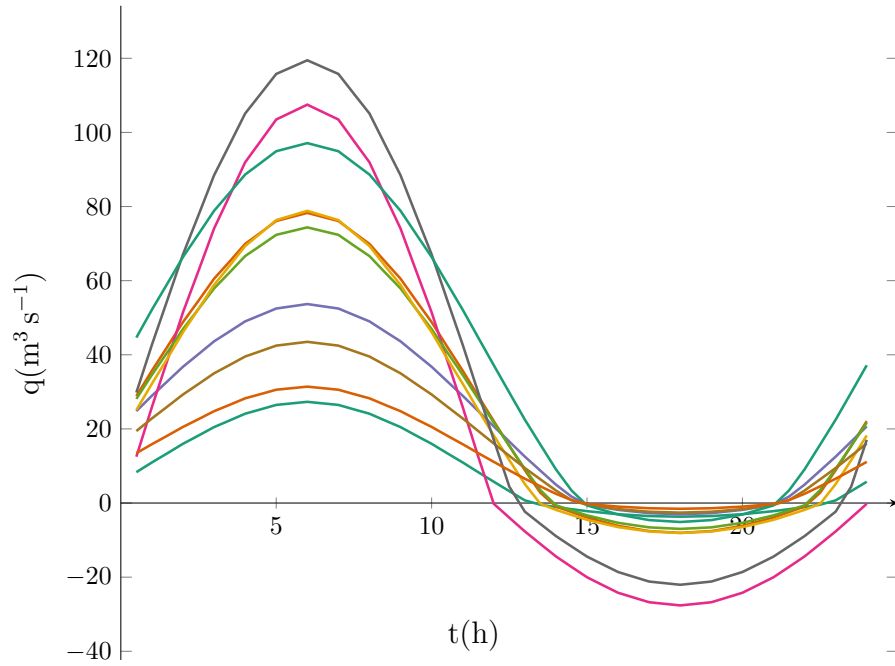
We, for instance, drop the Bernoulli coupling from now on. Besides the implementation issues, it is questionable whether the small errors observed could ever exceed inevitable modeling errors, that stem from, for example, not considering a pipeline as a three-dimensional object, but as a line.

Gas-power conversion

We now present the results of gas-to-power conversion and power-to-gas conversion. Because of the changing power demand (5.18) at



(a) Pressure at the gas-plant nodes over time.



(b) Gas flow for power conversion over time.

Figure 5.7: State variables at gas-plant nodes over time.

the loads, the gas-power conversion plants will go through a similar cycle. Over the course of a day, all of the plants go through a peak of power demand during which gas is consumed to power a generator. During the second half of the time horizon much less power is needed and so the gas-to-power mode is used to convert power back to gas. All the data of pressure and flow in the conversion plants is found in Table 5.8 and Table 5.9 on pages 114 and 115. The total volume of gas consumed by the power plants is obtained by integrating the

outflow over time: As it is only given ad discretization points, we use the trapezoidal rule for this task. In our case the amount of consumed gas is $2.3098 \times 10^7 \text{ m}^3$, the total volume of gas generated is $2.0522 \times 10^6 \text{ m}^3$.

Figure 5.7a shows the pressure evolution at those gas nodes attached to a conversion plant. It shows that the gas network cannot provide the peak power demand indefinitely as the pressure drops considerably during power generation (in the first 12 hours). But it is suitable to counter-balance high and low power demand over the course of a day as it recuperates during low power demand when gas is injected into the pipeline network by power-to-gas operation.

Figure 5.7b shows the amount of gas consumed ($q > 0$) by power generation and generated ($q < 0$) by power-to-gas operation. Note the kink at the switching times, where $q = 0$ which is due to the difference in efficiencies of the two processes.

Table 5.8: Pressure values at the gas-power conversion plants (time in hours, pressure in bar).

t	N7071	N7024	N230	N221	N7061	N7017	N213	N7001	N7039	N7057
0.5	71.05	68.77	69.15	53.77	51.88	74.85	51.15	51.06	50.18	73.85
1.0	70.72	67.58	68.34	53.21	51.72	74.67	51.13	50.83	49.93	73.57
1.5	70.24	66.09	67.32	52.11	51.25	74.33	50.85	50.29	49.41	73.17
2.0	69.63	64.22	66.07	50.44	50.48	73.82	50.29	49.41	48.58	72.63
2.5	68.91	62.20	64.71	48.41	49.47	73.18	49.47	48.27	47.50	71.96
3.0	68.06	59.80	63.14	45.84	48.21	72.40	48.40	46.83	46.15	71.16
3.5	67.13	57.44	61.57	43.09	46.77	71.53	47.14	45.21	44.61	70.27
4.0	66.09	54.74	59.81	39.86	45.12	70.55	45.69	43.35	42.85	69.27
4.5	65.01	52.36	58.18	36.72	43.38	69.51	44.12	41.43	40.99	68.22
5.0	63.86	49.73	56.42	33.22	41.51	68.40	42.44	39.34	38.99	67.10
5.5	62.71	47.80	54.96	30.18	39.66	67.29	40.71	37.33	36.98	65.98
6.0	61.54	45.78	53.45	27.02	37.76	66.16	38.94	35.26	34.93	64.84
6.5	60.43	44.76	52.35	24.79	35.98	65.08	37.21	33.41	32.99	63.76
7.0	59.36	43.79	51.29	22.81	34.26	64.02	35.51	31.61	31.11	62.71
7.5	58.38	43.86	50.69	22.15	32.74	63.06	33.93	30.13	29.45	61.75
8.0	57.48	43.97	50.16	21.96	31.37	62.17	32.47	28.81	27.94	60.87
8.5	56.72	44.86	50.06	22.95	30.27	61.40	31.18	27.87	26.73	60.12
9.0	56.07	45.73	50.04	24.20	29.39	60.74	30.07	27.15	25.76	59.47
9.5	55.58	47.06	50.35	26.02	28.82	60.23	29.20	26.80	25.13	58.96
10.0	55.23	48.29	50.73	27.77	28.50	59.84	28.56	26.67	24.80	58.58
10.5	55.04	49.69	51.33	29.59	28.49	59.61	28.19	26.84	24.82	58.35
11.0	55.00	50.98	51.98	31.17	28.73	59.52	28.06	27.21	25.14	58.25
11.5	55.12	52.26	52.75	32.55	29.26	59.58	28.17	27.80	25.76	58.31
12.0	55.39	53.43	53.57	33.63	30.03	59.78	28.54	28.56	26.66	58.50
12.5	55.81	54.52	54.44	34.36	30.96	60.13	29.22	29.51	27.81	58.84
13.0	56.37	55.52	55.35	35.18	32.02	60.61	30.14	30.55	29.15	59.31
13.5	57.04	56.42	56.26	36.12	33.16	61.20	31.24	31.71	30.60	59.89
14.0	57.76	57.21	57.15	37.19	34.38	61.83	32.48	33.00	32.15	60.56
14.5	58.52	58.00	57.98	38.33	35.64	62.49	33.82	34.39	33.72	61.27
15.0	59.30	58.81	58.78	39.56	36.96	63.20	35.23	35.83	35.29	62.01
15.5	60.09	59.63	59.57	40.83	38.32	63.93	36.67	37.29	36.81	62.75
16.0	60.89	60.46	60.38	42.15	39.71	64.68	38.13	38.77	38.32	63.51
16.5	61.69	61.29	61.20	43.47	41.11	65.44	39.60	40.25	39.82	64.29
17.0	62.51	62.13	62.03	44.81	42.53	66.22	41.09	41.73	41.33	65.07
17.5	63.33	62.96	62.86	46.12	43.94	67.00	42.57	43.19	42.81	65.86
18.0	64.15	63.80	63.69	47.44	45.34	67.78	44.04	44.65	44.29	66.65
18.5	64.97	64.62	64.52	48.70	46.72	68.55	45.49	46.07	45.73	67.44
19.0	65.79	65.43	65.34	49.96	48.09	69.33	46.92	47.48	47.15	68.23
19.5	66.59	66.22	66.14	51.15	49.41	70.08	48.32	48.83	48.52	69.00
20.0	67.39	67.00	66.94	52.31	50.70	70.83	49.67	50.15	49.85	69.76
20.5	68.16	67.75	67.70	53.41	51.93	71.56	50.97	51.41	51.12	70.51
21.0	68.91	68.49	68.46	54.48	53.12	72.26	52.22	52.62	52.34	71.24
21.5	69.61	69.16	69.15	55.46	54.22	72.94	53.37	53.73	53.43	71.92
22.0	70.25	69.78	69.73	56.38	55.22	73.57	54.41	54.73	54.38	72.54
22.5	70.79	70.28	70.15	57.19	56.07	74.15	55.30	55.59	55.16	73.09
23.0	71.19	70.51	70.37	57.87	56.75	74.60	56.05	56.29	55.77	73.52
23.5	71.44	70.42	70.36	58.41	57.23	74.89	56.62	56.80	56.18	73.82
24.0	71.55	70.03	70.14	58.79	57.50	75.04	56.96	57.02	56.33	73.98

Table 5.9: Flow values at the gas-power conversion plants (time in hours, in $\text{m}^3 \text{s}^{-1}$). Positive flow means gas is converted to power, negative flow means gas is generated with power.

t	N7071	N7024	N230	N221	N7061	N7017	N213	N7001	N7039	N7057
0.5	8.32	28.86	24.80	12.50	28.10	25.10	19.43	29.81	44.63	13.46
1.0	10.95	35.72	28.95	25.90	34.59	32.19	22.82	42.70	52.16	15.87
1.5	13.44	42.21	32.84	38.58	40.72	39.02	26.03	54.83	59.24	18.18
2.0	15.98	48.82	36.77	51.45	46.94	46.09	29.28	67.08	66.39	20.54
2.5	18.21	54.61	40.18	62.67	52.37	52.36	32.12	77.70	72.59	22.63
3.0	20.49	60.52	43.63	74.06	57.90	58.86	35.00	88.42	78.85	24.78
3.5	22.27	65.17	46.30	82.93	62.22	64.01	37.25	96.73	83.71	26.49
4.0	24.10	69.92	49.01	91.94	66.63	69.34	39.53	105.10	88.63	28.25
4.5	25.28	72.98	50.74	97.68	69.47	72.80	40.99	110.42	91.76	29.40
5.0	26.47	76.10	52.48	103.48	72.36	76.35	42.47	115.78	94.93	30.58
5.5	26.89	77.19	53.08	105.48	73.37	77.59	42.98	117.62	96.02	30.99
6.0	27.31	78.28	53.69	107.49	74.38	78.84	43.50	119.46	97.12	31.41
6.5	26.89	77.19	53.08	105.48	73.37	77.59	42.98	117.62	96.02	30.99
7.0	26.47	76.10	52.48	103.48	72.36	76.35	42.47	115.78	94.93	30.58
7.5	25.28	72.98	50.74	97.68	69.47	72.80	40.99	110.42	91.76	29.40
8.0	24.10	69.92	49.01	91.94	66.63	69.34	39.53	105.10	88.63	28.25
8.5	22.27	65.17	46.30	82.93	62.22	64.01	37.25	96.73	83.71	26.49
9.0	20.49	60.52	43.63	74.06	57.90	58.86	35.00	88.42	78.85	24.78
9.5	18.21	54.61	40.18	62.67	52.37	52.36	32.12	77.70	72.59	22.63
10.0	15.98	48.82	36.77	51.45	46.94	46.09	29.28	67.08	66.39	20.54
10.5	13.44	42.21	32.84	38.58	40.72	39.02	26.03	54.83	59.24	18.18
11.0	10.95	35.72	28.95	25.90	34.59	32.19	22.82	42.70	52.16	15.87
11.5	8.32	28.86	24.80	12.50	28.10	25.10	19.43	29.81	44.63	13.46
12.0	5.75	22.11	20.68	-0.20	21.71	18.22	16.06	17.05	37.18	11.11
12.5	3.21	15.44	16.59	-3.96	15.41	11.57	12.74	4.42	29.80	8.81
13.0	0.72	8.85	12.51	-7.66	9.19	5.11	9.44	-2.32	22.50	6.56
13.5	-0.45	2.76	8.73	-11.05	3.49	-0.21	6.39	-5.64	15.76	4.50
14.0	-1.10	-0.94	4.97	-14.40	-0.61	-1.84	3.37	-8.92	9.08	2.49
14.5	-1.65	-2.43	1.74	-17.23	-1.98	-3.21	0.79	-11.70	3.41	0.79
15.0	-2.19	-3.91	-0.42	-20.02	-3.33	-4.53	-0.51	-14.46	-0.64	-0.25
15.5	-2.61	-5.03	-1.13	-22.13	-4.35	-5.53	-1.08	-16.55	-1.87	-0.61
16.0	-3.01	-6.16	-1.84	-24.21	-5.34	-6.50	-1.64	-18.62	-3.09	-0.97
16.5	-3.27	-6.87	-2.29	-25.50	-5.96	-7.10	-2.00	-19.91	-3.85	-1.19
17.0	-3.52	-7.57	-2.73	-26.77	-6.57	-7.68	-2.35	-21.19	-4.60	-1.41
17.5	-3.60	-7.81	-2.88	-27.20	-6.78	-7.88	-2.47	-21.62	-4.86	-1.48
18.0	-3.69	-8.05	-3.03	-27.63	-6.98	-8.08	-2.59	-22.06	-5.12	-1.55
18.5	-3.60	-7.81	-2.88	-27.20	-6.78	-7.88	-2.47	-21.62	-4.86	-1.48
19.0	-3.52	-7.57	-2.73	-26.77	-6.57	-7.68	-2.35	-21.19	-4.60	-1.41
19.5	-3.27	-6.87	-2.29	-25.50	-5.96	-7.10	-2.00	-19.91	-3.85	-1.19
20.0	-3.01	-6.16	-1.84	-24.21	-5.34	-6.50	-1.64	-18.62	-3.09	-0.97
20.5	-2.61	-5.03	-1.13	-22.13	-4.35	-5.53	-1.08	-16.55	-1.87	-0.61
21.0	-2.19	-3.91	-0.42	-20.02	-3.33	-4.53	-0.51	-14.46	-0.64	-0.25
21.5	-1.65	-2.43	1.74	-17.23	-1.98	-3.21	0.79	-11.70	3.41	0.79
22.0	-1.10	-0.94	4.97	-14.40	-0.61	-1.84	3.37	-8.92	9.08	2.49
22.5	-0.45	2.76	8.73	-11.05	3.49	-0.21	6.39	-5.64	15.76	4.50
23.0	0.72	8.85	12.51	-7.66	9.19	5.11	9.44	-2.32	22.50	6.56
23.5	3.21	15.44	16.59	-3.96	15.41	11.57	12.74	4.42	29.80	8.81
24.0	5.75	22.11	20.68	-0.20	21.71	18.22	16.06	17.05	37.18	11.11

5.4 A NETWORK WITH VOLATILE POWER DEMAND

5.4.1 Stochastic power demand

We now come to our final simulation scenario. It will introduce stochastic power demand to model the uncertainty intrinsic to realistic power systems, which must react to customers extracting power without any central planning.

Stochastic power nodes in $\mathcal{V}_{\text{Power}}$

In order to incorporate uncertain power demands into our model we add a new kind of load node, the *stochastic PQ-node*. The type of uncertainty employed, the Ornstein–Uhlenbeck process, has a long history in modeling uncertain demands of various types and has also been used for electricity demand, see [Bar02; GKL21]. In [GKL21] a setting similar to ours was examined but applied to the Telegrapher’s equations, that were introduced in Chapter 3, instead of the power flow equations and only a single power plant was used.

The stochastic PQ-node, just like its deterministic cousin prescribes a real and reactive power demand as boundary conditions but now these demands are stochastic time-dependent quantities modeling the uncertainty of demand at this node. Of course this uncertainty is not total, as one may expect the demand to follow historic timelines of demand or some other estimate derived from knowledge about the season, weather or even current events like a sports tournament. This structure of uncertain fluctuation about a deterministic estimate suggests using a mean reverting stochastic process for the power demand, $(P_t)_{t \in [0, T]}$, that is, a process that is drawn back to some deterministic function $\mu(t)$ over time. If we further assume that fluctuations around μ are independent of the current time and also of the current value of P , a natural choice for the process is the Ornstein–Uhlenbeck process (OUP). It is characterized by the following stochastic differential equation,

$$dP_t = \theta (\mu(t) - P_t) dt + \sigma dW_t, \quad P_{t_0} = p_0, \quad (5.19)$$

where W_t is a one-dimensional Brownian motion, $\theta, \sigma > 0$ are the so-called *drift* and *diffusion* coefficients, and p_0 is the demand at the starting time t_0 . Note, that here we use the subscript t for the evaluation at time t instead of the derivative with respect to t . This ambiguity is unfortunate but this use of the subscript is customary for stochastic processes. However, as the stochastic process models power, where no PDEs enter our model, confusion is unlikely.

Whenever the current demand P_t differs from $\mu(t)$, the drift term enacts a force towards the deterministic demand estimate $\mu(t)$. This behavior is called *mean reversion*. The size of this force is characterized by the drift coefficient θ . In absence of diffusion (for $\sigma = 0$), the OUP degenerates to a deterministic ordinary differential equation, that is drawn to the mean exponentially. For $\sigma > 0$ on the other hand, this mean reversion is counteracted by fluctuations, whose size

is determined by σ . OUP realizations are similar to Figure 5.8 and Figure 5.9, although these realizations are capped at certain cut-offs, as described in Section 5.4.2.

Both the real power demand P as well as the reactive power demand Q are realized as an OUP in our setting.

Note that it is even possible to solve the stochastic differential Equation (5.19) explicitly via

$$P_t = p_0 e^{-\theta(t-t_0)} + \theta \int_{t_0}^t e^{-\theta(t-s)} \mu(s) ds + \sigma \int_{t_0}^t e^{-\theta(t-s)} dW_s. \quad (5.20)$$

From this explicit expression one can see that P_t is normally distributed with mean

$$\mu_t = p_0 e^{-\theta(t-t_0)} + \theta \int_{t_0}^t e^{-\theta(t-s)} \mu(s) ds$$

and variance

$$V = \sigma^2 \int_{t_0}^t e^{-2\theta(t-s)} ds.$$

The mathematical simplicity as well as the possibility to account for forecasts make the OUP a convenient candidate for modeling uncertainty in power demand, see also [BBKo8]. A drawback is the unbounded range, which can become a problem and which we will mitigate by introducing a cut-off.

5.4.2 Discretization

A general stochastic process is time-dependent and cannot simply be evaluated at every time step. Although this would be possible for the Ornstein–Uhlenbeck process as evidenced by Equation (5.20), we choose a discretization that can be extended to different stochastic processes, namely the explicit Euler-Maruyama method, see [SM96]. Due to the explicit nature, the time steps for this method must usually be chosen much finer than the time steps for the implicit box scheme, as we detail below. To make this distinction explicit we call the step size for the Euler-Maruyama method Δt_{stoch} . To choose the power boundary condition at a stochastic power node at time $t^* = t + \Delta t$, we make steps of size Δt_{stoch} according to

$$P(t + \Delta t_{\text{stoch}}) = P(t) + \theta \left(\hat{P}(t) - P(t) \right) \Delta t_{\text{stoch}} + \sigma S(\Delta t_{\text{stoch}}), \quad (5.21)$$

where \hat{P} takes on the role of the deterministic mean μ of Equation (5.19) and $S(\Delta t_{\text{stoch}})$ is a sample from a normal distribution with mean 0 and variance Δt_{stoch} . The same process is applied to get the discretized values of $Q(t)$. For stability in the mean (see again [SM96]), this discretization has the step size constraint

$$|1 - \theta \Delta t_{\text{stoch}}| < 1,$$

which for $\theta > 0$ yields

$$0 < \Delta t_{\text{stoch}} < \frac{2}{\theta}. \quad (5.22)$$

In addition we also restrict the stochastic power demand according to

$$\begin{aligned} (1 - c)\hat{P}(t) &\leq P(t) \leq (1 + c)\hat{P}(t) & \text{if } \hat{P}(t) > 0, \\ (1 - c)\hat{P}(t) &\geq P(t) \geq (1 + c)\hat{P}(t) & \text{if } \hat{P}(t) < 0, \end{aligned}$$

for some cut-off c with $0 \leq c \leq 1$. If the condition is violated, $P(t)$ is set to the boundary of the allowed interval. This cut-off prevents great outliers that would prevent our numerical methods from converging. Depending on the size of the chosen cut-off it may also be argued that such outliers are improbable and hence ignored. For real applications, the occurrence of such very drastic changes in power demand would possibly lead to black-outs, see [Pra+16].

It may be argued that a stochastic process, whose samples must sometimes be cast away to yield usable solutions, is a bad fit for its purpose. Unfortunately we are not aware of a process that has been shown to be especially accurate for power fluctuations. However, an alternative might be the Jacobi process, as recently proposed in [CK21], which stays within a predefined interval.

Samples of the OUP for a couple of choices for the cut-off can be found in Figure 5.8 and a zoomed in version in Figure 5.9. In these figures the influence of the cut-off is easily seen. Naturally a higher the cut-off means fewer instances of actually capping the stochastic process, which in turn means that the OUP is followed more faithfully for a high cut-off. Any cut-off $0 \leq c < 1$ can be chosen in principle, yet, the more volatile the process due to high σ , the more difficult numerical convergence becomes.

The process for the reactive power $Q(t)$ is the same.

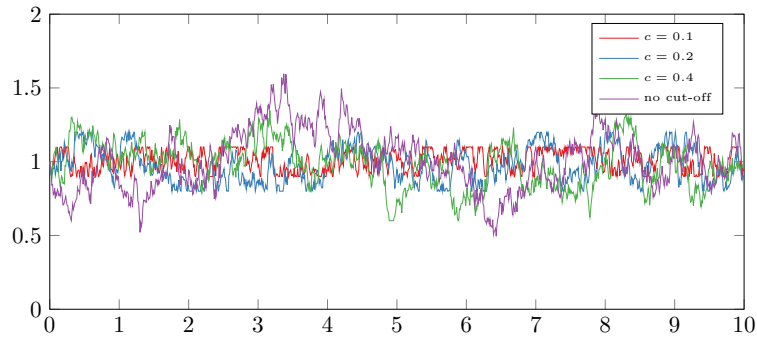


Figure 5.8: Ornstein-Uhlenbeck realizations for $\mu = 1.0, \theta = 3.0, \sigma = 0.45$ and different cut-off values c .

5.4.3 Network properties

Scenario description

Here we describe the considered scenario of a combined power and gas network. The data of the described networks of this section can be found in the git repository⁴, we created for this purpose.

⁴ https://github.com/eike-fokken/efficient_network-data.git

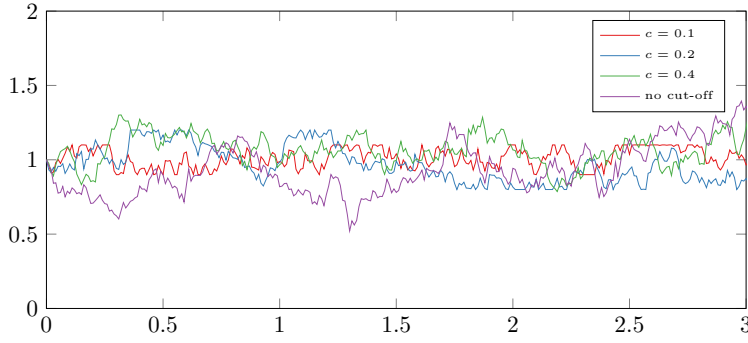


Figure 5.9: Zoomed-in version of Figure 5.8.

Specification of the power network

As starting point for the power network we use again the *IEEE-300-bus* system of the last example. This time we alter the IEEE-300 network in the following way.

- The power demand (real and reactive) is lowered by 10%.
- The former slack node N7049 is changed into a PV-node.
- The old PV-nodes given in Table 5.11 are turned into slack busses ($V\phi$ -nodes).
- All PQ-nodes are turned into stochastic PQ-nodes described in Section 5.4.1 and Section 5.4.2.

Once again, at the new $V\phi$ -nodes, power that is generated from gas burned in gas-fired power plants is injected into the power network according to Equation (5.16). An image of the power net can be found in Figure 5.6.

Specification of the gas network

As starting point for the gas network we again use the *GasLib-134* system (see [Sch+17]), which was depicted in Figure 5.5 on page 107. It is the same gas network as before, consisting of 86 pipelines, 3

Source node id	inflow [$\text{m}^3 \text{s}^{-1}$]
1	105.3282
20	280.6652
80	170.4607

Table 5.10: Inflow into the gas network

inflow nodes (*sources*) and 45 outflow nodes (*sinks*). The inflow of gas this time remains constant over time and is given in Table 5.10.

Also we increase the number of gas-power conversion plants and therefore, 17 of the sinks of the gas network, all gathered in Table 5.11, draw gas to be converted into power. The amount is set by the power network and is computed from the power flow equations. All other sinks do not consume gas in this scenario.

Specification of the Gas-Power connections

The two networks are connected through gas-power conversion plants, that turn gas into power, when power is needed, and turn power into gas, when surplus power is available. The gas-power conversion plants are arcs between the nodes listed in Table 5.11. For simplicity they all share the same efficiencies both for power generation and gas generation, that were used before, namely

$$E_{\text{GtP}} = 0.0796 \text{ m}^3 \text{ MJ}^{-1},$$

$$E_{\text{PtG}} = 0.0229 \text{ m}^3 \text{ MJ}^{-1}.$$

All further data concerning these plants is gathered in Table 5.11. In there, also a real power demand is given, which corresponds to the default demand in our setting, when no uncertainty is present.

Gasnode	Powernode	P[100MW]
ld2	N7017	2.2890
ld6	N7057	1.3952
ld10	N7071	0.7217
ld12	N7024	2.7771
ld13	N230	2.5978
ld23	N119	19.3000
ld24	N221	-0.0893
ld27	N187	11.4020
ld29	N7061	2.7269
ld31	N213	2.0176
ld33	N9051	-0.3581
ld35	N186	11.4020
ld36	N7001	2.1410
ld37	N9002	-0.0420
ld38	N7166	5.5300
ld39	N7003	12.1000
ld42	N7039	4.6702

Table 5.11: Start and end nodes of gas-power-conversion plants and the deterministic demands of real power.

Specification of the stochastic power demand nodes

As mentioned, all PQ-nodes of the IEEE-300-bus system are replaced by stochastic PQ-nodes. As mean function we choose the power demands given by the IEEE-300-bus problem, but lowered by 10 %. In addition, we choose $\theta = 3$ for the drift coefficient and for the stability constraint we choose

$$\Delta t_{\text{stoch}} = \frac{0.1}{\theta},$$

which results in rather high numbers of stochastic time steps but is unfortunately needed for convergence. For the cut-off we use

$c = 0.4$, which means that power demand will not deviate more than 40 % from the anticipated (mean) power demand. This choice was made to investigate high deviations while keeping the number of non-converging simulations small. The diffusion coefficient σ will be varied to compare different values.

5.4.4 Numerical results

In each run we simulate the combined network over a time horizon of 24 h (86 400 s) with the same time step size as before, $\Delta t = 0.5$ h (1800 s) and the same spatial step sizes in the pipeline as before ($\Delta x \approx 10$ km).

Steady-state vs. stochastic example

At first we simulate the network in a deterministic setting, which can be achieved by setting σ in (5.21) to zero. To keep the scenario simple, we choose steady-state initial conditions, which were generated by using arbitrary initial conditions and integrating them for a long time. The resulting end state is then used as initial conditions for our setting.

In this deterministic setting we find the (constant) power demands in the gas plants given in Table 5.11. To illustrate our results we will usually picture the situation of pipe p_{br71} , which is located to the lower right in Figure 5.5 connecting nodes 71 and 72. The steady-state solution in pipe p_{br71} remains constant over time as is fitting for a steady-state solution. The same is true for the flow, and also for all other pipes in the network.

Along with the deterministic setting, we simulate a scenario with $\theta = 3.0$ and $\sigma = 0.45$ for all PQ-nodes. The number of stochastic steps is set to at least 1000 which, due to stability constraints of Equation (5.22), is then automatically raised to 18000.

A comparison of the steady-state and stochastic pressure can be seen in Figure 5.10, while a comparison of the fluxes is given in Figure 5.11.

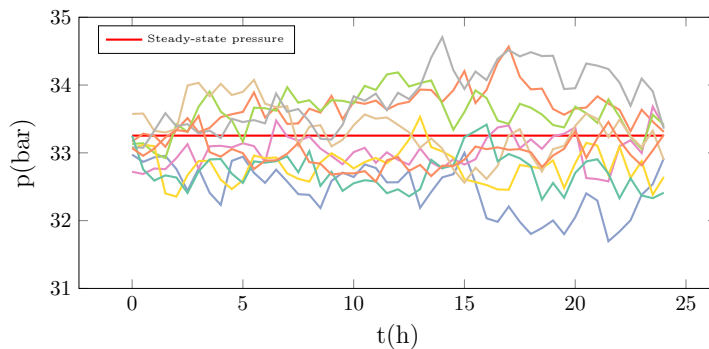


Figure 5.10: Pressure evolution in p_{br71} for deterministic and some realizations of stochastic demand.

In the power network we find for the PQ-node N_1 power demands over time like those in Figure 5.12 and Figure 5.13. Of course the situation is similar for all PQ-nodes.

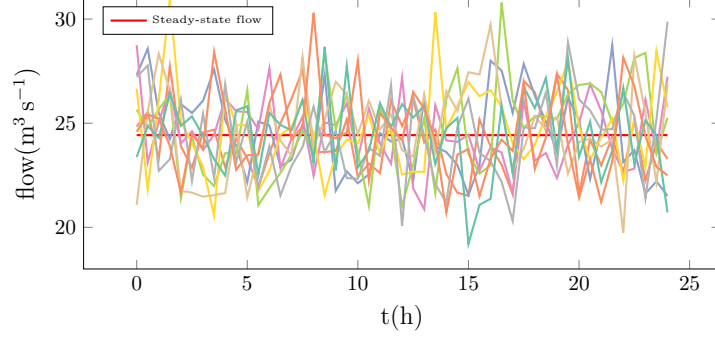


Figure 5.11: Flow evolution in p_{br71} for deterministic and some realizations of stochastic demand.

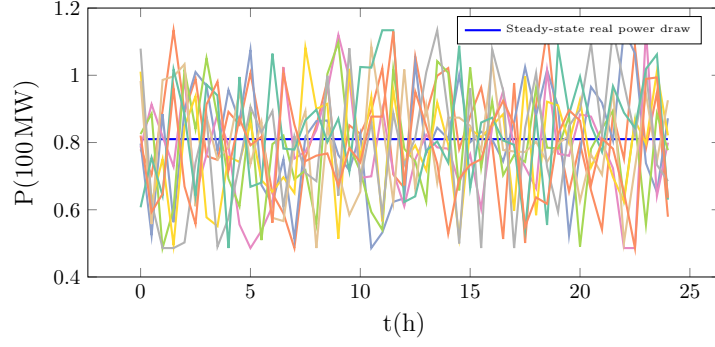


Figure 5.12: Real power demand in $N1$ for deterministic and stochastic demand with $\sigma = 0.45$.

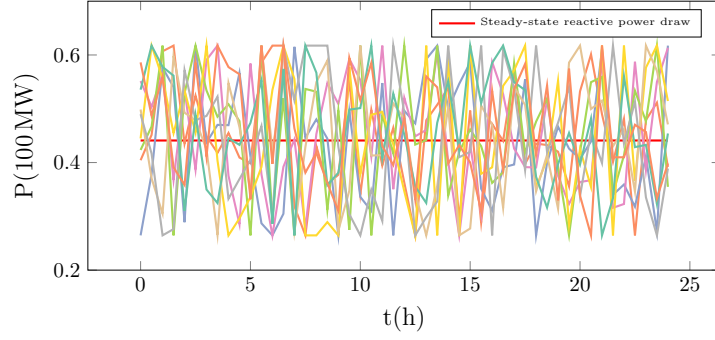


Figure 5.13: Reactive power demand in $N1$ for deterministic and stochastic demand with $\sigma = 0.45$.

Stochastic demand with variable noise

Now we examine repercussions of the uncertainty on the gas network. Therefore we make 100 runs for each $\sigma \in \{0.05, 0.1, 0.3, 0.45\}$ and compare the quantiles at 50 %, 75 % and 90 %. Taking an arbitrary point in time, $t = 12$ h, the quantiles for the pressure can be seen in Figure 5.14. For the flow the quantile comparison can be found in Figure 5.15.

For both quantities we see the expected expansion of quantile boundaries with higher diffusion σ .

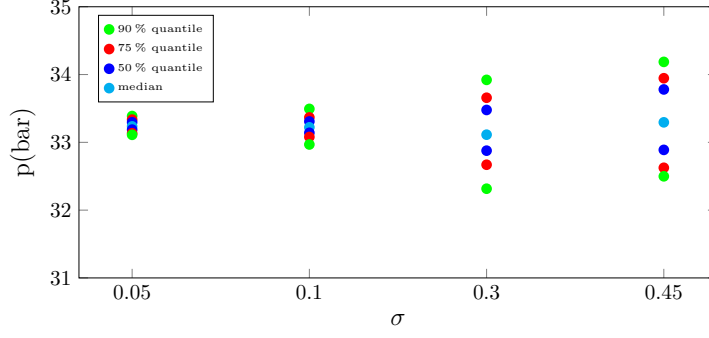


Figure 5.14: Comparison of pressure quantile boundaries at different σ at $t = 12$ h in pipeline p_{br71} .

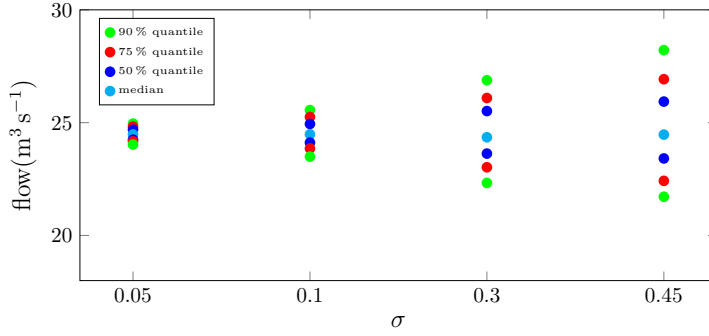


Figure 5.15: Comparison of flow quantile boundaries at different σ at $t = 12$ h in pipeline p_{br71} .

Comparison of deterministic and stochastic pressure prediction

Now we give an overview of the impact of the volatility in power demand on the network. Therefore we revisit the scenario with the highest volatility, that is with $\sigma = 0.45$ and consider again a time frame of 24 h. In Figure 5.16 on page 125 one can see the maximal deviation of real power demand from the steady-state solution. At a first glance this looks similar to Figure 5.6, just with colors cycled around. Although this impression turns out false, the similarity is due to the fact, that the load nodes have defined volatility as they follow their own Ornstein–Uhlenbeck process approximation defined in Equation (5.21). The PV-nodes on the other hand have zero volatility in real power. Yet the V ϕ -nodes must account for all remaining power demand and as such have the highest volatility. A similar picture can be found in Figure 5.17 on page 126, where the deviation of the reactive power is depicted. Here the PV-nodes do not have zero volatility, yet it seems that they also do not carry much volatility in Q .

At last we consider the possible impact of the volatility in the power network on the gas network. Therefore we show the maximal pressure deviation from the steady-state solution over the course of 24 h for $\sigma = 0.45$ in Figure 5.18 on page 126. It is easily seen that the lower part of the network experiences much higher pressure volatility than the upper part. This is expected, as on the one hand the upper part has higher pressure as the three gas sources are located there and on

the other hand many more gas-power conversion plants are located in the lower part, so that the volatility can add up.

5.5 SUMMARY AND OUTLOOK

In this chapter we have applied the theory of the last two chapters in three different scenarios. The first example showed the general applicability of our approach and showed plausible results for a small network.

The second scenario introduced power-to-gas plants and proved that our techniques can deal with bigger problems. Here we investigated the solution of a combined gas and power network where the gas network is used as an energy storage for surplus electrical power that can be tapped in times of deficient electrical power supply. Additionally we compared Bernoulli coupling and pressure coupling for this scenario and found little difference for our parameters. The data of these experiments have been gathered in a git repository⁵ for comparison with other approaches.

Lastly we introduced stochastic power demand and examined repercussions on the gas network given different levels of volatility in power demand. The data of these experiments can also be found in a git repository⁶ created for this purpose.

Although the different coupling conditions as proposed in [Rei15] do not introduce great changes in the results of realistic pipeline network simulations, it would be interesting to retrace the steps of [Rei15] in order to prove results similar to those of Section 4.2 and the bounds on the fuel demand in Equation (4.2).

⁵ <https://bitbucket.org/efokken/gas-power-benchmark>

⁶ https://github.com/eike-fokken/efficient_network-data.git

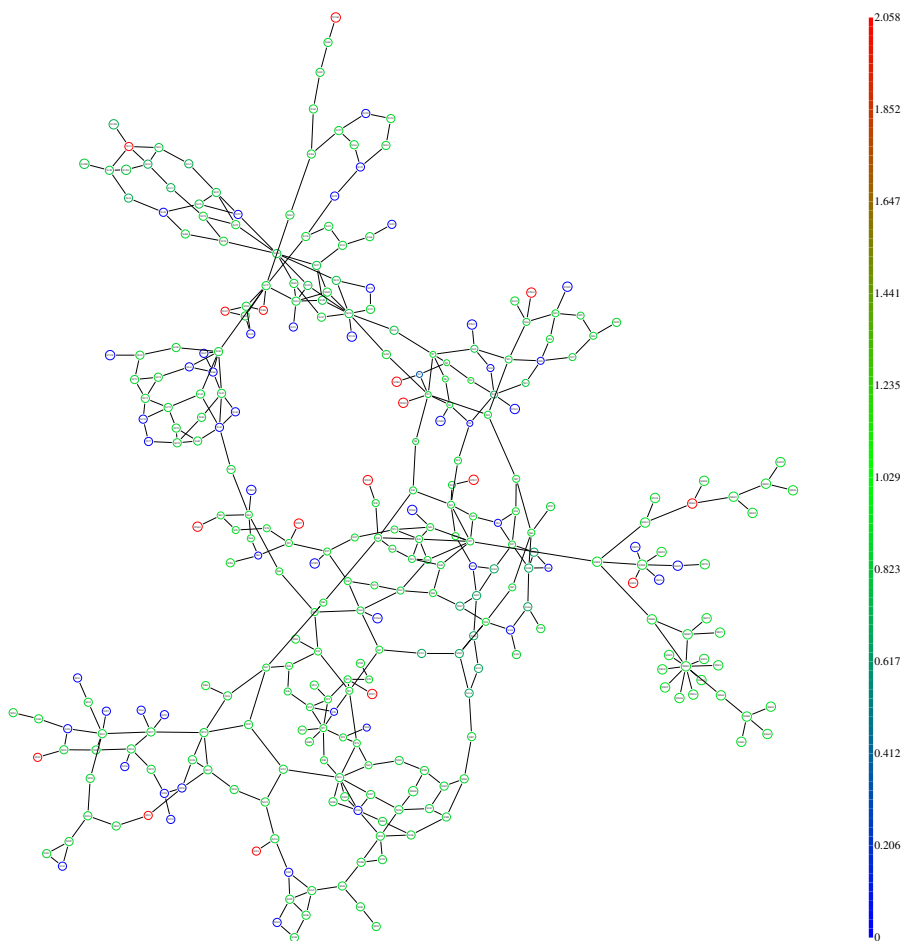


Figure 5.16: Heatmap of the maximal real power deviation over the course of 24 h, units are in 100 MW.

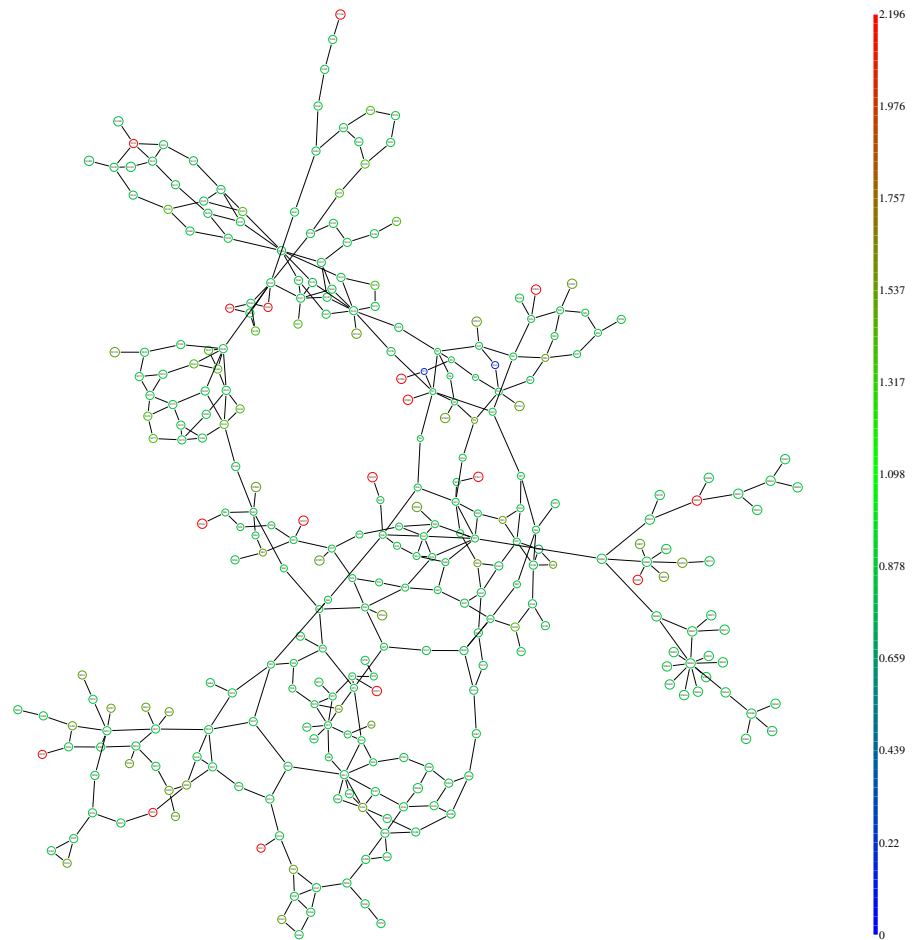


Figure 5.17: Heatmap of the maximal reactive power deviation over the course of 24h, units are in 100 MW.

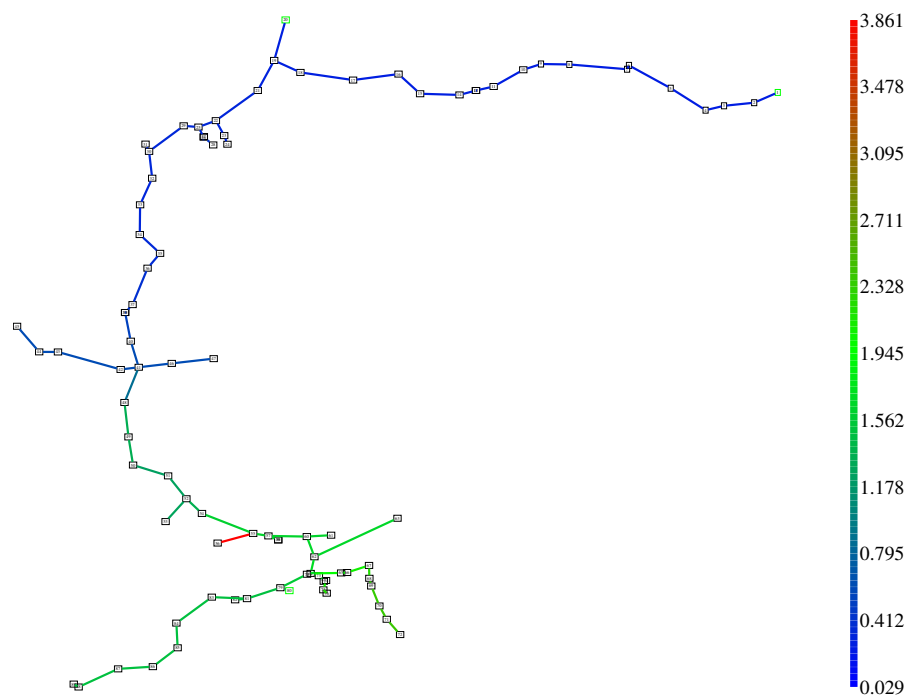


Figure 5.18: Heatmap of the maximal pressure deviation over the course of 24h, units are in bar.

Over the course of this thesis we have defined models for coupled gas and power networks, discretized them and simulated different scenarios. While this is valuable, it is often necessary to not only simulate, but also to find an optimal way of operating energy networks in order to achieve some goal, usually expressed as controlling the networks in such a way, that minimal costs are incurred, while still satisfying some demands.

Finding such an optimal operating procedure can often be cast as an optimal control problem, as defined in Section 2.3. This amounts to defining functions f, g, h , that model certain properties of the scenario in question, such that an optimal control of the network is given as the solution to

$$\begin{aligned} \mathbf{u} &= \operatorname{argmin} f(\mathbf{u}) \\ \text{s.t. } g(\mathbf{u}) &= 0 \\ h(\mathbf{u}) &\geq 0. \end{aligned}$$

In the case where all quantities of interest—in our case power and voltage in the power network, pressure and flux in the gas network—depend smoothly on all input quantities, such a problem can be solved with the tools of smooth optimization, see [NW06] for an introduction. Note that along with the discretizations of the last chapters we only consider control variables, that are themselves discrete (with a possibly different discretization lattice) so that our optimization falls under the so-called “first-discretize-then-optimize” approach. For the opposite approach see for instance [Ulbo2].

In order to compute minima of the above problem, we need to compute the derivatives of f, g and h . Note that first derivatives are sufficient to employ quasi-Newton optimization algorithms, see [NW06, Chapter 6]. Unfortunately the simulation implementation we used throughout Chapter 5 relies on solving non-linear systems of equations. Therefore the quantities of interest, usually called states, can be thought of as functions of initial and boundary conditions, that are only given implicitly. Although it is possible, via Newton’s method, to compute these functions, their derivatives are not readily available. Luckily, there are two well-known algorithms to produce these derivatives. On the one hand there is the so-called *direct approach*, which computes derivatives of the states directly, along the lines of the implicit function theorem, see for example [PGo8], in order to compute derivatives of the functions f, g and h . On the other hand there is the *adjoint approach*, see [Hero7; Kol11; PGo8], which sidesteps computing the state derivatives and directly computes the derivatives of f, g and h . In this chapter we will present and analyze both of these approaches for an optimal control problem that is discretized into distinct time steps.

Afterwards we will solve an optimal control problem on the combined gas-power network of Section 5.3 to showcase the optimization.

Before we come to the main part of this chapter, let us note that it is possible to also obtain second derivatives via a combination of the direct and adjoint method, detailed in [PGo8]. This would enable us to use Newton's method instead of quasi-Newton methods, but we forego this, as the quasi-Newton methods that rely on first derivatives only are performant enough and have a much reduced implementation overhead.

6.1 OPTIMIZATION WITH CONTROL AND STATE VARIABLES

When setting up an optimization problem for use with our networks, an immediate challenge is, that the objective and constraint functions do not solely depend on controls directly but also on state variables — in our case pressure, flow, power and voltage — that in turn depend on the controls.

We will now describe the theory of such problems in general and only later on come back to the gas and power networks.

Mathematically the dependence of state variables on the controls means that the state variables can be eliminated by replacing them with the function that computes them from the controls. Yet, often this function is only given implicitly by a set of equality constraints which the controls and states have to fulfill.

Such a problem is of the form

$$\begin{aligned} \mathbf{u} = \operatorname{argmin} \quad & f(\mathbf{u}, \mathbf{x}) \\ \text{s.t.} \quad & E(\mathbf{u}, \mathbf{x}) = 0 \\ & \mathbf{g}(\mathbf{u}, \mathbf{x}) = 0 \\ & \mathbf{h}(\mathbf{u}, \mathbf{x}) \geq 0, \end{aligned} \tag{6.1}$$

where all variables are now discretized, corresponding to our “first-discretize-then-optimize” approach. Let therefore $\mathbf{u} \in \mathbb{R}^{n_u}$, $\mathbf{x} \in \mathbb{R}^{n_s}$ and

$$\begin{aligned} f : \mathbb{R}^{n_u} \times \mathbb{R}^{n_s} &\rightarrow \mathbb{R}, \\ E : \mathbb{R}^{n_u} \times \mathbb{R}^{n_s} &\rightarrow \mathbb{R}^{n_s}, \\ \mathbf{g} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_s} &\rightarrow \mathbb{R}^{n_g}, \\ \mathbf{h} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_s} &\rightarrow \mathbb{R}^{n_h}. \end{aligned}$$

Additionally we require that the equation $E(\mathbf{u}, \mathbf{x}) = 0$ determines $\mathbf{x} = \mathbf{y}(\mathbf{u})$ uniquely as a function of \mathbf{u} . We also only consider problems where all these functions are smooth. By this replacement of \mathbf{x} , Problem (6.1) can be thought of as an optimization problem of type (2.30).

Therefore, if we can evaluate $y(\mathbf{u}) \in \mathbb{R}^{n_s}$ and the derivative of y with respect to \mathbf{u} , $y'(\mathbf{u}) = \frac{\partial y(\mathbf{u})}{\partial \mathbf{u}} \in \mathbb{R}^{n_s \times n_u}$, we can construct the functions

$$\begin{aligned}\tilde{f}(\mathbf{u}) &= f(\mathbf{u}, y(\mathbf{u})) \\ \tilde{f}'(\mathbf{u}) &= \frac{d\tilde{f}(\mathbf{u})}{d\mathbf{u}} = \partial_{\mathbf{u}} f(\mathbf{u}, y(\mathbf{u})) + \partial_x f(\mathbf{u}, y(\mathbf{u})) \cdot y'(\mathbf{u}) \\ \tilde{g}(\mathbf{u}) &= g(\mathbf{u}, y(\mathbf{u})) \\ \tilde{g}'(\mathbf{u}) &= \frac{d\tilde{g}(\mathbf{u})}{d\mathbf{u}} = \partial_{\mathbf{u}} g(\mathbf{u}, y(\mathbf{u})) + \partial_x g(\mathbf{u}, y(\mathbf{u})) \cdot y'(\mathbf{u}) \\ \tilde{h}(\mathbf{u}) &= h(\mathbf{u}, y(\mathbf{u})) \\ \tilde{h}'(\mathbf{u}) &= \frac{d\tilde{h}(\mathbf{u})}{d\mathbf{u}} = \partial_{\mathbf{u}} h(\mathbf{u}, y(\mathbf{u})) + \partial_x h(\mathbf{u}, y(\mathbf{u})) \cdot y'(\mathbf{u}).\end{aligned}\tag{6.2}$$

With these we can solve the problem (6.1) by solving (2.30) with the functions $\tilde{f}, \tilde{g}, \tilde{h}$.

To evaluate these functions at some point \mathbf{u} , we first must compute $y(\mathbf{u})$, which is straightforward to do by applying a Newton-type method to the system of equations $E(\mathbf{u}, x) = 0$ and setting $y(\mathbf{u}) = x$. Then we can evaluate $\tilde{f}(\mathbf{u}), \tilde{g}(\mathbf{u}), \tilde{h}(\mathbf{u})$. Yet, there are two alternatives to evaluate the derivatives $\tilde{f}'(\mathbf{u}), \tilde{g}'(\mathbf{u}), \tilde{h}'(\mathbf{u})$, both of which we now describe shortly. A more detailed explanation can be found in [PGo8] and [Cap11].

On the one hand we can compute $y'(\mathbf{u})$ via the so-called *direct method*, by noting first that $E(\mathbf{u}, y(\mathbf{u})) = 0$ for all \mathbf{u} implies

$$0 = \frac{d}{d\mathbf{u}} E(\mathbf{u}, y(\mathbf{u})) = \partial_{\mathbf{u}} E(\mathbf{u}, y(\mathbf{u})) + \partial_x E(\mathbf{u}, y(\mathbf{u})) \cdot y'(\mathbf{u})$$

and therefore

$$\partial_x E(\mathbf{u}, y(\mathbf{u})) y'(\mathbf{u}) = -\partial_{\mathbf{u}} E(\mathbf{u}, y(\mathbf{u})),\tag{6.3}$$

through which $y'(\mathbf{u})$ can be evaluated by a linear solve. Note, that $y'(\mathbf{u})$ is a matrix and therefore Equation (6.3) can be thought of as a collection of n_u usual linear equations, one of each column of $y'(\mathbf{u})$. This can be plugged into Equation (6.2) to evaluate the derivatives.

Note however that $y'(\mathbf{u})$ is an $n_s \times n_u$ matrix and hence the linear solve must be done for n_u right-hand sides. If there are many controls, it may be desirable to use another method.

This other method is the so-called *adjoint method*. Here we define new types of Lagrange functions, $\mathcal{L}_f, \mathcal{L}_g, \mathcal{L}_h$, one for the objective function, one for the equality constraints and one for inequality constraints¹. The procedure is the same for each of these and therefore we show it only for the inequality constraints,

$$\mathcal{L}_h(\mathbf{u}, x, \xi) = h(\mathbf{u}, x) + \xi^T E(\mathbf{u}, x),$$

where $\xi \in \mathbb{R}^{n_s \times n_h}$ is a matrix of Lagrange multipliers.

As before there holds $E(\mathbf{u}, y(\mathbf{u})) = 0$ and $\frac{d}{d\mathbf{u}} E(\mathbf{u}, y(\mathbf{u})) = 0$. For the Lagrange function evaluated at $x = y(\mathbf{u})$ this means

$$\mathcal{L}_h(\mathbf{u}, y(\mathbf{u}), \xi) = h(\mathbf{u}, y(\mathbf{u})) + \xi^T E(\mathbf{u}, y(\mathbf{u})) = h(\mathbf{u}, y(\mathbf{u}))$$

¹ Actually the optimization library we use, namely IPOPT [WBo6], only accepts inequality constraints as input and checks internally, whether pairs of inequality constraints of the form $g(x) \leq 0, g(x) \geq 0$ appear, which are then treated as equality constraints.

and

$$\begin{aligned}\frac{d}{du}\mathcal{L}_h(\mathbf{u}, \mathbf{y}(\mathbf{u}), \xi) &= \frac{d}{du}\mathbf{h}(\mathbf{u}, \mathbf{y}(\mathbf{u})) + \xi^T \frac{d}{du}E(\mathbf{u}, \mathbf{y}(\mathbf{u})) \\ &= \frac{d}{du}\mathbf{h}(\mathbf{u}, \mathbf{y}(\mathbf{u})).\end{aligned}$$

Therefore we find for the derivative $\tilde{\mathbf{h}}'(\mathbf{u}) = \frac{d}{du}\mathbf{h}(\mathbf{u}, \mathbf{y}(\mathbf{u}))$:

$$\begin{aligned}\tilde{\mathbf{h}}'(\mathbf{u}) &= \frac{d}{du}\mathcal{L}_h(\mathbf{u}, \mathbf{y}(\mathbf{u}), \xi) \\ &= \partial_{\mathbf{u}}\mathcal{L}_h(\mathbf{u}, \mathbf{y}(\mathbf{u}), \xi) + \partial_{\mathbf{x}}\mathcal{L}_h(\mathbf{u}, \mathbf{y}(\mathbf{u}), \xi)\mathbf{y}'(\mathbf{u}) \\ &= \partial_{\mathbf{u}}\mathbf{h}(\mathbf{u}, \mathbf{y}(\mathbf{u})) + \xi^T \partial_{\mathbf{u}}E(\mathbf{u}, \mathbf{y}(\mathbf{u})) \\ &\quad + \partial_{\mathbf{x}}\mathcal{L}_h(\mathbf{u}, \mathbf{y}(\mathbf{u}), \xi)\mathbf{y}'(\mathbf{u})\end{aligned}\tag{6.4}$$

This is easy to evaluate, if the last line vanishes. Luckily we are still free to choose ξ and therefore demand

$$0 = \partial_{\mathbf{x}}\mathcal{L}_h(\mathbf{u}, \mathbf{y}(\mathbf{u}), \xi) = \partial_{\mathbf{x}}\mathbf{h}(\mathbf{u}, \mathbf{y}(\mathbf{u})) + \xi^T \partial_{\mathbf{x}}E(\mathbf{u}, \mathbf{y}(\mathbf{u})),$$

which is equivalent to

$$\partial_{\mathbf{x}}E(\mathbf{u}, \mathbf{y}(\mathbf{u}))^T \xi = -\partial_{\mathbf{x}}\mathbf{h}(\mathbf{u}, \mathbf{y}(\mathbf{u}))^T, \tag{6.5}$$

from which ξ can be determined by a linear solve.

This time $\partial_{\mathbf{x}}\mathbf{h}(\mathbf{u}, \mathbf{y}(\mathbf{u}))^T$ is an $n_s \times n_h$ matrix (mind the transpose). To compute all derivatives in Equation (6.2), we also must follow these steps for \mathbf{f} , which yields one additional linear solve, and \mathbf{g} , which yields n_g additional solves, summing up to $n_g + n_h + 1$ linear solves.

Comparing the two methods, it seems plausible to favor the direct method, if $n_u < n_g + n_h + 1$, and the adjoint method in the other case.

One big drawback of both methods is the need to solve a linear system involving the (sometimes huge) matrix $\frac{\partial E}{\partial \mathbf{x}}$. To overcome this issue we will now describe algorithms to carry out direct and adjoint methods that exploit the structure that arises from the decomposition of our problems into separate time steps.

6.2 DERIVATIVES OF TIME-DEPENDENT SYSTEMS

The computation of derivatives up to now is valid for any optimization problem with objective and constraints given by Equation (6.2). Yet, the problems we solve have a special structure, they are discretized into time steps. Therefore it is on the one hand natural to look for opportunities to break apart the computations into computations for each time step and on the other hand it is clear that states and evaluations of \mathbf{f} , \mathbf{g} and \mathbf{h} at earlier time steps cannot depend on controls or states of later steps.

To break apart all quantities into the different time steps we need three different time lattices, one for states (and objective function), one for the constraints and one for the controls. Corresponding quantities are listed in Table 6.1.

Of course there must hold $m_s \cdot N^s = n_s$ and likewise for controls and constraints. In addition we demand $T^h \subset T^s$, so that every constraint time point must also be a state time point. The control

Table 6.1: Time lattices and corresponding quantities

	states s	controls u	constraints h
lattice sets	T^s	T^u	T^h
index sets	I^s	I^u	I^h
number of time points	$N^s = T^s $	$N^u = T^u $	$N^h = T^h $
time points	$t_i^s \in T^s,$ $i \in I^s$	$t_i^u \in T^u,$ $i \in I^u$	$t_i^h \in T^h,$ $i \in I^h$
number per time point	m_s	m_u	m_h
quantity at time point	$x_i \in \mathbb{R}^{m_s}$	$u_i \in \mathbb{R}^{m_u}$	$h_i \in \mathbb{R}^{m_h}$

time points are allowed to be different, but we demand that they enclose the state time points in the sense that $\min T^u \leq \min T^s$ and $\max T^u \geq \max T^s$.

As the state time lattice and the control time lattice can be disjoint, it doesn't make sense at first to speak of $u(t_i^s)$, a control at a state time point. We define this quantity by linear interpolation. Let us make that explicit. For a state index $i \in I^s$ we define $\tilde{i}, \hat{i} \in I^u$ by

$$\begin{aligned} \tilde{i} &= \operatorname{argmax}_{i \in I^u} t_i^u \\ &\text{s.t. } t_i^u \leq t_i^s \\ \hat{i} &= \operatorname{argmin}_{i \in I^u} t_i^u \\ &\text{s.t. } t_i^u \geq t_i^s. \end{aligned}$$

So $t_{\tilde{i}}^u$ and $t_{\hat{i}}^u$ are the closest control time points that surround t_i^s . Actually there always holds $\hat{i} = \tilde{i} + 1$. Then there is a unique $\lambda_i \in [0, 1]$, such that

$$t_i^s = (1 - \lambda_i)t_{\tilde{i}}^u + \lambda_i t_{\hat{i}}^u.$$

With this we can define controls at state time points,

$$\tilde{u}_i = (1 - \lambda_i)u_{\tilde{i}} + \lambda_i u_{\hat{i}} \in \mathbb{R}^{m_u}. \quad (6.6)$$

We will always write these state-time controls with a tilde, to distinguish them from the control variables defined on the control time lattice T^u . Also the matrix $\frac{\partial \tilde{u}}{\partial u}$ will be useful. It is a block matrix with block-size $m_u \times m_u$. Its structure is given as follows, where the indices indicate the respective block.

$$\begin{aligned} \frac{\partial \tilde{u}}{\partial u} &\in \mathbb{R}^{n_s \times n_u} \\ \left(\frac{\partial \tilde{u}}{\partial u} \right)_{i, \tilde{i}} &= (1 - \lambda_i) \operatorname{Id}_{m_u} \\ \left(\frac{\partial \tilde{u}}{\partial u} \right)_{i, \hat{i}} &= \lambda_i \operatorname{Id}_{m_u} \\ \left(\frac{\partial \tilde{u}}{\partial u} \right)_{i, j} &= 0, \text{ for } j \neq \tilde{i}, \hat{i}. \end{aligned} \quad (6.7)$$

We now explain the reason for the structure of the constraint time lattice. The constraints we are interested in, are, in continuous time given by some function $\hat{h}(u, x)$, that should fulfill $\hat{h}(u(t), x(t)) \geq 0$ for

every time $t \in [0, T]$. For the discretized time we instead demand the inequality to hold at all lattice time points $t \in T^s$. But to reduce the numerical cost, we may opt to check the constraints only at a subset $T^h \subset T^s$. Therefore the continuous constraint $\hat{\mathbf{h}}(\mathbf{u}(t), \mathbf{x}(t)) \geq 0$ for $t \in [0, T]$ is transformed into the discretized constraint

$$\begin{aligned} \mathbf{h} &= (\mathbf{h}_i)_{i \in I^h} \\ \mathbf{h}_i &= \hat{\mathbf{h}}(\tilde{\mathbf{u}}_i, \mathbf{x}_i) \in \mathbb{R}^{m_h}, \end{aligned} \quad (6.8)$$

where $\mathbf{h}_i \in \mathbb{R}^{m_h}$. Because the index structure is a bit non-standard, we provide an example.

Example 6.1. Let $I^s = \{1, \dots, 5\}$ and $I^h = \{1, 3, 5\}$. In that case, \mathbf{h} is given as $\mathbf{h} = (\mathbf{h}_1, \mathbf{h}_3, \mathbf{h}_5)$, while \mathbf{h}_2 and \mathbf{h}_4 are not defined.

This of course leads to a very sparse structure of $\frac{\partial \mathbf{h}}{\partial \mathbf{x}}$ and $\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{u}}}$, namely

$$\begin{aligned} \frac{\partial \mathbf{h}}{\partial \mathbf{x}} &\in \mathbb{R}^{n_h \times n_s} \\ \frac{\partial \mathbf{h}_i}{\partial \mathbf{x}_i} &\in \mathbb{R}^{m_h \times m_s} \\ \left(\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right)_{i,i} &= \begin{cases} \frac{\partial \mathbf{h}_i}{\partial \mathbf{x}_i} & \text{if } i \in I^h \\ 0 & \text{else.} \end{cases} \end{aligned} \quad (6.9)$$

and likewise for $\frac{\partial \mathbf{h}}{\partial \tilde{\mathbf{u}}}$. We will write this out for our example above.

Example 6.2. For the setting of Example 6.1 there holds

$$\frac{\partial \mathbf{h}}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial \mathbf{h}_1}{\partial \mathbf{x}_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial \mathbf{h}_3}{\partial \mathbf{x}_3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\partial \mathbf{h}_5}{\partial \mathbf{x}_5} \end{pmatrix}.$$

Note again the index structure. $\frac{\partial \mathbf{h}_3}{\partial \mathbf{x}_3}$ for example is in the second row, not in the third.

The objective function is evaluated at every state time point but apart from that similar to the constraints. We consider continuous objective functions $\hat{\mathbf{f}}$ of the form

$$\hat{\mathbf{f}}(\mathbf{u}, \mathbf{x}) = \int_0^T \mathbf{f}_{\text{kernel}}(\mathbf{u}(t), \mathbf{x}(t)) dt. \quad (6.10)$$

It is discretized with the trapezoidal rule at every state time point yielding for the discretized objective:

$$\tilde{\mathbf{f}}(\mathbf{u}, \mathbf{x}) = \sum_{i=0}^{N^s} \alpha_i \mathbf{f}_{\text{kernel}}(\tilde{\mathbf{u}}_i, \mathbf{x}_i), \quad (6.11)$$

where α_i is the weight of the trapezoidal rule, namely

$$\begin{aligned} \alpha_0 &= \frac{1}{2}(t_1 - t_0) \\ \alpha_{N^s} &= \frac{1}{2}(t_{N^s} - t_{N^s-1}) \\ \alpha_i &= \frac{1}{2}(t_{i+1} - t_{i-1}), \text{ for } 0 < i < N^s. \end{aligned}$$

Its derivatives $\frac{\partial \tilde{f}}{\partial \mathbf{x}} \in \mathbb{R}^{1 \times n_s}$ and $\frac{\partial \tilde{f}}{\partial \tilde{\mathbf{u}}} \in \mathbb{R}^{1 \times n_s}$ on the other hand are usually dense and given by

$$\begin{aligned} \left(\frac{\partial \tilde{f}}{\partial \mathbf{x}} \right)_i &= \alpha_i \partial_{\mathbf{x}} f_{\text{kernel}}(\tilde{\mathbf{u}}_i, \mathbf{x}_i) \in \mathbb{R}^{m_s}, \\ \left(\frac{\partial \tilde{f}}{\partial \tilde{\mathbf{u}}} \right)_i &= \alpha_i \partial_{\tilde{\mathbf{u}}} f_{\text{kernel}}(\tilde{\mathbf{u}}_i, \mathbf{x}_i) \in \mathbb{R}^{m_u}. \end{aligned}$$

The only quantities left to discretize are the equality constraints E that determine the state variables \mathbf{x} for given control variables \mathbf{u} . These can also be decomposed into E_i , one for every state time point, such that $E = (E_1, \dots, E_{N_s})$. The components of E_i have the structure

$$E_i = E_i(\mathbf{u}_i, \mathbf{x}_i, \mathbf{x}_{i-1}) \in \mathbb{R}^{m_s}.$$

This structure means that the state variable \mathbf{x}_i can depend directly only on the preceding state variable (called \mathbf{x}_{i-1}) and the control at the current time point, $\tilde{\mathbf{u}}_i$ (remember that $\tilde{\mathbf{u}}_i$ itself depends on the two control variables $\mathbf{u}_i, \mathbf{u}_i$). For the derivatives of E_i we define

$$\begin{aligned} A_i &= \frac{\partial E_i}{\partial \mathbf{x}_{i-1}} \in \mathbb{R}^{m_s \times m_s} \\ B_i &= \frac{\partial E_i}{\partial \mathbf{x}_i} \in \mathbb{R}^{m_s \times m_s} \\ C_i &= \frac{\partial E_i}{\partial \tilde{\mathbf{u}}_i} \in \mathbb{R}^{m_s \times m_u}, \end{aligned}$$

similar to [Kol11, Section 5.2.2] and find

$$\frac{\partial E}{\partial \mathbf{x}} = \begin{pmatrix} B_1 & & & \\ A_2 & B_2 & & \\ & \ddots & \ddots & \\ & & A_{N_s} & B_{N_s} \end{pmatrix}, \quad \frac{\partial E}{\partial \tilde{\mathbf{u}}} = \begin{pmatrix} C_1 & & \\ & \ddots & \\ & & C_{N_s} \end{pmatrix}, \quad (6.12)$$

where all omitted entries are 0 and of course

$$\frac{\partial E}{\partial \mathbf{u}} = \frac{\partial E}{\partial \tilde{\mathbf{u}}} \frac{\partial \tilde{\mathbf{u}}}{\partial \mathbf{u}} \in \mathbb{R}^{n_s \times n_u}$$

from (6.7). The structure of this last matrix is given as follows:

$$\begin{aligned} \frac{\partial E}{\partial \mathbf{u}} &\in \mathbb{R}^{n_s \times n_u} \\ \left(\frac{\partial E}{\partial \mathbf{u}} \right)_{i,\hat{i}} &= (1 - \lambda_i) C_i \\ \left(\frac{\partial E}{\partial \mathbf{u}} \right)_{i,\hat{i}} &= \lambda_i C_i \\ \left(\frac{\partial E}{\partial \mathbf{u}} \right)_{i,j} &= 0, \text{ for } j \neq \hat{i}, \hat{i}. \end{aligned}$$

With this we can examine the direct and adjoint method in detail.

6.2.1 The discretized direct method

For the direct method we have to solve the system (6.3), namely

$$\frac{\partial E}{\partial \mathbf{x}} \mathbf{Y} = -\frac{\partial E}{\partial \mathbf{u}}, \quad (6.13)$$

for $Y = y'(\mathbf{u}) \in \mathbb{R}^{n_s \times n_u}$. This can be done block-row by block-row, starting from the first and using the structure of the matrices. We have gathered the steps in Algorithm 6.1². There Y_i denotes the i -th block-row of Y and the inverse algorithmically means a linear solve, usually by LU-decomposition. This procedure results in a special structure

Algorithm 6.1 Computation of $Y = \frac{dy}{du}$

```

 $Y = 0$ 
 $Y_{1,i} += (1 - \lambda_1) C_1$ 
 $Y_{1,i} += \lambda_1 C_1$ 
 $Y_1 = -B_1^{-1} Y_1$ 
for  $i = 1$  to  $N_s$  do
   $Y_i = A_i Y_{i-1}$ 
   $Y_{i,i} += (1 - \lambda_i) C_i$ 
   $Y_{i,i} += \lambda_i C_i$ 
   $Y_i = -B_i^{-1} Y_i$ 
end for

```

of Y , similar to that of a lower block-triangular matrix. If $\frac{\partial \tilde{\mathbf{u}}}{\partial \mathbf{u}} = \text{Id}$, it actually is a block-triangular matrix. The structure is characterized by

$$Y_{i,j} = 0 \text{ for } j > i.$$

This structure can be exploited to make the algorithm numerically cheaper. In addition it encodes the physical property that future controls cannot influence states in the past with the exception of an influence via the interpolation (6.6).

What still needs to be done is, plugging $\frac{dy}{du}$ into Equation (6.2), which is straight-forward. As the derivatives $\frac{\partial \mathbf{h}}{\partial \mathbf{x}}$ and $\frac{\partial \mathbf{h}}{\partial \mathbf{u}}$, found in and below Equation (6.9), are extremely sparse, computing $\frac{d\mathbf{h}}{du}$ from $\frac{dy}{du}$ is numerically cheap. Also storage of the whole matrix $\frac{dy}{du}$ can be avoided, using each row of it directly in the loop of Algorithm 6.1 to compute the corresponding row in $\frac{d\mathbf{h}}{du}$.

The most expensive part of the direct method is the solution of System (6.13), which is more expensive, the more columns its right-hand side has, which in turn is given by the number of controls \mathbf{u} in the optimization problem. For the direct method we can therefore summarize that choosing it is the better, the fewer controls are in the system.

6.2.2 The discretized adjoint method

The adjoint method also benefits from the decomposition of the system in time steps. This time we have to solve the system (6.5), namely we have to solve

$$\frac{\partial E}{\partial \mathbf{x}}^T \xi = -\frac{\partial \mathbf{h}}{\partial \mathbf{x}}^T, \quad (6.14)$$

² Note that in algorithms, an equality sign “=” denotes an assignment, meaning, that the right-hand side of it should be computed and the left-hand side should then be replaced by the result of that computation. One and the same quantity may appear on both sides and the rule applies in that case, too. Also $a+ = b$ shall mean $a = a + b$.

where $\xi \in \mathbb{R}^{n_s \times n_h}$. The matrix $\frac{\partial \mathbf{E}}{\partial \mathbf{x}}^T$ is just the transpose of the matrix in (6.12), but we present it here, so the algorithm for solving Equation (6.13) is easier to understand:

$$\frac{\partial \mathbf{E}}{\partial \mathbf{x}}^T = \begin{pmatrix} B_1^T & A_2^T & & \\ & B_2^T & \ddots & \\ & & \ddots & A_m^T \\ & & & B_m^T \end{pmatrix} \quad (6.15)$$

With this, (6.15) can be solved block-row by block-row, this time starting from the last row. A description of this can be found in [Kol11, Section 5.2.2], especially Equation 5.15 therein.

The full solution algorithm is given in Algorithm 6.2, where ξ_i means the i -th block-row of ξ . The solution ξ will again exhibit a

Algorithm 6.2 Computation of ξ

```

 $\xi = 0$ 
 $\xi_{N_s, N_s} = - \left( B_{N_s}^T \right)^{-1} \frac{\partial \mathbf{h}_{N_s}}{\partial \mathbf{x}_{N_s}}^T$ 
for  $i = N_s - 1$  to  $1$  do
     $\xi_i = A_{i+1}^T \xi_{i+1}$  ▷ A
    if  $\left( \frac{\partial \mathbf{h}}{\partial \mathbf{x}}^T \right)_{i,j} \neq 0$  then
         $\xi_{i,j} += \left( \frac{\partial \mathbf{h}}{\partial \mathbf{x}}^T \right)_{i,j}$ 
    end if
     $\xi_i = - \left( B_i^T \right)^{-1} \xi_i.$  ▷ B
end for

```

special structure, which this time is upper block-triangular. This can be exploited in the lines marked A and B, by not taking a full row of ξ but only the part of it, that can be non-zero at all.

We still need to compute $\frac{d\mathbf{h}}{d\mathbf{u}}$ from ξ via (6.4), which can be done block-column by block-column and is shown in Algorithm 6.3. Here the interpolation (6.6) is also incorporated. The structure of ξ again forces a block-trigonal structure on $\frac{d\mathbf{h}}{d\mathbf{u}}$, which is *lower* trigonal, because ξ appears transposed. This again encodes the impossibility of future controls influencing past constraints.

Algorithm 6.3 Updating columns of $\frac{d\mathbf{h}}{d\mathbf{u}}$ and $\frac{d\mathbf{h}}{d\mathbf{u}}$

```

 $\frac{d\mathbf{h}}{d\mathbf{u}_i} = (\xi_i)^T \cdot \frac{\partial \mathbf{E}}{\partial \mathbf{u}_i}$ 
 $\frac{d\mathbf{h}_i}{d\mathbf{u}_i} += \frac{\partial \mathbf{h}_i}{\partial \mathbf{u}_i}$ 
 $\frac{d\mathbf{h}_i}{d\mathbf{u}_i} += (1 - \lambda_i) \frac{d\mathbf{h}_i}{d\mathbf{u}_i}$ 
 $\frac{d\mathbf{h}_i}{d\mathbf{u}_i} += \lambda_i \frac{d\mathbf{h}_i}{d\mathbf{u}_i}$ 

```

Like in the direct case (stated after Algorithm 6.1), storage of the whole matrix ξ can be avoided by evaluating Algorithm 6.3 inside the for-loop of Algorithm 6.2. We will now give an example to show the structure of ξ .

Example 6.3. As in Example 6.2, we consider a problem with $N_s = 5$ time steps and $n_h = 3$ constraint evaluations, which take place on time steps 1, 3, 5, that is, $I^h = \{1, 3, 5\}$. The matrices in the linear system to be solved have the form

$$\frac{\partial E}{\partial x}^T = \begin{pmatrix} B_1^T & A_2^T & & & \\ & B_2^T & A_3^T & & \\ & & B_3^T & A_4^T & \\ & & & B_4^T & A_5^T \\ & & & & B_5^T \end{pmatrix},$$

$$\xi = \begin{pmatrix} \xi_{1,1} & \xi_{1,3} & \xi_{1,5} \\ \xi_{2,1} & \xi_{2,3} & \xi_{2,5} \\ \xi_{3,1} & \xi_{3,3} & \xi_{3,5} \\ \xi_{4,1} & \xi_{4,3} & \xi_{4,5} \\ \xi_{5,1} & \xi_{5,3} & \xi_{5,5} \end{pmatrix},$$

and

$$\frac{\partial h}{\partial x}^T = \begin{pmatrix} \frac{\partial h_1}{\partial x_1}^T & & & & \\ 0 & & & & \\ & \frac{\partial h_3}{\partial x_3}^T & & & \\ & 0 & & 0 & \\ & & & \frac{\partial h_5}{\partial x_5}^T \end{pmatrix}.$$

With these the condition of the if-statement in Algorithm 6.2 is satisfied exactly for $\xi_{3,3}, \xi_{1,1}$ and of course at the start of the algorithm for $\xi_{5,5}$. Therefore the structure of ξ is given as

$$\xi = \begin{pmatrix} \xi_{1,1} & \xi_{1,3} & \xi_{1,5} \\ 0 & \xi_{2,3} & \xi_{2,5} \\ 0 & \xi_{3,3} & \xi_{3,5} \\ 0 & 0 & \xi_{4,5} \\ 0 & 0 & \xi_{5,5} \end{pmatrix}.$$

We have presented the adjoint method for the inequality constraints h . IPOPT, our optimization library (see [WBo6]), only accepts inequality constraints, so these are all constraints we must consider. Algorithms 6.2 and 6.3 can be used verbatim for the objective function f by replacing h with f . The only differences are, that $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial u}$ have only one row and that the condition in the if-statement is true in every iteration of the loop.

For the adjoint method, the most expensive part of the evaluation of $\frac{dh}{du}$ is—similar to the direct method—the solution of the linear system (6.14). Again this is more expensive, the more columns the right-hand side has. In this case this means, the adjoint method is better, the fewer constraints are in the problem.

Choosing a method to implement

As just explained both the direct and the adjoint method have their advantages. In our software *grazer* we implement only the adjoint

method, mostly because its performance is high enough that there is no need for further optimization. One reason for this is that in order to use any method for computing the derivative, we first must solve $E(\mathbf{u}, \mathbf{x}) = 0$. This is done via Newton's method and usually involves many more linear solves than those needed for the derivative computation in this chapter. Hence, the number of time steps needed (with a constraint in every time step) to yield a considerable cost of evaluating (6.14) in comparison to the cost of solving $E(\mathbf{u}, \mathbf{x}) = 0$ is difficult to reach. Still it would be interesting to implement the direct method as an alternative for problems of few controls but many constraints.

6.3 ADDITIONAL MODELS FOR OPTIMAL CONTROL PROBLEMS

In order to apply the adjoint method to an optimal control problem, we revisit the controlled components of Section 5.3.1.

6.3.1 Compressors and control valves

Compressors and control valves were introduced in Section 5.3.1. Both are modeled similarly and come equipped with a control function \mathbf{u} that influences the pressure. For the compressor we introduce a quadratic cost integrated over time, such that the kernel function in Equation (6.10) for the compressor is given by

$$\mathbf{f}_{\text{kernel}}^{\text{comp}}(\mathbf{u}(t)) = \mathbf{u}(t)^2.$$

After discretizing the integral, as in Equation (6.11), the total cost of compressor usage for a single compressor over the time horizon is given by

$$\tilde{\mathbf{f}}^{\text{comp}}(\mathbf{u}, \mathbf{x}) = \sum_{i=0}^{N^s} \alpha_i (\mathbf{u}_i^{\text{comp}})^2,$$

where the α_i are the weights of the integral discretization and hence contain the time intervals between two discretization points, see Equation (6.11). The cost functions of multiple compressors or of different cost incurring components must of course be added together to yield the final cost.

For compressors this means that a pressure increase, mediated by the model equation

$$p_{\text{out}}(t) = p_{\text{in}}(t) + \mathbf{u}(t),$$

must be balanced with the cost of the usage.

Control valves on the other hand are controlled, but in our model incur no cost. This means that they can be used freely in order to satisfy inequality constraints or reduce costs of other components.

6.3.2 Pressure-constrained gas nodes

In addition to controlled components that may contribute to the cost function of an optimal control problem, we also have components that

model certain bounds on variables and therefore contribute to the vector \mathbf{h} of inequalities (6.8).

Therefore a gas node may demand a certain pressure range, specifically it may demand

$$p_{\min}(t) \leq p(t) \leq p_{\max}(t), \quad (6.16)$$

where the lower and upper bound must be specified as additional conditions of the optimal control problem. These translate to two inequalities per node.

6.4 DISCRETIZATION

As explained in Section 6.2, the discretization of the control-related equations is different from the model equations that were introduced in Chapter 5.

The controls themselves are discretized with some control step size (which does not have to be uniform, although it is in our examples) and controls at each state time point are evaluated according to Equation (6.6).

The objective function is evaluated at every time point of the state discretization, that is, the discretization of the model equations in Chapter 5.

Lastly constraints like those of the pressure-constrained gas nodes in Equation (6.16) are evaluated at a subset of the state time points. Here we can see, that guaranteeing the inequality constraints at every time point may lead to a great numerical expense, when using the adjoint method.

6.5 A SMALL OPTIMAL CONTROL PROBLEM

6.5.1 *Network properties*

We solve an optimal control problem on a network very similar to that of Section 5.2. The whole network can be seen in Figure 6.1. The only difference is the insertion of a compressor station right before the branching junction at S17.

Table 6.2: Parameters of the gas network.

Pipe	From	To	Length ℓ_e [km]
P10	S4	S20	20.322
P20	S5	S0	20.635
P21	S17	S4	10.586
P22	S17	S8	10.452
P24	S8	S20	19.303
P25	S20	S25	66.037

The gas network is again a small part of GasLib-40 and the pipeline parameters are given in Table 6.2, which is almost the same as Table 5.2,

but the edge between S5 and S17 was moved to the left to make room for the compressor. The parameters for the power grid stay the same, see Table 5.3 and Table 5.4.

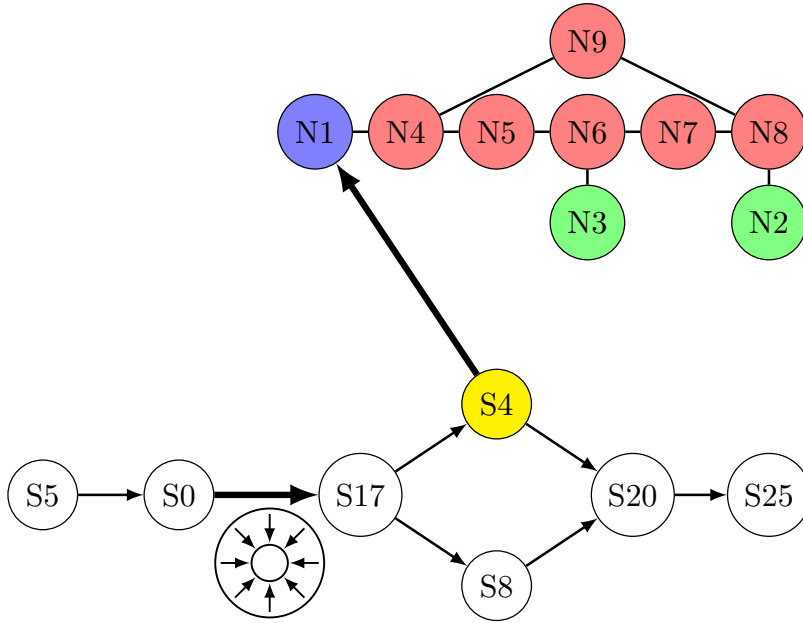


Figure 6.1: Coupled gas-power network with a compressor between S0 and S17.

6.5.2 Numerical results

For this network, we simulate a sudden increase in power demand within the power grid and study its effect on the gas network. The compressor station is supposed to compensate part of the pressure losses in the gas network such that a pressure bound of $p_{\min} = 41$ bar is satisfied at the node S25 at all times, while power consumption of the compressor is minimized.

Our time horizon is 12h and we discretize with a step size of $\Delta t = 15$ min. We leave all power demands and supplies in the power network at their initial settings in Table 5.4, except for the demand of node N5, which increases its active and reactive power demand sharply after one hour, exactly as in Section 5.2.3, see Figure 5.2.

To guarantee the pressure bound of 41 bar at the outflow node of the gas network we must use the compressor, as can be seen in Figure 6.2. We see that the pressure bound is violated after approximately 4 hours, unless we employ the compressor, which, due to its associated cost controls the pressure to just barely satisfy the bound.

6.6 A LARGE OPTIMAL CONTROL PROBLEM

Finally, we show results of an optimization task on a large network.

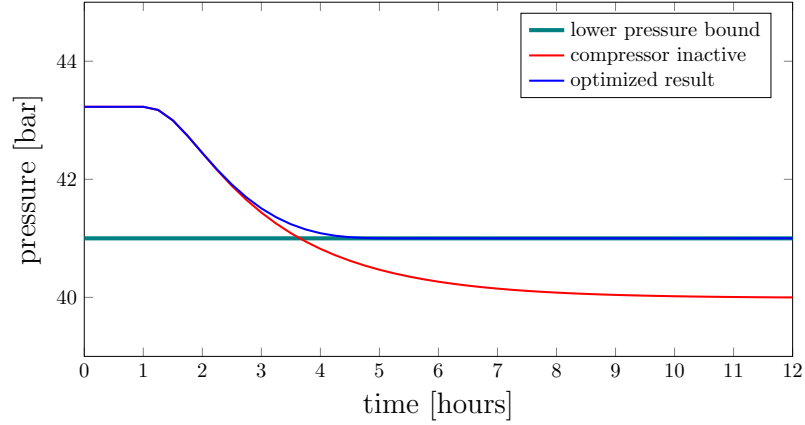


Figure 6.2: Pressure for active and in-active compressor.

6.6.1 Network properties

As underlying network we use that of Section 5.4, whose gas and power network parts are pictured in Figures 5.5 and 5.6, respectively. The GasLib-134 network actually contains two controllable components, a compressor between the nodes 29/30, slightly to the lower left of the top-most (green) source in Figure 5.5 and a control valve between the nodes 65/66, in the “lower right leg” at the bottom of Figure 5.5. Both of these were inactive in the preceding simulation scenarios. The power network is operated at deterministic conditions, that is with the stochastic parameter σ set to zero, as already done in Section 5.4.4.

6.6.2 Numerical results

We take the steady state initial conditions, also used in Section 5.4.4 as a starting point but add two continuous constraints in order to make the compressor and the valve actually do some work. At the sink `ld_22`, which is located at the end of the “right arm” in the middle of Figure 5.5, we impose a lower pressure bound of $p_{\min}(t = 0) = 70$ bar and $p_{\min}(t = 24 \text{ h}) = 90$ bar and interpolate linearly in between. At this sink the upper bound p_{\max} is set to infinity.

In addition we impose an upper pressure bound at sink `ld_40` of 90 bar at $t = 0$, 70 bar at $t = 24 \text{ h}$ and again interpolate linearly in between. This sink is located in the “lower right leg” at the bottom of Figure 5.5.

As mentioned, using the valve is free but compressor costs should be minimized. The control is discretized with the same discretization already used by the states, yielding 49 time points. Also the constraints are evaluated at every state time step. While our algorithm is capable of using coarser resolutions of both constraints and controls, the problem at hand is again small enough to compute a solution in acceptable time³. The control of the valve is constrained to not exceed 40 bar to keep the optimization routine from trying controls that are

³ On a work station finding a minimum of this problem took approximately a minute.

too high to yield a solution of the simulation. The compressor control is capped at 120 bar, although this bound is never attained.

With this data we compute the optimal controls shown in Figure 6.3 and Figure 6.4. The compressor control in Figure 6.3 ramps up as the lower pressure bound in ld_{22} rises. On the other hand, the valve control in Figure 6.4 stays at zero until this is not sufficient anymore to satisfy the decreasing upper pressure bound in ld_{40} at which point the control rises up to the maximal value, staying there until the end. As the valve control incurs no cost, this is one of many possible configurations.

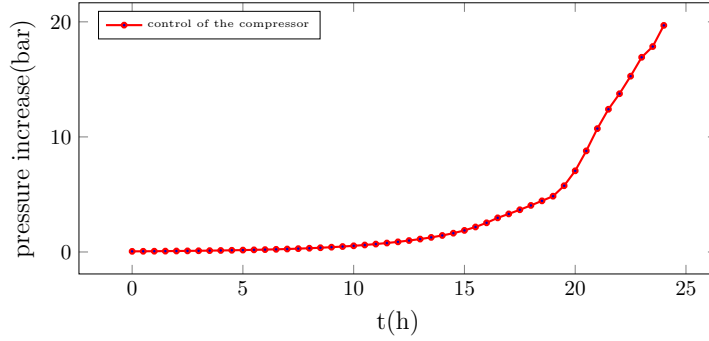


Figure 6.3: Computed optimal control of the compressor at nodes 29/30.

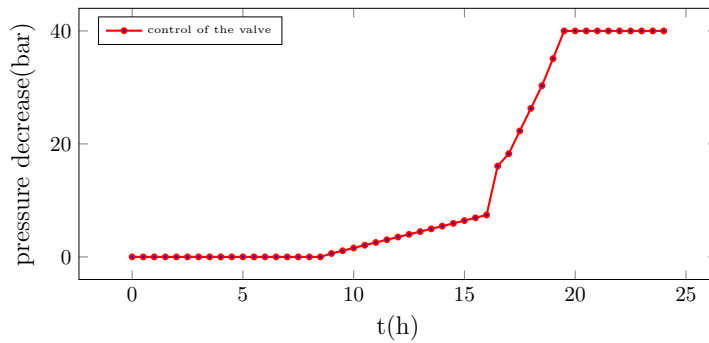


Figure 6.4: Computed optimal control of the valve at nodes 65/66.

A comparison of pressure evolution at the two sinks ld_{22} and ld_{40} is given in Figure 6.5 and Figure 6.6. It can be seen that the compressor increases the pressure just enough to satisfy the lower pressure bound as its usage is penalized, while the (free to use) valve at first matches the upper pressure bound exactly but later on over-compensates it rather strongly.

6.7 SUMMARY AND OUTLOOK

In this chapter, we introduced the direct and adjoint methods for computing derivatives of objective and constraint functions, as well as refinements of both for optimization problems discretized in time. Two optimal control problems were solved by computing the needed derivatives via the adjoint method.

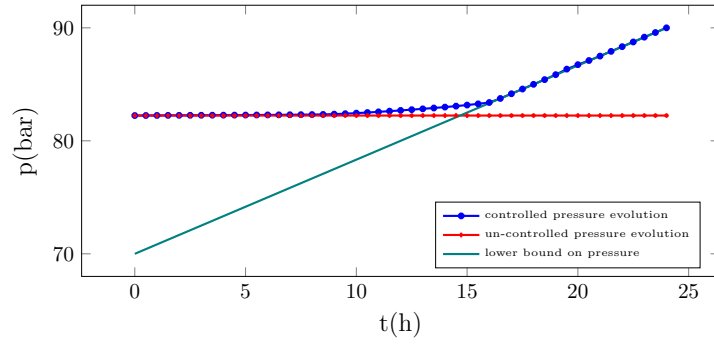


Figure 6.5: Comparison of controlled and uncontrolled pressure at ld_22.

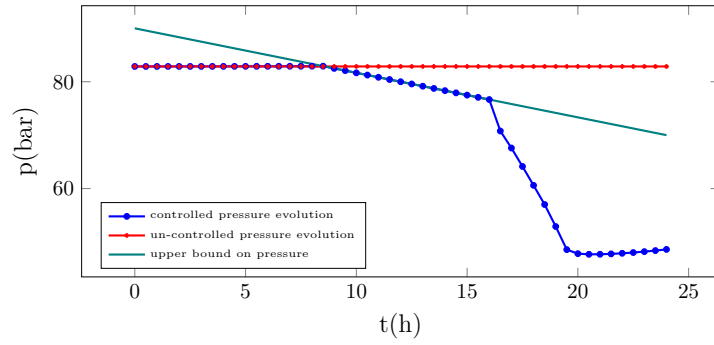


Figure 6.6: Comparison of controlled and uncontrolled pressure at ld_40.

Possible extensions include optimization in the power network itself in the sense of optimal power flow, see for example [Eng+19; Guo+19; MFH18; Müh+19; XA18] as well as the computation of second derivatives as mentioned in the beginning of this chapter.

SUMMARY AND CONCLUSION

In this thesis we have investigated the mathematical properties of power networks and gas networks as well as the coupling between them.

In Chapter 3 we have analyzed two models for electricity networks, on the one hand the Telegrapher's equations and on the other the power flow equations. We showed with the help of Lyapunov stability theory, that for periodic boundary conditions the former model converges to a periodic solution (Section 3.1) which in turn can be described with the help of the power flow equations (Section 3.2). In addition, we examined a second order numerical method for the solution of the Telegrapher's equations on a network and showed that it mimics the Lyapunov stability of the analytical solution (Section 3.3). The numerical experiments in Section 3.4 provided justification for choosing the power flow equations over the Telegrapher's equation model.

In Chapter 4 we examined the isentropic Euler equations and provided conditions on the pressure function, that guarantee well-posedness of the corresponding Riemann problems for all sub-sonic initial conditions (Section 4.1.1). Additionally, we showed that the usual family of γ -laws with $1 \leq \gamma \leq 3$ can be extended to $-1 < \gamma \leq 3$ while still guaranteeing well-posedness of these Riemann problems, but also that for any γ -law with $\gamma \notin [-1, 3]$ there are sub-sonic initial conditions that define an ill-posed Riemann problem. The findings were extended to Riemann problems at junctions of a gas network (Section 4.1.3) and finally to junctions with external gas sinks and sources, where a power grid can be interfaced with the gas network (Section 4.2).

Chapter 5 was dedicated to the application of the findings of the previous chapters to coupled gas and power networks. To this end, models for different components of such networks were presented (Section 5.2.1) and discretized, which were then applied to a small network (Section 5.2.3). Then we examined a bigger network, the GasLib-134 network, that also incorporates a compressor station and a control valve (Section 5.3) and compared the simple coupling via equality of pressure to the physically more accurate Bernoulli coupling, finding little difference but much higher implementation effort, which may often not be justified (Sections 5.3.1 and 5.3.4). In this chapter we also introduced a simple model for gas-power conversion with different efficiencies for gas-to-power and power-to-gas conversion (Section 5.3.1). In a numerical study we showcased the repercussions of changing power demand on the gas network in Section 5.3.4.

In the final part Section 5.4 of Chapter 5 we extended the power model further to include uncertain power demands modeled by the

Ornstein–Uhlenbeck process (Section 5.4.1) and compared scenarios of different volatility in the numerical study (Section 5.4.4).

Lastly in Chapter 6 we enhanced both the model and the underlying software *grazer* with optimization capabilities, presenting in detail the direct and adjoint methods for solving constrained non-linear optimal control problems with control and state variables (Section 6.1) as well as certain improvements that can be made by exploiting the decomposition of the problem into time steps (Section 6.2). In numerical studies (Sections 6.5.2 and 6.6.2) the optimization capabilities of our software were applied to both the small network and the GasLib-134 network already examined in the preceding chapter.

All in all we have presented and analyzed models for coupled gas and power networks and applied them to simulation and optimization problems on a network of realistic size. The last distinguishing feature of this work is that the software that was developed during this work is accessible for application and extension by both academic and industrial users.

BIBLIOGRAPHY

- [Age19] D. E. Agency. *Technology Data - Generation of Electricity and District Heating*. Tech. rep. Energistyrelsen (Danish Energy Agency), 2019. URL: https://ens.dk/sites/ens.dk/files/Analyser/technology_data_catalogue_for_el_and_dh.pdf (visited on 02/20/2022).
- [Aïd+09] R. Aïd et al. "A structural risk-neutral model of electricity prices". In: *Int. J. Theor. Appl. Finance* 12.7 (2009), pp. 925–947. ISSN: 0219-0249. URL: <https://doi.org/10.1142/S021902490900552X>.
- [Ama97] D. Amadori. "Initial-boundary value problems for non-linear systems of conservation laws". In: *Nonlinear Differential Equations and Applications NoDEA* 4 (Jan. 1997), pp. 1–42. DOI: 10.1007/PL00001406.
- [And15] G. Andersson. *Lecture Notes on Power System Analysis, Power Flow Analysis, Fault Analysis, Power System Dynamics and Stability*. Sept. 2015.
- [AtG] *Atomgesetz*. Tech. rep. G 5702 Teil I, Nr. 43. Bundesanzeiger Verlag GmbH, Aug. 2011.
- [ATK21] L. Anderson, M. Turner, and T. Koch. *Generative deep learning for decision making in gas networks*. 2021. arXiv: 2102.02125 [math.OC].
- [Bal77] J. M. Ball. "Strongly Continuous Semigroups, Weak Solutions, and the Variation of Constants Formula". In: *Proceedings of the American Mathematical Society* 63.2 (1977), pp. 370–373. ISSN: 00029939, 10886826. URL: <http://www.jstor.org/stable/2041821>.
- [Bar02] M. T. Barlow. "A diffusion model for electricity prices". In: *Math. Finance* 12.4 (2002), pp. 287–298. ISSN: 1467-9965. DOI: 10.1111/j.1467-9965.2002.tb00125.x.
- [BBKo8] F. Benth, J. Benth, and S. Koekebakker. *Stochastic modelling of electricity and related markets*. Vol. 11. Advanced Series on Statistical Science & Applied Probability. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008. ISBN: 978-981-281-230-8. URL: <https://doi.org/10.1142/9789812812315>.
- [BC16] G. Bastin and J. Coron. *Stability and boundary stabilization of 1-D hyperbolic systems*. Progress in Nonlinear Differential Equations and Their Applications. Springer International Publishing, 2016. ISBN: 9783319320625. DOI: 10.1007/978-3-319-32062-5.

- [BCH14] D. Bienstock, M. Chertkov, and S. Harnett. “Chance-Constrained Optimal Power Flow: Risk-Aware Network Control under Uncertainty”. In: *SIAM Review* 56.3 (2014), pp. 461–495. DOI: 10.1137/130910312.
- [Beu+19] D. J. Beulertz et al. “Development of a Modular Framework for Future Energy System Analysis”. In: *2019 54th International Universities Power Engineering Conference (UPEC) : proceedings : 3rd-6th September 2019, Bucharest, Romania / IEEE*. 54th International Universities Power Engineering Conference, Bucharest (Romania). IEEE, Sept. 2019. DOI: 10.1109/UPEC.2019.8893472.
- [Bez+17] J. Bezanson et al. “Julia: A fresh approach to numerical computing”. In: *SIAM review* 59.1 (2017), pp. 65–98. URL: <https://doi.org/10.1137/141000671>.
- [BGH11] J. Brouwer, I. Gasser, and M. Herty. “Gas Pipeline Models Revisited: Model Hierarchies, Nonisothermal Models, and Simulations of Networks”. In: *Multiscale Modeling & Simulation* 9.2 (2011), pp. 601–623. DOI: 10.1137/100813580.
- [BHH16] M. K. Banda, A.-S. Häck, and M. Herty. “Numerical Discretization of Coupling Conditions by High-Order Schemes”. In: *Journal of Scientific Computing* 69.1 (Apr. 2016), pp. 122–145. DOI: 10.1007/s10915-016-0185-x.
- [BHKo6a] M. Banda, M. Herty, and A. Klar. “Coupling conditions for gas networks governed by the isothermal Euler equations”. In: *Networks & Heterogeneous Media* 1 (2006), pp. 295–314. DOI: 10.3934/nhm.2006.1.295.
- [BHKo6b] M. Banda, M. Herty, and A. Klar. “Gas flow in pipeline networks”. In: *Networks & Heterogeneous Media* 1.1 (2006), pp. 41–56. DOI: 10.3934/nhm.2006.1.41.
- [BHS18] T. Brown, J. Hörsch, and D. Schlachtberger. “PyPSA: Python for Power System Analysis”. In: *Journal of Open Research Software* 6.4 (1 2018). DOI: 10.5334/jors.188.
- [Bie15] D. Bienstock. *Electrical transmission system cascades and vulnerability*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2015. DOI: 10.1137/1.9781611974164.
- [BK14] R. Borsche and J. Kall. “ADER schemes and high order coupling on networks of hyperbolic conservation laws”. In: *Journal of Computational Physics* 273 (Sept. 2014), pp. 658–670. DOI: 10.1016/j.jcp.2014.05.042.
- [BMZ15] A. Baeza, P. Mulet, and D. Zorío. “High Order Boundary Extrapolation Technique for Finite Difference Methods on Complex Domains with Cartesian Meshes”. In: *Journal of Scientific Computing* 66 (May 2015). DOI: 10.1007/s10915-015-0043-2.

- [Brao4] H. Brakelmann. *Netzverstärkungs-Trassen zur Übertragung von Windenergie: Freileitung oder Kabel?* Oct. 2004. URL: <https://www.wind-energie.de/fileadmin/redaktion/dokumente/pressemitteilungen/2018/20180103-brak-studie-2004.pdf> (visited on 04/05/2023).
- [Bre+14] A. Bressan et al. “Flow on networks: recent results and perspectives”. In: *European Mathematical Society-Surveys in Mathematical Sciences* 1.1 (2014), pp. 47–111. DOI: 10.4171/EMSS/2.
- [Bro] G. O. Brown. “The History of the Darcy-Weisbach Equation for Pipe Flow Resistance”. In: *Environmental and Water Resources History*, pp. 34–43. DOI: 10.1061/40650(2003)4.
- [Bro+18] T. Brown et al. “Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable European energy system”. In: *Energy* 160 (2018), pp. 720–739. ISSN: 0360-5442. DOI: <https://doi.org/10.1016/j.energy.2018.06.222>.
- [Cap11] P. Caplan. “Numerical computation of second derivatives with applications to optimization problems”. Unpublished academic report. 2011. URL: <http://web.mit.edu/pcaplan/www/SecondDerivative2012.pdf>.
- [CBL15] M. Chertkov, S. Backhaus, and V. Lebedev. “Cascading of fluctuations in interdependent energy infrastructures: Gas-grid coupling”. In: *Applied Energy* 160 (2015), pp. 541–551. DOI: 10.1016/j.apenergy.2015.09.085.
- [CGo8] R. Colombo and M. Garavello. “On the Cauchy Problem for the p-System at a Junction”. In: *SIAM Journal on Mathematical Analysis* 39.5 (2008), pp. 1456–1471. DOI: 10.1137/060665841.
- [CG10] R. M. Colombo and G. Guerra. “On general balance laws with boundary”. In: *Journal of Differential Equations* 248.5 (2010), pp. 1017–1043. ISSN: 0022-0396. DOI: 10.1016/j.jde.2009.12.002.
- [Chao8] K. Chayakulkheeree. “Application of distributed slack bus power flow to competitive environments”. In: *2007 Australasian Universities Power Engineering Conference, AUPEC*. Jan. 2008, pp. 1–6. ISBN: 978-0-646-49488-3. DOI: 10.1109/AUPEC.2007.4548081.
- [CHSo8] R. M. Colombo, M. Herty, and V. Sachers. “On 2×2 Conservation Laws at a Junction”. In: *SIAM Journal on Mathematical Analysis* 40.2 (2008), pp. 605–622. DOI: 10.1137/070690298.
- [CK21] S. Coskun and R. Korn. “Modeling the Intraday Electricity Demand in Germany”. In: *Mathematical Modeling, Simulation and Optimization for Power Engineering and Management*. Feb. 2021, pp. 3–23. ISBN: 978-3-030-62731-7. DOI: 10.1007/978-3-030-62732-4_1.

- [CL16] G. Cerbe and B. Lendt. *Grundlagen der Gastechnik: Gasbeschaffung – Gasverteilung – Gasverwendung*. Carl Hanser Verlag GmbH & Company KG, 2016. ISBN: 9783446449664.
- [Cle+16] T. Clees. et al. “MYNTS: Multi-physics NeTwork Simulator”. In: *Proceedings of the 6th International Conference on Simulation and Modeling Methodologies, Technologies and Applications - SIMULTECH*, INSTICC. SciTePress, 2016, pp. 179–186. ISBN: 978-989-758-199-1.
- [Daf09] C. Dafermos. *Hyperbolic Conservation Laws in Continuum Physics*. Vol. 325. Jan. 2009, p. 626. DOI: 10.1007/978-3-642-04048-1.
- [Egg18] H. Egger. “A robust conservative mixed finite element method for isentropic compressible flow on pipe networks”. In: *SIAM Journal on Scientific Computing* 40.1 (2018), A108–A129. DOI: 10.1137/16M1094373.
- [EK17] H. Egger and T. Kugler. “Damped wave systems on networks: exponential stability and uniform approximations”. In: *Numerische Mathematik* 138.4 (Oct. 2017), pp. 839–867. DOI: 10.1007/s00211-017-0924-4.
- [EN01] K.-J. Engel and R. Nagel. “One-Parameter Semigroups for Linear Evolution Equations”. In: *Semigroup Forum* 63 (June 2001), pp. 278–280. DOI: 10.1007/s002330010042.
- [Eng+19] A. Engelmann et al. “Toward distributed OPF using ALADIN”. In: *IEEE Transactions on Power Systems* 34.1 (Jan. 2019), pp. 584–594. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2018.2867682.
- [ES05] K. Ehrhardt and M. Steinbach. “Nonlinear Optimization in Gas Networks”. In: *Modeling, Simulation and Optimization of Complex Processes*. Dec. 2005, pp. 139–148. ISBN: 3-540-23027-0. DOI: 10.1007/3-540-27170-8_11.
- [Eva10] L. Evans. *Entropy and Partial Differential Equations, Lecture Notes*. Mar. 2010.
- [FG21] E. Fokken and S. Göttlich. *On the relation of powerflow and Telegrapher’s equations: continuous and numerical Lyapunov stability*. 2021. arXiv: 2101.12662 [math.NA].
- [FGH22] E. Fokken, S. Göttlich, and M. Herty. “Efficient simulation of coupled gas and power networks under uncertain demands”. In: *European Journal of Applied Mathematics* (2022). Cambridge University Press, pp. 1–27. DOI: 10.1017/S0956792522000079.
- [FGK19] E. Fokken, S. Göttlich, and O. Kolb. “Modeling and simulation of gas networks coupled to power grids”. In: *Journal of Engineering Mathematics* (Nov. 2019). Springer Netherlands. ISSN: 1573-2703. DOI: 10.1007/s10665-019-10026-6.

- [FGK20] E. Fokken, S. Göttlich, and O. Kolb. “Optimal Control of Compressor Stations in a Coupled Gas-to-Power Network”. In: *Advances in Energy System Optimization*. Ed. by V. Bertsch et al. Cham: Springer International Publishing, 2020, pp. 67–80. ISBN: 978-3-030-32157-4. DOI: 10.1007/978-3-030-32157-4_5.
- [FMT12] U. S. Fjordholm, S. Mishra, and E. Tadmor. “Arbitrarily High-order Accurate Entropy Stable Essentially Nonoscillatory Schemes for Systems of Conservation Laws”. In: *SIAM Journal on Numerical Analysis* 50.2 (2012), pp. 544–573. DOI: 10.1137/110836961.
- [Fok+21] E. Fokken et al. “Modeling and Simulation of Sector-Coupled Energy Networks: A Gas-Power Benchmark”. In: *Mathematical Modeling, Simulation and Optimization for Power Engineering and Management*. Ed. by S. Göttlich, M. Herty, and A. Milde. Cham: Springer International Publishing, 2021, pp. 263–284. ISBN: 978-3-030-62732-4. DOI: 10.1007/978-3-030-62732-4_12.
- [FR16] S. Frank and S. Rebennack. “An introduction to optimal power flow: Theory, formulation, and examples”. In: *IIE Transactions* 48.12 (2016), pp. 1172–1197. DOI: 10.1080/0740817X.2016.1189626.
- [GH19] S. Gerster and M. Herty. “Discretized feedback control for systems of linearized hyperbolic balance laws”. In: *Math. Control Relat. Fields* 9.3 (2019), pp. 517–539. ISSN: 2156-8472. DOI: 10.3934/mcrf.2019024.
- [GHM17] M. Gugat, M. Herty, and S. Müller. “Coupling conditions for the transition from supersonic to subsonic fluid states”. In: *Networks and Heterogeneous Media* 12.3 (2017), pp. 371–380. DOI: 10.3934/nhm.2017016.
- [GHM21] S. Göttlich, M. Herty, and A. Milde. *Mathematical Modeling, Simulation and Optimization for Power Engineering and Management*. Ed. by S. Göttlich, M. Herty, and A. Milde. Mathematics in Industry. Springer International Publishing, 2021. ISBN: 978-3-030-62732-4. DOI: 10.1007/978-3-030-62732-4.
- [GHS16] S. Göttlich, M. Herty, and P. Schillen. “Electric transmission lines: Control and numerical discretization”. In: *Optim. Control Appl. Meth.* 37.5 (2016), pp. 980–995. DOI: 10.1002/oca.2219.
- [GJ+10] G. Guennebaud, B. Jacob, et al. *Eigen v3*. 2010. URL: <http://eigen.tuxfamily.org>.
- [GKL19] S. Göttlich, R. Korn, and K. Lux. “Optimal control of electricity input given an uncertain demand”. In: *Mathematical Methods of Operations Research* (Aug. 2019). ISSN: 1432-5217. DOI: 10.1007/s00186-019-00678-6.

- [GKL21] S. Göttlich, O. Kolb, and K. Lux. “Chance-constrained optimal inflow control in hyperbolic supply systems with uncertain demand”. In: *Optimal Control Applications and Methods* 42.2 (2021), pp. 566–589. DOI: <https://doi.org/10.1002/oca.2689>.
- [GPR18] M. Gugat, V. Perrollaz, and L. Rosier. “Boundary stabilization of quasilinear hyperbolic systems of balance laws: exponential decay for small source terms”. In: *Journal of Evolution Equations* 18.3 (May 2018), pp. 1471–1500. DOI: 10.1007/s00028-018-0449-z.
- [GPT19] S. Göttlich, A. Potschka, and C. Teuber. “A partial outer convexification approach to control transmission lines”. In: *Comput. Optim. Appl.* 72.2 (2019), pp. 431–456. ISSN: 0926-6003. DOI: 10.1007/s10589-018-0047-6.
- [GR14] E. Godlewski and P.-A. Raviart. *Numerical Approximation of Hyperbolic Systems of Conservation Laws*. Springer Publishing Company, Incorporated, 2014. ISBN: 9781461207146.
- [GS01] J. D. D. Glover and M. S. Sarma. *Power System Analysis and Design*. 3rd. USA: Brooks/Cole Publishing Co., 2001. ISBN: 0534953670.
- [GS17] S. Goettlich and P. Schillen. “Numerical Discretization of Boundary Control Problems for Systems of Balance Laws: Feedback Stabilization”. In: *European Journal of Control* 35 (Feb. 2017). DOI: 10.1016/j.ejcon.2017.02.002.
- [GS96] S. Gottlieb and C.-W. Shu. “Total Variation Diminishing Runge-Kutta Schemes.” In: *Mathematics of Computation* 67 (Aug. 1996), pp. 73–85. DOI: 10.1090/S0025-5718-98-00913-2.
- [GSC16] J. Grainger, W. Stevenson, and G. Chang. *Power System Analysis*. McGraw-Hill series in electrical and computer engineering: Power and energy. McGraw-Hill Education, 2016. ISBN: 9781259 008351.
- [GT18] S. Göttlich and C. Teuber. “Space mapping techniques for the optimal inflow control of transmission lines”. In: *Optim. Methods Softw.* 33.1 (2018), pp. 120–139. ISSN: 1055-6788. DOI: 10.1080/10556788.2016.1278542.
- [Gug+12] M. Gugat et al. “Well-posedness of networked hyperbolic systems of balance laws”. In: *Constrained Optimization and Optimal Control for Partial Differential Equations*. Ed. by G. Leugering et al. Basel: Springer Basel, 2012, pp. 123–146. ISBN: 978-3-0348-0133-1. DOI: 10.1007/978-3-0348-0133-1_7.
- [Gug14] M. Gugat. “Boundary feedback stabilization of the telegraph equation: decay rates for vanishing damping term”. In: *Systems Control Lett.* 66 (2014), pp. 72–84. ISSN: 0167-6911. DOI: 10.1016/j.sysconle.2014.01.007.

- [Guo+19] Y. Guo et al. “Data-Based Distributionally Robust Stochastic Optimal Power Flow – Part I: Methodologies”. In: *IEEE Transactions on Power Systems* 34.2 (Mar. 2019), pp. 1483–1492. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2018.2878385.
- [GZ19] V. Gyrya and A. Zlotnik. “An explicit staggered-grid method for numerical simulation of large-scale natural gas pipeline networks”. In: *Applied Mathematical Modelling* 65 (2019), pp. 34–51. DOI: <https://doi.org/10.1016/j.apm.2018.07.051>.
- [HBO16] S. Heinen, D. Burke, and M. O’Malley. “Electricity, gas, heat integration via residential hybrid heating technologies – An investment model assessment”. In: *Energy* 109 (2016), pp. 906–919. ISSN: 0360-5442. DOI: 10.1016/j.energy.2016.04.126.
- [HDS13] K. Heuck, K. Dettmann, and D. Schultz. *Elektrische Energieversorgung*. Springer Vieweg, 2013. ISBN: 978-3-8348-1699-3. DOI: 10.1007/978-3-8348-2174-4.
- [Hero7] M. Herty. “Modeling, simulation and optimization of gas networks with compressors”. In: *Networks and Heterogeneous Media* 2.1 (2007), pp. 81–97. DOI: 10.3934/nhm.2007.2.81.
- [HHW20] Y. Holle, M. Herty, and M. Westdickenberg. *New Coupling Conditions for Isentropic Flow on Networks*. 2020. arXiv: 2004.09184 [math.AP].
- [HMS19] M. Herty, S. Müller, and A. Sikstel. “Coupling of compressible Euler equations”. In: *Vietnam Journal of Mathematics* (June 2019). ISSN: 2305-2228. DOI: 10.1007/s10013-019-00353-7.
- [HR02] H. Holden and N. Risebro. *Front Tracking for Hyperbolic Conservation Laws*. Springer Berlin, 2002. ISBN: 978-3-642-62797-2.
- [IEA19] IEA. *Tracking Energy Integration*. Tech. rep. Paris, France: International Energy Agency (IEA), 2019. URL: <https://www.iea.org/reports/tracking-energy-integration>.
- [Jer+20] M. Jereminov et al. “Evaluating Feasibility Within Power Flow”. In: *IEEE Transactions on Smart Grid* 11.4 (2020), pp. 3522–3534. DOI: 10.1109/TSG.2020.2966930.
- [KBoo] B. Kim and R. Baldick. “A comparison of distributed optimal power flow algorithms”. In: *IEEE Transactions on Power Systems* 15.2 (2000), pp. 599–604. DOI: 10.1109/59.867147.
- [KLB10] O. Kolb, J. Lang, and P. Bales. “An implicit box scheme for subsonic compressible flow with dissipative source term”. In: *Numer. Algorithms* 53.2 (2010), pp. 293–307. DOI: 10.1007/s11075-009-9287-y.

- [KMNo3] M. Kramar, D. Mugnolo, and R. Nagel. "Semigroups for Initial-Boundary Value Problems". In: *Evolution Equations: Applications to Physics, Industry, Life Sciences and Economics*. Ed. by M. Iannelli and G. Lumer. Basel: Birkhäuser Basel, 2003, pp. 275–292. ISBN: 978-3-0348-8085-5. DOI: 10.1007/978-3-0348-8085-5_19.
- [Kol11] O. Kolb. "Simulation and Optimization of Gas and Water Supply Networks". PhD thesis. TU Darmstadt, 2011.
- [Kol14] O. Kolb. "On the full and global accuracy of a compact third order WENO scheme". In: *SIAM Journal on Numerical Analysis* 52.5 (2014), pp. 2335–2355. ISSN: 0036-1429. DOI: 10.1137/130947568.
- [Kre07] E. Kreyszig. *Introductory Functional Analysis with Applications*. Wiley classics library. Wiley India Pvt. Limited, 2007. ISBN: 9788126511914.
- [KSB09] R. Kiesel, G. Schindlmayr, and R. H. Börger. "A two-factor model for the electricity forward market". In: *Quant. Finance* 9.3 (2009), pp. 279–287. ISSN: 1469-7688. URL: <https://doi.org/10.1080/14697680802126530>.
- [LeVo2] R. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2002. ISBN: 9780521009249.
- [LeV92] R. LeVeque. *Numerical methods for conservation laws*. Lectures in Mathematics ETH Zürich, Department of Mathematics Research Institute of Mathematics. Springer, 1992. ISBN: 9783764327231.
- [Low14a] S. Low. "Convex relaxation of optimal power flow—Part I: Formulations and equivalence". In: *IEEE Transactions on Control of Network Systems* 1.1 (2014), pp. 15–27. DOI: 10.1109/TCNS.2014.2309732.
- [Low14b] S. Low. "Convex relaxation of optimal power flow—Part II: Exactness". In: *IEEE Transactions on Control of Network Systems* 1.2 (2014), pp. 177–189. DOI: 10.1109/TCNS.2014.2323634.
- [LPL19] H. Li, L. Pan, and Q. Liu. "A Linear Power Flow Solution for Distribution Power System Including PV Bus and ZIP Load". In: *Journal of Electrical Engineering & Technology* 14 (July 2019). DOI: 10.1007/s42835-019-00222-w.
- [LS02] J. J. Lucia and E. S. Schwartz. "Electricity Prices and Power Derivatives: Evidence from the Nordic Power Exchange". In: *Review of Derivatives Research* 5.1 (2002), pp. 5–50. ISSN: 1573-7144. DOI: 10.1023/A:1013846631785.
- [LS80] T.-P. Liu and J. Smoller. "On the vacuum state for the isentropic gas dynamics equations". In: *Advances in Applied Mathematics* 1.4 (1980), pp. 345–359. ISSN: 0196-8858. DOI: 10.1016/0196-8858(80)90016-0.

- [Mat+17] C. Matke et al. “Structure analysis of the German transmission network using the open source model SciGRID”. In: *Advances in Energy System Optimization*. Springer, 2017, pp. 177–188. URL: <http://www.scigrid.de/>.
- [Men15] E. Menon. *Transmission Pipeline Calculations and Simulations Manual*. Jan. 2015. ISBN: 9781856178303.
- [MFH18] T. Mühlpfordt, T. Faulwasser, and V. Hagenmeyer. “A Generalized Framework for Chance-Constrained Optimal Power Flow”. In: *Sustainable Energy, Grids and Networks* 16 (2018), pp. 231–242. ISSN: 2352-4677. DOI: <https://doi.org/10.1016/j.segan.2018.08.002>.
- [MMo1] A. Maffucci and G. Miano. *Transmission lines and lumped circuits*. Jan. 2001. ISBN: 0121897109. DOI: 10.1016/B978-0-12-189710-9.X5000-1.
- [Mol+17] D. Molzahn et al. “A survey of distributed optimization and control algorithms for electric power systems”. In: *IEEE Transactions on Smart Grid* 8.6 (2017), pp. 2941–2962. DOI: 10.1109/TSG.2017.2720471.
- [Müh+19] T. Mühlpfordt et al. “Chance-Constrained AC Optimal Power Flow – A Polynomial Chaos Approach”. In: *IEEE Transactions on Power Systems* 34.6 (2019), pp. 4806–4816. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2019.2918363.
- [Nic16] S. Nicaise. “Control and stabilization of 2×2 hyperbolic systems on graphs”. In: *Mathematical Control and Related Fields* 7 (Dec. 2016), pp. 53–72. DOI: 10.3934/mcrf.2017004.
- [NKS18] A. Naumann, O. Kolb, and M. Semplice. “On a third order CWENO boundary treatment with application to networks of hyperbolic conservation laws”. In: *Applied Mathematics and Computation* 325 (2018), pp. 252–270. DOI: <https://doi.org/10.1016/j.amc.2017.12.041>.
- [NTS19] NTS. *NTS linepack*. Tech. rep. NationalGrid UK, 2019. URL: <https://www.nationalgrid.com/uk/gas-transmission/balancing/nts-linepack> (visited on 02/20/2022).
- [NWo6] J. Nocedal and S. J. Wright. *Numerical Optimization*. 2e. New York, NY, USA: Springer, 2006. ISBN: 978-0387-30303-1.
- [OMa+20] C. O’Malley et al. “Natural gas system dispatch accounting for electricity side flexibility”. In: *Electric Power Systems Research* 178 (2020), p. 106038. DOI: 10.1016/j.epsr.2019.106038.
- [ONe14] M. E. O’Neill. *PCG: A Family of Simple Fast Space-Efficient Statistically Good Algorithms for Random Number Generation*. Tech. rep. HMC-CS-2014-0905. Claremont, CA: Harvey Mudd College, Sept. 2014. URL: <https://www.pcg-random.org/>.

- [Pfe+15] M. Pfetsch et al., eds. *Evaluating gas network capacities*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2015. DOI: 10.1137/1.9781611973693.
- [PGo8] D. I. Papadimitriou and K. C. Giannakoglou. "Direct, adjoint and mixed approaches for the computation of Hessian in airfoil design problems". In: *International Journal for Numerical Methods in Fluids* 56.10 (2008), pp. 1929–1943. DOI: <https://doi.org/10.1002/flid.1584>.
- [Pra+16] S. Prasad et al. "Blackout: Its Causes and its Prevention". In: *International Journal Of Engineering Research & Technology (IJERT)* 4 (2 2016). DOI: 10.17577/IJERTCONV4IS02017.
- [Rei14] G. Reigstad. "Numerical network models and entropy principles for isothermal junction flow". In: *Networks and Heterogeneous Media* 9.1556-1801 (2014), p. 65. ISSN: 1556-1801. DOI: 10.3934/nhm.2014.9.65.
- [Rei15] G. Reigstad. "Existence and uniqueness of solutions to the generalized Riemann problem for isentropic flow". In: *SIAM Journal on Applied Mathematics* 75.2 (2015), pp. 679–702. DOI: 10.1137/140962759.
- [Rud87] W. Rudin. *Real and Complex Analysis*. Mathematics series. McGraw-Hill, 1987. ISBN: 9780071002769.
- [Sch+15] S. Schiebahn et al. "Power to gas: Technological overview, systems analysis and economic assessment for a case study in Germany". In: *International Journal of Hydrogen Energy* 40.12 (2015), pp. 4285–4294. DOI: 10.1016/j.ijhydene.2015.01.123.
- [Sch+17] M. Schmidt et al. "GasLib—A Library of Gas Network Instances". In: *Data* 2.4 (2017). ISSN: 2306-5729. DOI: 10.3390/data2040040.
- [SJ76] P. K. Swamee and A. K. Jain. "Explicit Equations for Pipe-Flow Problems". In: *Journal of the Hydraulics Division* 102.5 (1976), pp. 657–664. DOI: 10.1061/JYCEAJ.0004542.
- [SM96] Y. Saito and T. Mitsui. "Stability Analysis of Numerical Schemes for Stochastic Differential Equations". In: *SIAM Journal on Numerical Analysis* 33.6 (1996), pp. 2254–2267. ISSN: 00361429. URL: <http://www.jstor.org/stable/2158468>.
- [SSoo] E. Schwartz and J. E. Smith. "Short-Term Variations and Long-Term Dynamics in Commodity Prices". In: *Management Science* 46.7 (2000), pp. 893–911. ISSN: 0025-1909. DOI: 10.1287/mnsc.46.7.893.12034.
- [Stro6] K. Strunz. "Developing benchmark models for studying the integration of distributed energy resources". In: *2006 IEEE Power Engineering Society General Meeting*. 2006, 2 pp. DOI: 10.1109/PES.2006.1709568.

- [TA14] D. Trimis and S. Anger. "Potenzial der thermisch integrierten Hochtemperaturelektrolyse und Methanisierung für die Energiespeicherung durch Power-to-Gas (PtG)". In: *gfw Gas* 155.1-2 (2014), pp. 50–59.
- [Tad87] E. Tadmor. "The numerical viscosity of entropy stable schemes for systems of conservation laws". In: *Mathematics of Computation* 49.179 (1987), pp. 91–103. DOI: 10.2307/2008251.
- [TM22] D. Turizo and D. K. Molzahn. "Invertibility Conditions for the Admittance Matrices of Balanced Power Systems". In: *IEEE Transactions on Power Systems* (2022), pp. 1–12. DOI: 10.1109/tpwrs.2022.3206285.
- [Ulbo2] S. Ulbrich. *Optimal Control of Nonlinear Hyperbolic Conservation Laws with Source Terms*. Feb. 2002.
- [Wag14] A. Wagner. "Residual Demand Modeling and Application to Electricity Pricing". In: *The Energy Journal* Volume 35. Number 2 (2014), pp. 45–73. DOI: 10.5547/01956574.35.2.3.
- [WB06] A. Wächter and L. Biegler. "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming". In: *Mathematical Programming* 106.1 (Mar. 2006), pp. 25–57. ISSN: 1436-4646. DOI: 10.1007/s10107-004-0559-y.
- [XA18] W. Xie and S. Ahmed. "Distributionally Robust Chance Constrained Optimal Power Flow with Renewables: A Conic Reformulation". In: *IEEE Transactions on Power Systems* 33.2 (Mar. 2018), pp. 1860–1867. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2017.2725581.
- [ZA96] J. Zhou and M. Adewumi. "Simulation of transients in natural gas pipelines". English (US). In: *SPE Production and Facilities* 11.4 (Nov. 1996), pp. 202–207. DOI: 10.1109/ICCIS.2010.11.
- [Zab19] A. Zablocki. *Fact Sheet: Energy Storage*. Tech. rep. Environmental and Energy Study Institute (EESI), 2019. URL: <https://www.eesi.org/papers/view/energy-storage-2019> (visited on 02/20/2022).
- [Zen+16] Q. Zeng et al. "Steady-state analysis of the integrated natural gas and electric power system with bi-directional energy conversion". In: *Applied Energy* 184.C (2016), pp. 1483–1492. URL: <https://EconPapers.repec.org/RePEc:eee:appene:v:184:y:2016:i:c:p:1483-1492>.
- [Zlo+16] A. Zlotnik et al. "Control policies for operational coordination of electric power and natural gas transmission systems". In: *2016 American Control Conference (ACC)*. July 2016, pp. 7478–7483. DOI: 10.1109/ACC.2016.7526854.

- [Zlo+17] A. Zlotnik et al. "Coordinated Scheduling for Interdependent Electric Power and Natural Gas Infrastructures". In: *IEEE Transactions on Power Systems* 32.1 (2017), pp. 600–610. DOI: 10.1109/TPWRS.2016.2545522.
- [ZMT11] R. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas. "MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education". In: *IEEE Transactions on Power Systems* 26.1 (Feb. 2011), pp. 12–19. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2010.2051168.

APPENDIX: ON THE IMPLEMENTATION OF GRAZER

A.1 OVERVIEW

For most of the computations in this thesis we use the network simulation and optimization tool *grazer*¹ written specifically for this purpose. In this appendix we give a broad overview of *grazer*. More on *grazer* can be found in its documentation, which is distributed along with its source code.

grazer is an open source software suite developed by the author at the Chair of Scientific Computing at the University of Mannheim. For the purpose of long-term usability the following design goals have been chosen:

- Easy installation
- Full C++17-standard compliance with tested support for compilers GCC-9, Clang-9 and later versions thereof.
- Few external dependencies
- High test coverage
- Clean warning profile
- Open Source License (MIT)

grazer depends on external code, namely on Eigen, ([GJ+10]), N. Lohmann's json library², googletest³, pcg-random ([ONe14]), IPOPT ([WBo6]) and CLI11⁴.

Support for MSVC++ was dropped, when optimization was incorporated, as IPOPT is more difficult to use in this case, yet it is planned to re-support MSVC++ again.

A.2 INSTALLATION

In order to build *grazer* on a UNIX-like operating system (for example Linux or MacOS) you need four pieces of software: CMake⁵, Git⁶ and a C++17 capable C++ compiler, e.g. clang⁷ or gcc⁸. The fourth requirement is installation of IPOPT. If you don't supply an IPOPT installation the script coinbrew⁹ is automatically called from CMake

¹ <https://github.com/eike-fokken/grazer>

² <https://github.com/nlohmann/json>

³ <https://github.com/google/googletest>

⁴ <https://github.com/CLIUtils/CLI11>

⁵ <https://cmake.org/>

⁶ <https://git-scm.com/>

⁷ <https://clang.llvm.org/>

⁸ <https://gcc.gnu.org/>

⁹ <https://coin-or.github.io/coinbrew/>

when building *grazer*. In this case, all dependencies of coinbrew must also be installed.

Installation of *grazer* can be done by executing

```
git clone https://github.com/eike-fokken/grazer.git
cd grazer
git submodule update --init --recursive --depth=1
cmake -DCMAKE_BUILD_TYPE=Release -S . -B release
```

Afterwards there is a *grazer* binary in `.../grazer/release/src/Grazer` called *grazer*.

A.3 USAGE

Up to now *grazer* is a command-line application usable from a shell, that is controlled by a number of input json files. In the future it is planned to also support a python interface.

grazer is used by pointing it to a directory with input json files.

```
grazer run data/one_pipeline
```

will run the problem defined in the directory `data/one_pipeline` for example. The problem directory contains a directory `problem`, which holds the json files

```
problem_data.json,
topology.json,
boundary.json,
initial.json
control.json.
```

For optimization problems, the files

```
upper_bounds.json,
lower_bounds.json,
constraint_upper_bounds.json,
constraint_lower_bounds.json
```

are also needed to provide box constraints for the controls and for the constraints as in Equation (6.16). Note that the layout of `topology.json` was heavily inspired by the layout of GasLib files, see [Sch+17]. It is probably best to browse through the data directory of the *grazer* git repository to get an understanding of these files.

After solving the problem, an output directory will be generated in `data/one_pipeline/output`. This again a json file, so it can be read with almost all software. For ease of use, some helper programs, compiled alongside *grazer*, can be found in `release/helper_functions/`. For example calling

```
generate_printing_csv \
data/one_pipeline/output/<output_directory>/states.json \
p_br1
```

will extract the json data into a csv file for usage with plotting tools.

In addition, json schemas can be generated and inserted into the jsons (up to now with the exception of `problem_data.json`) with

```
grazer schema make-full-factory data/one_pipeline
grazer schema insert-key data/one_pipeline
```

This has the advantage that json-aware editors help the users to only write jsons that are valid inputs for *grazer* which cuts down on bug searches.

As a final note on the usage, be aware that although *grazer* runs only sequentially, the output directory names are chosen “atomically”, meaning that two instances of *grazer* running in parallel will not interfere with each others output. This is especially useful when executing many runs of stochastic problems in a Monte-Carlo method as was done in Section 5.4.

Note that parts of the API are still subject to change. For an up-to-date explanation check out the user guide in docs/userguide.tex in the repository.

A.4 A ROUGH OVERVIEW OF THE INNER WORKINGS

On execution, *grazer* will read the input files, configure the Newton solver according to settings in `problem_data.json`, construct a representation of the network from `topology.json`, set initial and boundary values from the respective files, configure the optimization, set constraints and then start solving the problem.

For simulations, the problem is solved time step per time step. In each time step the model equations described in Chapter 5 and their derivatives are evaluated to find a solution of them with Newton’s method. Note that therefore along with the model equations themselves, their derivatives are implemented in *grazer*. If Newton’s method yields a solution, it is saved and the next time step is started. If no solution can be found, the user is notified and all data computed in prior time steps is written to the output files. If all time steps can be solved, all data is written out.

For optimizations on the other hand, each optimization iteration contains a full simulation run. After this run, the derivatives for IPOPT are computed according to Chapter 6.

If a stochastic component is present in the network, a pseudo random number generator must be initialized with a seed. These are generated automatically or taken from the `boundary.json` file, if a seed is present in there.

Note that it is up to now not supported to use stochastic components in an optimization problem.