TU WIEN TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

On Decay Rates in Linear Kinetic Equations with Defects

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften unter der Leitung von

Univ.Prof. Dipl.-Ing. Dr.techn. Anton Arnold

E101 - Institut für Analysis und Scientific Computing, TU Wien

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformation

von

Dipl.-Ing. Tobias Wöhrer

Matrikelnummer: 0725539

Die Dissertation haben begutachtet:

Univ.Prof. Dr. Anton Arnold
Institut für Analysis und Scientific Computing, TU Wien

Univ.-Prof. Dr. Shi Jin
Institute of Natural Sciences, Shanghai Jiao Tong Universität

Univ.-Prof. Dr. Christian Schmeiser
Fakultät für Mathematik, Universität Wien

Wien, am 10. Mai 2020.

Kurzfassung

Die vorliegende Arbeit widmet sich der Analyse des Langzeitverhaltens von Lösungen linearer kinetischer Gleichungen mit *Defekten*. Dabei stehen zwei Modelle im Mittelpunkt: Die degenerierte Fokker–Planck Gleichung und die Goldstein–Taylor Gleichung (ein Transport-Relaxationsmodell von BGK Typ), welche beide hypokoerzive Dynamiken vorweisen. Unser spezieller Fokus sind die "defekten Fälle" dieser Modelle. Die Terminologie orientiert sich hierbei an endlichdimensionalen gewöhnlichen Differentialgleichungen (GDG) mit ähnlichem Verhalten. Dort impliziert eine nicht diagonalisierbare lineare Systemmatrix, im Englischen als "*defective matrix*" bezeichnet, ein Abklingverhalten, das einem exponentiellen Term multipliziert mit einem Polynom entspricht. Um explizite Abschätzungen für das Langzeitverhalten von Lösungen der genannten Gleichungen zu erlangen, konstruieren wir neue Lyapunov Funktionale für Entropiemethoden und kombinieren Resultate nicht-symmetrischer Spektraltheorie.

Die Arbeit ist in drei Kapitel gegliedert:

Im ersten Kapitel beweisen wir scharfes asymptotisches Langzeitverhalten mittels einer Familie von Entropien für *defekte Fokker–Planck Gleichungen* auf $\mathbb{R}^d$ mit hypokoerzivem Verhalten und zeigen, dass das Abklingverhalten dem einer defekten GDG entspricht. Die Neuheit unserer Methodik liegt dabei in der Kombination von Spektraltheorie und nicht-symmetrischer Hyperkontraktivität, eine explizite Glättungseigenschaft des Fokker–Planck Propagators, die wir für degenerierte Diffusion beweisen.

Im zweiten Kapitel werden explizite *Lypanunov Funktionale für lineare GDG* konstruiert, die scharfe Abklingraten, inklusive den defekten Fällen, liefern. Zur Anwendung dieser Methode betrachten wir drei Evolutionsgleichungen: Die lineare Konvektions-Diffusionsgleichung, die Goldstein–Taylor Gleichung und die Fokker–Planck Gleichung.

Die Erweiterung der Gleichungen mit einem zusätzlichen Parameter, der Unsicherheiten in der praktischen Bestimmung von Gleichungskoeffizienten beschreibt, und einer linearen Sensitivitätsanalyse dieses Parameters führt zu defekten GDG. Die Anwendung unserer Lyapunov Funktional Methode liefert scharfe Abschätzungen des Langzeitverhaltens von charakteristischer defekter Form. Dabei ist es essenziell, dass durch das Auftreten des Unsicherheitsparameters die Abschätzungen gleichmäßig im *nicht-defekten Limes* sind.

Im dritten Kapitel wird ein Entropiefunktional konstruiert, um das Langzeitverhalten der *Goldstein–Taylor Gleichung am Torus mit ortsabhängigem Relaxationskoeffizienten* zu analysieren. Die Verwendung dieses Funktionals führt zu scharfen Abklingraten für konstante Relaxation und gibt explizite Raten im Falle einer örtlich variierenden Relaxation. Um das Erweiterungspotential unserer Methode für verwandte Modelle zu demonstrieren, beweisen wir explizite Abklingraten für ein auf drei Geschwindigkeiten erweitertes Goldstein–Taylor Model.

Abstract

This thesis is devoted to the analysis of the long-time behaviour of solutions to linear kinetic equations with defects. The two main models of interest are the degenerate Fokker–Planck equation and the Goldstein–Taylor system (a two velocity transport-relaxation model of BGK-type), which both exhibit hypocoercive dynamics. The thesis focuses on the defective cases that occur in these models, which, much like finite dimensional defective ODEs, imply a polynomial times exponential decay of solutions. To obtain explicit estimates on the decay behaviour of solutions, we construct tools for entropy methods and utilise spectral theory in a non-symmetric setting.

The thesis is divided into three parts:

In the first part, we establish sharp long-time asymptotic behaviour for a family of entropies to *defective Fokker–Planck equations* on $\mathbb{R}^d$ that exhibit hypocoercive dynamics, and we show that their decay rate is an exponential multiplied by a polynomial in time. The novelty of our study lies in the combination of spectral theory and non-symmetric hypercontractivity, a long-time smoothing property of the Fokker–Planck propagator that we extend to include degenerate diffusion.

In the second part, we review the *Lyapunov functional method for linear ODEs* and give an explicit construction of such functionals that yield sharp decay estimates, including an extension to defective ODE systems. As an application, we consider three evolution equations, namely the linear convection-diffusion equation, the Goldstein–Taylor equation and the Fokker–Planck equation.

Adding an uncertain parameter to the equations and analysing their linear sensitivity with respect to this parameter leads to defective ODE systems. By applying the Lyapunov functional framework, we prove sharp long-time behaviour of the typical defective form. The appearance of the uncertain parameter in the three applications makes it important to have decay estimates that are uniform in the *non-defective limit*.

Finally, in the last part, we construct an entropy functional to analyse the long-time behaviour of the *Goldstein–Taylor equation* on the one-dimensional torus with space-dependent relaxation. Utilising this functional yields sharp decay rates to equilibrium for constant relaxation, and explicit decay rates, when the relaxation varies in space. To demonstrate the potential of extending our entropy method to related models, we prove exponential decay with an explicit rate for a three-velocity Goldstein–Taylor model.

Acknowledgement

First and foremost I want to express my gratitude to my advisor Anton Arnold for the very interesting topics, exemplary support and guidance in writing this thesis. I am very thankful for the countless motivating mathematical discussions we had at the university and international conferences, that make the topics come to life and difficult challenges more approachable.

I am also very thankful to Prof. Shi Jin for the collaboration on three continents, interesting new insights, support and making me feel at home in his work group during my stay in Shanghai.

I am very grateful to Amit for our fruitful collaborations and his guidance throughout my doctoral studies. From the beginning, which made my introduction into kinetic topics so much smoother, until the final stressful weeks of finishing the thesis and helping with proof reading.

Great thanks to Beatrice for the collaboration, intense scientific discussions and friendship.

I am thankful for the engaging environment of our work group at TU Wien, which I painfully miss in times of the Corona virus. Thanks to Giacomo, who was always willing to discuss small questions, which, more often than not, turned into bigger discussions. Thanks also to Franz, Boris, Dimitri and Kirian who always had an open ear if needed.

I thank Gaurav and Franz for final proof reading and Oliver for friendship, talk feedback and many great scientific discussions.

A very special thanks goes to my parents, Gudrun and Heinz, who gave me the strong and stable support throughout the years of my mathematical education that made everything possible.

Finally, I wholeheartedly thank Liuying for her selfless support and strength.

Contents

Introduction

To sum up, the aim of mathematical physics is not only to facilitate for the physicist the numerical calculation of certain constants or the integration of certain differential equations. It is besides, it is above all, to reveal to him the hidden harmony of things in making him see them in a new way.

— Henri Poincaré

Mathematical physics takes on the challenge of expressing real-world phenomena through simple and exact mathematical language. In the field of kinetic theory, one guiding principle is the second law of thermodynamics. It states that, in a closed system, heat, i.e. microscopic kinetic energy, always flows from hotter to colder regions as time passes. The mechanics of this process is happening on a microscopic level, where kinetic energy is transferred from one particle to the other upon collision, as in a game of billiard with billions of balls. Due to the immense number of particle collisions, the path of an individual particle exceeds the capacities of direct computation. However, if one takes a step back from the particle point of view and observes the macroscopic properties of the system, an ordered structure reveals itself. The total of all local differences of kinetic energy in the system, called *entropy*, will decrease[1] in each time step. This happens until the whole system reaches a uniform temperature, the thermodynamic equilibrium.

One method to analyse the behaviour of solutions to partial differential equations (PDEs), which arise from a statistical consideration of particle models, is to follow the second law of thermodynamics and use the monotonicity of an entropy that is associated to the equation. The objective of the so-called *entropy method* is to construct a Lyapunov functional that decreases along the evolution of solutions to the PDE. Here, we are specifically devoted to tools that capture the long-time behaviour of the system as precisely and explicitly as possible.

The two main models of interest in the present thesis are the *linear degenerate Fokker–Planck equation* and the *Goldstein–Taylor equation* (a two velocity BGK model[2]). Our

[1] For physicists it is customary to define entropy with the opposite sign. Hence, in the physical context entropy is said to *increase*.

[2] Named after the physicists Bhatnagar, Gross and Krook, '54.

focus lies on understanding the possible appearance of *defects* in the equations. This terminology is analogous to defective eigenvalues, which may appear in the modal decomposition of these equations. The defectiveness in these cases manifests itself in the long time behaviour of the solutions to the equations: The sharp exponential convergence to equilibrium is not purely exponential but rather has the form of an exponential term slowed down by a polynomial factor. We will proceed with providing an overview of the models and our specific settings.

Defective Fokker–Planck Equations

In its simplest form, the linear *Fokker–Planck equation* (FPE) for $x \in \mathbb{R}$ is given as

$$\partial_t f(x,t) = \operatorname{div}(D\nabla f(x,t) + Cxf(x,t)), \quad t \geq 0, \tag{1}$$

where $D > 0$ is the diffusion coefficient and $C > 0$ is the drift coefficient that corresponds to a quadratic confinement potential. The solution $f(x,t)$ describes the probability density function of a statistically average particle under the influence of two forces. A deterministic force that pushes the particle in a certain direction (here, to the origin), corresponding to the drift term, and a fluctuating force due to particle collisions corresponding to the diffusion term.

The research on this fundamental equation has a long history, starting with the statistical analysis of particle fluctuations as Brownian motion. A summarising survey of the equation can be found in [6]. Higher dimensional versions of (1), as well as nonlinear extensions of it, and its long time behaviour have been extensively investigated in the last few decades. One elegant way to estimate the decay to equilibrium for the FPE is the so-called Bakry-Émery method (see [3]). In the setting of the FPE for $x \in \mathbb{R}^d$, however, this methodology works only when diffusion is present in all directions.

Here, we are interested in sharp decay estimates for FPEs that exhibit defective behaviour. We focus on two generalisations of (1) where it arises:

- An extension of (1) to $x \in \mathbb{R}^d$, where the drift coefficient becomes a constant-in-space drift matrix $\boldsymbol{C} \in \mathbb{R}^{d\times d}$ with spectral gap[3] $\mu > 0$. We investigate the case, where $\boldsymbol{C}$ has defective eigenvalues in its spectral gap. Said differently, $\boldsymbol{C}$ is not diagonalisable on the appropriate eigenspace, and has a non-trivial Jordan normal form. As an additional difficulty, we only assume the diffusion matrix to be positive semi-definite, which hence allows degenerate diffusions as in the example of the linear kinetic Fokker–Planck equation.

- We are further interested in including uncertainty to (1) by imposing a drift coefficient $C(z) > 0$ that depends on an uncertain parameter $z \in \mathbb{R}$. We raise the

[3] The smallest real part of all eigenvalues of $\boldsymbol{C}$.

question: How sensitive is the long-time behaviour of solutions to modelling uncertainty in the drift coefficient? Analysing the solution dependence on z leads to sensitivity equations that again exhibit a defective structure.

Defectiveness

The main challenge in the above mentioned defective cases lies in the deviation from the purely exponential decay behaviour of solutions. We shall look at an explicit example of this phenomenon: For a given symmetric positive definite diffusion matrix $\boldsymbol{D} \in \mathbb{R}^{d\times d}$ and a drift matrix $\boldsymbol{C} \in \mathbb{R}^{d\times d}$ with spectral gap $\mu > 0$, consider the following Fokker–Planck equation:

$$\begin{aligned} \partial_t f(x,t) &= \operatorname{div}(\boldsymbol{D}\nabla f(x,t) + \boldsymbol{C}xf(x,t)) =: Lf(x,t), \quad x\in\mathbb{R}^d,\ t\geq 0, \\ f(x,0) &= f_\infty(x) + f_1(x), \end{aligned} \tag{2}$$

where f_∞ is the equilibrium of the equation, and $f_1 \in V_1$, where V_1 is a finite dimensional L^2-subspace, which includes the eigenfunctions of the FP operator L corresponding to its spectral gap (cf. Fig. 1).

One can show that the evolution of the semigroup e^{Lt} on V_1 is equivalent to the evolution of the semigroup $e^{-\boldsymbol{C}^T t}$ (with respect to the coefficients of each element of V_1 in a standard basis). Thus, since $f(x,t) - f_\infty(x) = e^{Lt} f_1(x)$, we see that in order to understand the long time behaviour of the solution, we only need to consider the ODE $\dot{x} = -\boldsymbol{C}^T x$. From the above, we can conclude that if $\boldsymbol{C}$ has an eigenvalue with real-part μ that is *defective of order*[4] $n\in\mathbb{N}$, then

$$\|f(x,t) - f_\infty(x)\|_{L^2} \leq \mathscr{C}(1+t^n)e^{-\mu t}. \tag{3}$$

The exponential decay with rate μ is slowed down by the polynomial in time of order n due to the non-trivial Jordan normal form of $\boldsymbol{C}$.

One difficulty that arises in the defective cases is that entropy methods commonly rely on finding a time independent entropy functional $E[\cdot]$ and $\alpha > 0$, such that any solution f of the equation satisfies

$$\frac{d}{dt}E[f(\cdot,t)] \leq -\alpha E[f(\cdot,t)], \quad t\geq 0.$$

Gronwall's Lemma then directly implies *purely* exponential decay in entropy with rate $\alpha > 0$. To recover decay of the form presented in (3), which is natural in the defective setting, different, and more complicated techniques are required — such as allowing the entropy functional to be explicitly dependent on time. For the d-dimensional defective FPE, we shall take an alternative approach to the entropy methods: Combining spectral

[4] An eigenvalue is *defective of order n* if the difference between its algebraic multiplicity and its geometric multiplicity is n. This corresponds to a Jordan block of size $n+1$.

properties of the (in general non-symmetric) FP operator L in L^2 together with a non-symmetric hypercontractivity result, which asserts that solutions eventually belong to the appropriate L^2 space, to achieve our sharp decay estimates.

For arbitrary L^2 initial data, our strategy is splitting the solution into two parts:

$$f(x,t) = f_1(x,t) + f_2(x,t),$$

with a finite dimensional part, f_1, corresponding to the discussion above with decay (3), and an orthogonal remainder f_2 that lies in an infinite dimensional subspace, which converges to equilibrium significantly faster. See Fig. 1 for the correspondence of eigenvalues of L and the subspaces partition of L^2.

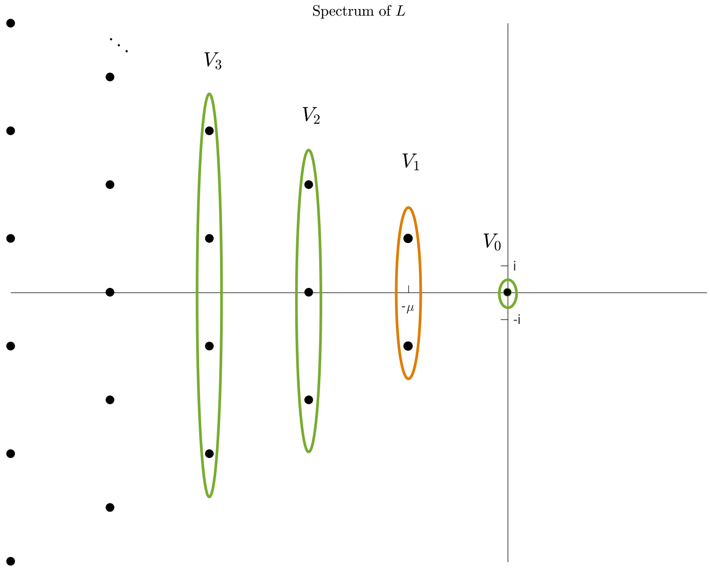


Figure 1: The black dots represent the spectrum of L. The grouping depicts the correspondence of eigenvalues with the subspace partition $L^2 = \bigoplus_{i=0}^{\infty} V_i$. The solution part f_1 corresponds to V_1 and f_2 corresponds to $V_1^{\perp}$.

To extend the decay estimates to include even more general initial data of only L^p-integrability for $1 < p < 2$, we prove, there is an explicit waiting time $t_p > 0$ after which the solution is L^2-integrable. In the standard case of diffusion in all directions, this property is called hypercontractivity. It states the equivalence between the L-associated Log-Sobolev inequality constants and the explicit waiting time until the L^p initial data, $1 < p < 2$, reaches L^2 integrability. In our most general setting, there is no naturally associated Log-Sobolev inequality if the diffusion is degenerate yet a smoothing property of solutions is still present. Thus, if we start with initial datum in an appropriate L^p space, with $1 < p < 2$, we only need to wait t_p time, before being able to use our decay estimate for L^2 datum, yielding sharp decay rates.

A further technical challenge arises, if we include an uncertain parameter $z \in \mathbb{R}$ in the FPE (as well as the GT equation, which we will discuss shortly). Proving decay estimates that are uniform in this parameter require particular care when there is a transition from defective to non-defective regimes, which we shall call *non-defective limits.* The reason is that the underlying geometric structure changes drastically in such a transition (as well as between different orders of defectiveness). To explain what we mean, we consider the ODE system

$$\dot{y} = -\boldsymbol{A}_\varepsilon y \quad \text{with} \quad \boldsymbol{A}_\varepsilon := \begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix}, \tag{4}$$

which is defective of order 1, if and only if $\varepsilon \neq 0$. For $\varepsilon \neq 0$ its corresponding Jordan transformation matrix reads

$$\boldsymbol{V}_\varepsilon := \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\varepsilon} \end{pmatrix}.$$

For fixed $\varepsilon \neq 0$, a standard calculation shows that

$$|y(t)|_2 \leq |\boldsymbol{V}_\varepsilon|_2 |\boldsymbol{V}_\varepsilon^{-1}|_2 |e^{-\boldsymbol{J}t}|_2 |y(0)|_2 \leq |\boldsymbol{V}_\varepsilon|_2 |\boldsymbol{V}_\varepsilon^{-1}|_2 c(1+t) e^{-t} |y(0)|_2, \tag{5}$$

where $\boldsymbol{J}$ is the Jordan normal form of $\boldsymbol{A}_\varepsilon$. For $\varepsilon \to 0$, the factor $|\boldsymbol{V}_\varepsilon|_2 |\boldsymbol{V}_\varepsilon^{-1}|_2$ in (5) becomes unbounded of order ε^{-1} (even though the true decay of the solution improves to $e^{-t}|y(0)|_2$ in the limit). This is due to the discontinuity of the Jordan transformation at the transition from defectiveness to non-defectiveness.

We circumvent this problem by providing *a time dependent Lyapunov functional framework for finite dimensional ODEs.* They are of form

$$|y|^2_{\boldsymbol{P}(t)} = y^H \boldsymbol{P}(t) y, \quad y \in \mathbb{C}^d, \tag{6}$$

where $\boldsymbol{P}(t) \in \mathbb{C}^{d\times d}$ is an explicit positive definite matrix for all $t \geq 0$. For the example (4), this framework yields an explicit matrix $\boldsymbol{P}_\varepsilon(t)$ such that

$$|y(t)|_{\boldsymbol{P}_\varepsilon(t)} = e^{-t} |y(0)|_{\boldsymbol{P}_\varepsilon(0)}.$$

As the norm itself is time-dependent, we relate this estimate back to the Euclidean norm, which leads to

$$|y(t)|_2 \leq \mathscr{C}_{\boldsymbol{P}_\varepsilon}(1+t)|y(0)|_2,$$

with an explicit constant $\mathscr{C}_{\boldsymbol{P}_\varepsilon} > 0$. The advantage of the above is that the constant, which appears in this estimate, can be chosen to be bounded in the non-defective limit $\varepsilon \to 0$.

An analogous problem to (4) appears for FPE of form (1) with uncertain parameter $z \in \mathbb{R}$ in the drift coefficient $C(z)$ (and the GT equation discussed below with uncertainty in the relaxation coefficient). Projected onto the first eigenfunction of the FP operator $L(z)$ (in analogy to the subspace V_1 in the non-symmetric case above), the first order sensitivity equations w.r.t. z reduce to an ODE where non-defective limits appear. Our framework of time-dependent norms $|\cdot|_{\boldsymbol{P}(z,t)}$ then provide sharp decay estimates for the sensitivity equations which are *uniform in the uncertain paramter.*

Goldstein–Taylor Model with Space-Dependent Relaxation

The *Goldstein–Taylor* (GT) system on the one-dimensional torus $x \in \mathbb{T}^1$ is given by

$$\begin{aligned} \partial_t f_+(x,t) + \partial_x f_+(x,t) &= \frac{\sigma}{2}(f_-(x,t) - f_+(x,t)), \\ \partial_t f_-(x,t) - \partial_x f_-(x,t) &= -\frac{\sigma}{2}(f_-(x,t) - f_+(x,t)), \end{aligned} \tag{7}$$

for time $t \geq 0$, where $f_\pm(x,t)$ represents the distribution of particles in the system that travel with velocity ± 1, respectively. The relaxation term that appears on the right-hand side corresponds to “collisions” of particles in the system with relaxation rate $\sigma > 0$. The GT model is a two velocity BGK model and as such encapsulates the core dynamics of these type of models, which are of hypocoercive nature, a topic which will be further discussed below.

For constant $\sigma > 0$, the GT model can be solved using straightforward methods (spatial Fourier expansion) and can serve as a first toy model to construct tools for more complex settings. Our aim is to understand the long time behaviour of the solution to (7), when the relaxation coefficient includes uncertainty or varies in space. Here, we focus on constructing entropy functionals that yield explicit decay estimates which can be generalised to closely related models, e.g. multi-velocity GT models.

- One goal is to perform uncertainty quantification for the GT model with a first order sensitivity analysis. As in the FP case above, the resulting sensitivity equations again exhibit a defective structure and can be treated via modified norms as Lyapunov functionals.

- Furthermore, we develop an entropy functional for GT models with space-dependent relaxation $\sigma(x) > 0$ that yields explicit decay estimates. The main feature of the entropy and method we find is the possible extension to models of similar nature. In comparison to sharp decay rates for the equation when σ is not constant, obtained in [4], our rate is not optimal, yet the methodology used by authors in [4] applies only to (7). We provide explicit decay estimates for an extension of the GTE to a three velocity model to emphasise our methods potential for extensions to similar settings.

Hypocoercivity

One main difficulty the above presented models have in common is the presence of *hypocoercive dynamics*, a topic which received growing attention since Villani’s monograph in 2009, see [7]. In contrast to coercive evolution equations, hypocoercive equations do not exhibit a global force driving solutions to equilibrium. It is rather an interplay of two effects, one conservative and one degenerate dissipative, that results in exponential decay. The abstract setting is the following.

Let A be a *coercive operator* on a Hilbert space $\mathscr{H}$ with scalar product $\langle \cdot, \cdot \rangle_{\mathscr{H}} \to \mathbb{C}$, i.e.

$$\operatorname{Re} \langle Ag, g \rangle_{\mathscr{H}} \geq \mu \|g\|_{\mathscr{H}}^2, \quad g \in \mathscr{H}, \tag{8}$$

with $\mu > 0$. Then, solutions to the initial value problem

$$\partial_t f = -Af, \quad f(0) = f_0 \in \mathscr{H},$$

satisfy

$$\partial_t \|f\|_{\mathscr{H}}^2 = -2 \operatorname{Re} \langle Af, f \rangle_{\mathscr{H}} \leq -2\mu \|f\|_{\mathscr{H}}^2.$$

As an immediate consequence, we have that

$$\|f(t)\|_{\mathscr{H}} \leq e^{-\mu t} \|f_0\|_{\mathscr{H}}, \quad t \geq 0.$$

For operators L that are coercive only on a subspace $\tilde{\mathscr{H}} \subset \mathscr{H}$, one cannot, in general, deduce exponential decay of solutions. However, decay of form

$$\|f(t)\|_{\mathscr{H}} \leq \mathscr{C} e^{-\mu t} \|f_0\|_{\mathscr{H}},$$

with $\mathscr{C} \geq 1$ is still possible if L has a *hypocoercive* form. That is

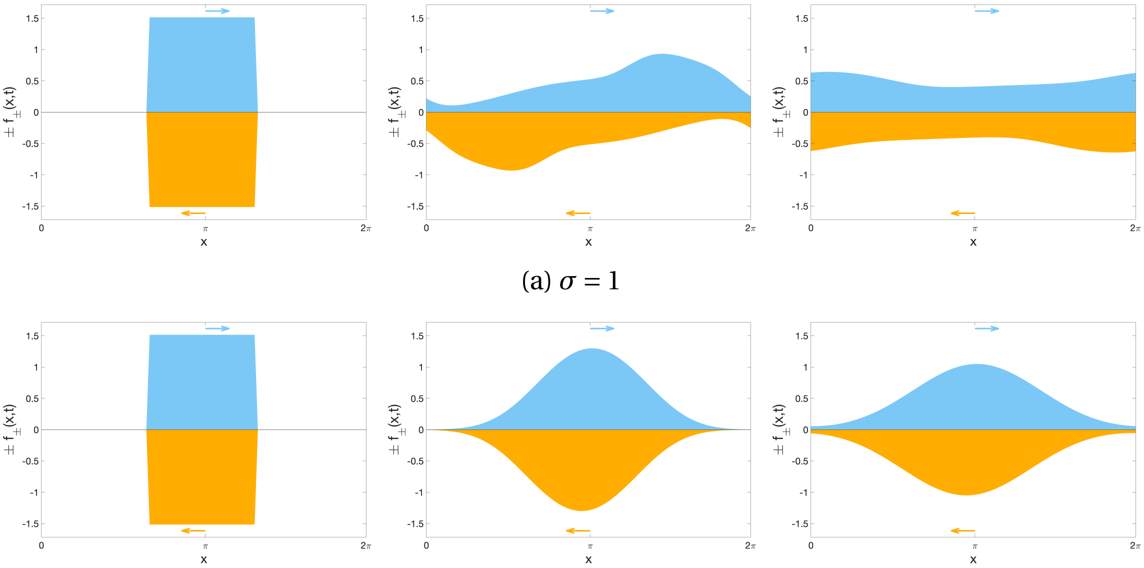
$$L = A + T, \quad \text{with} \quad T^* = -T, \tag{9}$$

where A is symmetric and coercive on a subspace $\tilde{\mathscr{H}}$ and T, typically a transport operator for kinetic equations, "mixes" the coercive subspace with its orthogonal. In this abstract forumlation, the necessary mixing properties are expressed in commutator relations involving A and T.

The GT model (7) provides a good illustration of hypocoercive dynamics: The transport terms, corresponding to the operator T in (9), represent a "horizontal force", shifting the particle mass to the left and right on the torus. It competes with relaxation, corresponding to the operator A in (9), that acts as a "vertical force" on the mass densities, reducing the local mass difference between the two particle types. In combination, every initial mass distribution gets "flattened out" over time to approach a uniform distribution along the torus, which is the unique global equilibrium. See Fig. 2a for plots of the solution behaviour for $\sigma = 1$.

The strength of the relaxation term, measured by σ, directly influences the long-time behaviour of the GT model: A constant relaxation rate of $\sigma \in (0, 2)$ translates into an exponential convergence rate $\frac{\sigma}{2}$. If $\sigma = 2$, the system is defective, resulting in a convergence behaviour of order $(1 + t)e^{-t}$. For relaxation rates $\sigma > 2$, a slowing down of the exponential rate to $\frac{\sigma}{2} - \sqrt{\frac{\sigma^2}{4} - \frac{1}{4}}$ occurs.

The reason behind the slowing in the last case is that locally the mass is balanced very quickly between the two species, giving the transport term little time to "spread it"



(a) $\sigma = 1$

(b) $\sigma = 10$

Figure 2: The evolution according to the Goldstein–Taylor model with relaxation σ. Here, the initial mass is distributed equally between the two species around $x = \pi$. Depicted are the points in time $t = 0,2,4$ from left to right. The images are created from an implicit Euler scheme simulation.

across the torus. For example, when one considers concentrated and balanced initial particle masses, the high relaxation rate results in a large amount of mass frequently changing direction. This prevents an effective shift of mass away from the initial region, see Fig. 2b.

While there are general strategies to incorporate hypocoercivity into an entropy method approach, see [5], explicit and precise decay estimates for many models still need to be fine-tuned. In dealing with space-dependent relaxation for the GT model, we first develop an entropy functional that captures the sharp decay for all cases of constant relaxation. This functional is pseudodifferential in x. Then, we use this functional to obtain explicit results for space-dependent relaxation in a somewhat "perturbative" approach.

A similar hypocoercive interplay of "forces" can occur in the FP setting when the diffusion is degenerate (and thus L is no longer coercive). To achieve such interplay, the drift term of the equation must mix the non-diffusive directions with the diffusive ones, causing the operator to always be non-symmetric. Using the above described methodology of solution splitting in this non-symmetric setting is the main technical challenge.

For the uncertainty quantification of both the FP and GT equations, hypocoercivity arises on a modal level as *hypocoercive ODEs*. Let us consider linear ODEs $\dot{y} = -\boldsymbol{A}y$ that are non-defective for simplicity[5]. The analogue to hypocoercive operators in this setting are matrices $\boldsymbol{A}$ with spectral gap $\mu_{\boldsymbol{A}} > 0$, but whose symmetric part is only positive semi-definite. Using the Euclidean norm as a Lyapunov functional, i.e.

$$\frac{d}{dt}|y|_2^2 = -y^H \underbrace{(A^H + A)}_{2\boldsymbol{A}_{\mathsf{symm}}} y \le 0$$

does not provide any decay rate due to the non-trivial kernel of $\boldsymbol{A}_{\mathsf{symm}}$.

By considering an appropriate positive definite matrix $\boldsymbol{P}$, one can geometrically transform the variable space, and consequently the ODE. The transformed ODE has a new system matrix, $\tilde{\boldsymbol{A}}$, with a positive definite symmetric part that has $\mu_{\boldsymbol{A}}$ as its spectral gap:

$$\tilde{\boldsymbol{A}} := \sqrt{\boldsymbol{P}}\boldsymbol{A}\sqrt{\boldsymbol{P}}^{-1}, \quad \tilde{\boldsymbol{A}}_{\mathsf{symm}} \ge \mu_{\boldsymbol{A}}\boldsymbol{I}.$$

For a geometric interpretation of the transformation induced by $\boldsymbol{P}$, see Fig. 3. Subsequently, an entropy method in $\boldsymbol{P}$-norm yields sharp exponential decay. Indeed, denoting $\tilde{y} := \sqrt{\boldsymbol{P}}y$, we have that

$$\frac{d}{dt}|y|_{\boldsymbol{P}}^2 = \frac{d}{dt}|\tilde{y}|_2^2 = -2\tilde{y}^H\tilde{\boldsymbol{A}}_{\mathsf{symm}}\tilde{y} \le -2\mu_{\boldsymbol{A}}|\tilde{y}|_2^2 = -2\mu_{\boldsymbol{A}}|y|_{\boldsymbol{P}}^2.$$

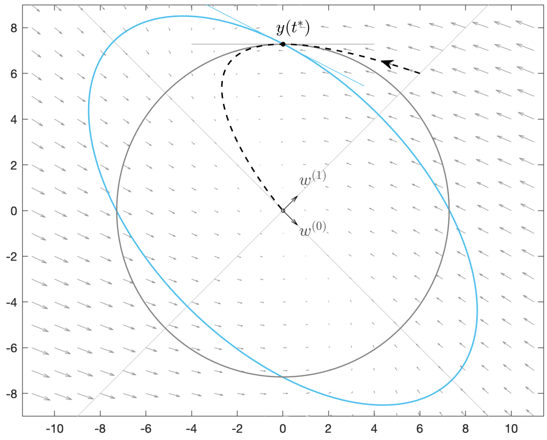


Figure 3: The dashed line shows the solution trajectory $y(t)$. At the marked point $y(t^*)$, the solution is tangential to the Euclidean level curve. This implies non-strict decay in the Euclidean norm. The ellipse represents a level curve of the $\boldsymbol{P}$-norm. It modifies the geometry such that the solution is never tangential to the level curves of $|\cdot|_{\boldsymbol{P}}$.

[5] The more involved defective cases that require norms depending on time are discussed in Chapter 2.

Structure & Authorship

The thesis is divided into three chapters.

Chapter 1 is devoted to linear Fokker–Planck equations on $\mathbb{R}^d$ of form (2) with degenerate diffusion and defective drift. We collect the necessary spectral information of the in general non-symmetric FP operator to be able to split the solution into two orthogonal parts. In combination we prove sharp long-time behaviour of solutions in L^2 and subsequently extend the decay estimates to a family of the more general L^p-entropies, $1 < p \leq 2$ the associated p-Fisher information functionals.

The content of this chapter is a joint work with Anton Arnold and Amit Einav. The results were published in [1].

In **Chapter 2**, we review the *Lyapunov functional method for linear ODEs* and give an explicit construction of such functionals that yields sharp decay estimates, including an extension to defective ODE systems. As an application, we consider three evolution equations, namely the linear convection-diffusion equation, the Goldstein–Taylor equation and the Fokker–Planck equation with an added uncertain parameter. Analysing its linear sensitivity leads to defective ODE systems. By applying the Lyapunov functional framework, we prove sharp long time behaviour of the typical defective form.

The content of this chapter is a joint work with Anton Arnold and Shi Jin. The results were published in [2].

In **Chapter 3**, we construct a spatial entropy functional to analyse the long time behaviour of the *Goldstein–Taylor equation* on the torus with space-dependent relaxation. Utilising this functional yields sharp decay rates to equilibrium for constant relaxation, and explicit decay rates, when the relaxation varies in space. We further prove explicit decay for a three velocity BGK model.

The content of this chapter is a joint work with Anton Arnold, Amit Einav and Beatrice Signorello.

Bibliography

[1] Arnold, A., Einav, A. and Wöhrer, T.: *On the rates of decay to equilibrium in degenerate and defective Fokker–Planck equations.* J. Differential Equations, vol. 264 (11), 6843–6872, (2018).

[2] Arnold, A., Jin, S. and Wöhrer, T.: *Sharp Decay Estimates in Local Sensitivity Analysis for Evolution Equations with Uncertainties: from ODEs to Linear Kinetic Equations* J. Differential Equations, vol. 268 (3), 1156–1204, (2020).

[3] Bakry, D., Émery, M.: *Diffusions hypercontractives*, Séminaire de probabiltés de Strasbourg **19** (1985), 177–206.

[4] Bernard, É., Salvarani, F.: *Optimal Estimate of the Spectral Gap for the Degenerate Goldstein-Taylor Model.* J Stat Phys 153: 363. https://doi.org/10.1007/s10955-013-0825-6 (2013); Erratum (2020).

[5] Dolbeault, J., Mouhot, C. and Schmeiser, C.: *Hypocoercivity for linear kinetic equations conserving mass*, Trans. Amer. Math. Soc., vol. 367, 3807–3828 (2015).

[6] Risken, H.: *The Fokker–Planck equation. Methods of solution and applications.*, Springer-Verlag (1989).

[7] Villani, C.: *Hypocoercivity*, American Mathematical Soc., (2009).

1 On The Rates of Decay to Equilibrium in Degenerate and Defective Fokker–Planck Equations

1.1 Introduction

1.1.1 Background

The study of Fokker–Planck equations (sometimes also called Kolmogorov forward equations) has a long history - going back to the early 20th century. Originally, Fokker and Planck used their equation to describe Brownian motion in a PDE form, rather than its usual SDE representation.

In its most general form, the Fokker–Planck equation reads as

$$\partial_t f(t,x) = \sum_{i,j=1}^{d} \partial_{x_i x_j} \big(D_{ij}(x) f(t,x)\big) - \sum_{i=1}^{d} \partial_{x_i} \big(A_i(x) f(t,x)\big), \tag{1.1.1}$$

with $t > 0, x \in \mathbb{R}^d$, and where $D_{ij}(x), A_i(x)$ are real valued functions, with the matrix $\boldsymbol{D}(x) = \big(D_{ij}(x)\big)_{i,j=1,\dots,d}$ being positive semidefinite.

The Fokker–Planck equation has many usages in modern mathematics and physics, with connection to statistical physics, plasma physics, stochastic analysis and mathematical finance. For more information about the equation we refer the reader to [19]. Here we will consider a very particular form of (1.1.1) that allows degeneracies and defectiveness to appear.

1.1.2 The Fokker–Planck Equation in our Setting

In this chapter we will focus our attention on Fokker–Planck equations of the form:

$$\partial_t f(t,x) = Lf(t,x) := \operatorname{div}\big(\boldsymbol{D}\nabla f(t,x) + \boldsymbol{C}x f(t,x)\big), \qquad t > 0, x \in \mathbb{R}^d, \tag{1.1.2}$$

with appropriate initial conditions, where the matrix $\boldsymbol{D}$ (the *diffusion* matrix) and $\boldsymbol{C}$ (the *drift* matrix) are assumed to be constant and real valued.

In addition to the above, we will also assume the following:

(A) $\boldsymbol{D}$ is a positive semidefinite matrix with

$$1 \le r := \operatorname{rank}(\boldsymbol{D}) \le d.$$

(B) All the eigenvalues of $\boldsymbol{C}$ have positive real part (this is sometimes called *positive stable*).

(C) There exists no non-trivial $\boldsymbol{C}^T$-invariant subspace of $\operatorname{Ker}(\boldsymbol{D})$ (this is equivalent to *hypoellipticity* of (1.1.2), cf. [12]).

Each of these conditions has a significant impact on the equation:

- Condition (A) allows the possibility that our Fokker–Planck equation is degenerate ($r < d$).
- Condition (B) implies that the drift term confines the system. Hence it is crucial for the existence of a non-trivial steady state to the equation, and
- Condition (C) tells us that, when $\boldsymbol{D}$ is degenerate, $\boldsymbol{C}$ compensates for the lack of diffusion in the appropriate direction and “pushes” the solution back to where diffusion happens.

Equations of the form (1.1.2), with emphasis on the degenerate structure (and hence $d \ge 2$), have been extensively investigated recently (see [2],[17]) and were shown to retain much of the structure of their non-degenerate counterpart. When it comes to the question of long-time behaviour, it has been shown in [2] that under Conditions (A)–(C) there exists a unique equilibrium state f_∞ to (1.1.2) with unit mass (it was actually shown that the kernel of L is one dimensional) and that the convergence rate to it can be explicitly estimated by the use of the so called *(relative) entropy functionals.* Based on [3, 5], and denoting by $\mathbb{R}^+ := \{x > 0 \mid x \in \mathbb{R}\}$ and $\mathbb{R}^+_0 := \mathbb{R}^+ \cup \{0\}$, we introduce these entropy functionals:

Definition 1.1.1. We say that a function ψ is a *generating function for an admissible relative entropy* if $\psi \not\equiv 0$, $\psi \in C\left(\mathbb{R}^+_0\right) \cap C^4\left(\mathbb{R}^+\right)$, $\psi(1) = \psi'(1) = 0$, $\psi'' > 0$ on $\mathbb{R}^+$ and

$$\left(\psi'''\right)^2 \le \frac{1}{2}\psi''\psi''''. \tag{1.1.3}$$

For such a ψ, we define the *admissible relative entropy* $e_\psi\left(\cdot|f_\infty\right)$ to the Fokker–Planck equation (1.1.2) with unit mass equilibrium state f_∞, as the functional

$$e_\psi\left(f|f_\infty\right) := \int_{\mathbb{R}^d} \psi\left(\frac{f(x)}{f_\infty(x)}\right) f_\infty(x)\,dx, \tag{1.1.4}$$

for any non-negative f with unit mass.

Remark 1.1.2. *It is worth to note a few things about Definition 1.1.1:*

- *As ψ is only defined on $\mathbb{R}_0^+$ the admissible relative entropy can only be used for non-negative functions f. This, however, is not a problem for equation* (1.1.2) *as it propagates non-negativity.*
- *Assumption* (1.1.3) *is equivalent to the concavity of $\frac{1}{\psi''}$ on $\mathbb{R}^+$.*
- *Important examples of generating functions include $\psi_1(y) := y\log y - y + 1$ (the Boltzmann entropy) and $\psi_2(y) := \frac{1}{2}(y-1)^2$.*
  Note that for $f \in L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$

  $$e_2(f|f_\infty) = \frac{1}{2}\|f - f_\infty\|^2_{L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)}.$$

  This means that up to some multiplicative constant, e_2 is the square of the (weighted) L^2 norm.

A detailed study of the rate of convergence to equilibrium of the relative entropies for (1.1.2) when $r < d$ was completed recently in [2]. Denoting by $L^1_+\left(\mathbb{R}^d\right)$ the space of non-negative L^1 functions on $\mathbb{R}^d$, the authors have shown the following:

Theorem 1.1.3. *Consider the Fokker–Planck equation* (1.1.2) *with diffusion and drift matrices $\boldsymbol{D}$ and $\boldsymbol{C}$ which satisfy Conditions (A)–(C). Let*

$$\mu := \min\left\{\mathrm{Re}\,(\lambda) \mid \lambda \text{ is an eigenvalue of } \boldsymbol{C}\right\}. \tag{1.1.5}$$

Then, for any admissible relative entropy e_ψ and a solution $f(t)$ to (1.1.2) *with initial datum $f_0 \in L^1_+\left(\mathbb{R}^d\right)$, of unit mass and such that $e_\psi(f_0|f_\infty) < \infty$ we have that:*

(i) If all the eigenvalues from the set

$$\{\lambda \mid \lambda \text{ is an eigenvalue of } \boldsymbol{C} \text{ and } \mathrm{Re}(\lambda) = \mu\} \tag{1.1.6}$$

are non-defective [1], then there exists a fixed geometric constant $c \geq 1$, that doesn't depend on f, such that

$$e_\psi(f(t)|f_\infty) \leq c e_\psi(f_0|f_\infty) e^{-2\mu t}, \quad t \geq 0.$$

(ii) If one of the eigenvalues from the set (1.1.6) *is defective, then for any $\varepsilon > 0$ there exists a fixed geometric constant c_ε, that doesn't depend on f, such that*

$$e_\psi(f(t)|f_\infty) \leq c_\varepsilon e_\psi(f_0|f_\infty) e^{-2(\mu-\varepsilon)t}, \quad t \geq 0. \tag{1.1.7}$$

[1] An eigenvalue is *defective* if its geometric multiplicity is strictly less than its algebraic multiplicity. We will call the difference between these numbers the *defect* of the eigenvalue.

The loss of the exponential rate $e^{-2\mu t}$ in part (ii) of the above theorem is to be expected, however it seems that replacing it by $e^{-2(\mu-\varepsilon)t}$ is too crude. Indeed, if one considers the much related, finite dimensional, ODE equivalent

$$\dot{x} = -\boldsymbol{B}x$$

where the matrix $\boldsymbol{B} \in \mathbb{R}^{d\times d}$ is positive stable and has, for example, a defect of order 1 in an eigenvalue with real part equal to $\mu > 0$ (defined as in (1.1.5)). Then one notices that

$$\|x(t)\|^2 \leq c\|x_0\|^2 \left(1+t^2\right) e^{-2\mu t}, \quad t \geq 0,$$

i.e. the rate of decay is worsened by a multiplication of a polynomial of the order twice the defect of the "minimal eigenvalue".

The goal of this chapter is to show that the above is also the case for our Fokker–Planck equation.

We will mostly focus our attention on the family of relative entropies $e_p\left(\cdot|f_\infty\right)$, with $1 < p \leq 2$, which are generated by

$$\psi_p(y) := \frac{y^p - p(y-1) - 1}{p(p-1)}.$$

Notice that ψ_1 can be understood as the limit of the above family as p goes to 1.

An important observation about the above family, that we will use later, is the fact that *the generating function for $p = 2$, associated to the entropy e_2, is actually defined on $\mathbb{R}$ and not only $\mathbb{R}^+$.* This is not surprising as we saw the connection between e_2 and the L^2 norm. This means that we are allowed to use e_2 even when we deal with functions without a definite sign.

Our main theorem for this chapter is the following:

Theorem 1.1.4. *Consider the Fokker–Planck equation* (1.1.2) *with diffusion and drift matrices $\boldsymbol{D}$ and $\boldsymbol{C}$ which satisfy Conditions (A)–(C). Let μ be defined as in* (1.1.5) *and assume that one, or more, of the eigenvalues of $\boldsymbol{C}$ with real part μ are defective. Denote by $n > 0$ the maximal defect of these eigenvalues. Then, for any $1 < p \leq 2$, the solution $f(t)$ to* (1.1.2) *with unit mass initial datum $f_0 \in L^1_+\left(\mathbb{R}^d\right)$ and finite p-entropy, i.e. $e_p\left(f_0|f_\infty\right) < \infty$, satisfies*

$$e_p\left(f(t)|f_\infty\right) \leq \begin{cases} c_2 e_2\left(f_0|f_\infty\right)\left(1+t^{2n}\right)e^{-2\mu t}, & p = 2,\\ c_p\left(p(p-1)e_p(f_0|f_\infty)+1\right)^{\frac{2}{p}}\left(1+t^{2n}\right)e^{-2\mu t}, & 1 < p < 2, \end{cases}$$

for $t \geq 0$, where $c_p > 0$ is a fixed geometric constant, that doesn't depend on f_0, and f_∞ is the unique equilibrium with unit mass.

The main idea, and novelty, of this work is in combining elements from Spectral Theory and the study of our p-entropies. We will give a detailed study of the geometry of the operator L in the $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$ space and deduce, from its spectral properties, the result for e_2. Since the other entropies, e_p for $1 < p < 2$, lack the underlying geometry of the L^2 space that e_2 enjoys, we will require additional tools: We will show a quantitative result of *hypercontractivity for non-symmetric Fokker–Planck operators* that will assure us that after a certain, *explicit* time, any solution to our equation with finite p-entropy will belong to $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$. This, together with the dominance of e_2 over e_p for functions in $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$ will allow us to "push" the spectral geometry of L to solutions with initial datum that only has finite p-entropy.

We have recently become aware that the long-time behaviour of Theorem 1.1.4 has been shown in a preprint by Monmarché, [15]. However, the method he uses to show this result is a generalised entropy method (more on which can be found in §1.5), while we have taken a completely different approach to the matter.

The structure of the chapter is as follows: In §1.2 we will recall known facts about the Fokker–Planck equation (degenerate or not). §1.3 will see the spectral investigation of L and the proof of Theorem 1.1.4 for $p = 2$. In §1.4 we will show our non-symmetric hypercontractivity result and conclude the proof of our Theorem 1.1.4. Lastly, in §1.5 we will recall another important tool in the study of Fokker–Planck equations — the Fisher information — and show that Theorem 1.1.4 can also be formulated for it, due to the hypoelliptic regularisation of the equation.

1.2 The Fokker–Planck Equation

This section is mainly based on recent work of Arnold and Erb (see [2]). We will provide here, mostly without proof, known facts about degenerate (and non-degenerate) Fokker–Planck equations of the form (1.1.2).

Theorem 1.2.1. *Consider the Fokker–Planck equation* (1.1.2)*, with diffusion and drift matrices* $\boldsymbol{D}$ *and* $\boldsymbol{C}$ *that satisfy Conditions (A)–(C), and an initial datum* $f_0 \in L^1_+\left(\mathbb{R}^d\right)$. *Then*

(i) There exists a unique classical solution $f \in C^\infty\left(\mathbb{R}^+ \times \mathbb{R}^d\right)$ *to the equation. Moreover, if* $f_0 \neq 0$ *it is strictly positive for all* $t > 0$.

(ii) For the above solution $\int_{\mathbb{R}^d} f(t,x)dx = \int_{\mathbb{R}^d} f_0(x)dx$.

(iii) If in addition $f_0 \in L^p\left(\mathbb{R}^d\right)$ *for some* $1 < p \leq \infty$, *then* $f \in C\left([0,\infty), L^p\left(\mathbb{R}^d\right)\right)$.

Theorem 1.2.2. *Assume that the diffusion and drift matrices,* $\boldsymbol{D}$ *and* $\boldsymbol{C}$, *satisfy Conditions (A)–(C). Then, there exists a unique stationary state* $f_\infty \in L^1\left(\mathbb{R}^d\right)$ *to* (1.1.2) *satisfying* $\int_{\mathbb{R}^d} f_\infty(x)dx = 1$. *Moreover,* f_∞ *is of the form:*

$$f_\infty(x) = c_{\boldsymbol{K}} e^{-\frac{1}{2}x^T \boldsymbol{K}^{-1} x}, \tag{1.2.1}$$

where the covariance matrix $\boldsymbol{K} \in \mathbb{R}^{d\times d}$ is the unique, symmetric and positive definite solution to the continuous Lyapunov equation

$$2\boldsymbol{D} = \boldsymbol{C}\boldsymbol{K} + \boldsymbol{K}\boldsymbol{C}^T,$$

and where $c_{\boldsymbol{K}} > 0$ is the appropriate normalization constant. In addition, for any $f_0 \in L^1_+\left(\mathbb{R}^d\right)$ with unit mass, the solution to the Fokker–Planck equation (1.1.2) *with initial datum f_0 converges to f_∞ in relative entropy (as referred to in Theorem 1.1.3).*

Remark 1.2.3. *In the case where $f_0 \in L^1_+\left(\mathbb{R}^d\right)$ is not of unit mass, it is immediate to deduce that the solution to the Fokker–Planck equation with initial datum f_0 converges to $\left(\int_{\mathbb{R}^d} f_0(x)dx\right) f_\infty(x)$.*

Corollary 1.2.4. The Fokker–Planck operator L can be rewritten as

$$Lf = \operatorname{div}\left(f_\infty(x)\boldsymbol{C}\boldsymbol{K}\nabla\left(\frac{f(t,x)}{f_\infty(x)}\right)\right) \tag{1.2.2}$$

(cf. Theorem 3.5 in [2]).

A surprising, and useful, property of (1.1.2) is that the diffusion and drift matrices associated to it can always be simplified by using a change of variables. The following can be found in [1]:

Theorem 1.2.5. *Assume that the diffusion and drift matrices satisfy Conditions (A)–(C). Then, there exists a linear change of variable that transforms* (1.1.2) *to itself with new diffusion and drift matrices $\boldsymbol{D}$ and $\boldsymbol{C}$ such that*

$$\boldsymbol{D} = \operatorname{diag}\{d_1, d_2, \ldots, d_r, 0, \ldots, 0\} \tag{1.2.3}$$

with $d_j > 0$, $j = 1, \ldots, r$ and $\boldsymbol{C}_s := \frac{\boldsymbol{C}+\boldsymbol{C}^T}{2} = \boldsymbol{D}$. In these new variables the equilibrium f_∞ is just the standard Gaussian with $\boldsymbol{K} = \boldsymbol{I}$.

The above matrix normalisation has additional impact on the calculation of the adjoint operator:

Corollary 1.2.6. Let $\boldsymbol{C}_s = \boldsymbol{D}$. Then:

(i)

$$\left(L_{\boldsymbol{D},\boldsymbol{C}}\right)^* = L_{\boldsymbol{D},\boldsymbol{C}^T},$$

where L^* denotes the (formal) adjoint of L, considered w.r.t. $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$. The domain of L will be discussed in §1.3.

(ii) The kernels of L and L^* are both spanned by $\exp(-\frac{|x|^2}{2})$. This is not true in general, i.e. for a Fokker–Planck operator L without the matrix normalisation assumption.

Proof. (i) Under the normalising coordinate transformation of Theorem 1.2.5 we see from (1.2.2) that

$$\begin{aligned}\int_{\mathbb{R}^d} f(x) L_{\boldsymbol{D},\boldsymbol{C}} g(x) f_\infty^{-1}(x)dx &= -\int_{\mathbb{R}^d} f_\infty(x)\nabla\left(\frac{f(x)}{f_\infty(x)}\right)^T \boldsymbol{C}\nabla\left(\frac{g(x)}{f_\infty(x)}\right)dx \\ &= \int_{\mathbb{R}^d} \operatorname{div}\left(f_\infty(x)\boldsymbol{C}^T\nabla\left(\frac{f(x)}{f_\infty(x)}\right)\right)g(x)f_\infty^{-1}(x)dx.\end{aligned} \tag{1.2.4}$$

(ii) follows from (1.2.1) and $\boldsymbol{K} = \boldsymbol{I}$. □

From this point onwards we will always assume that Conditions (A)–(C) hold, and that we are in the coordinate system where $\boldsymbol{D}$ is of form (1.2.3) and equals $\boldsymbol{C}_s$.

1.3 The Spectral Study of L

The main goal of this section is to explore the spectral properties of the Fokker–Planck operator L in $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$, and to see how one can use them to understand rates of convergence to equilibrium for e_2. The crucial idea we will implement here is that, since $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$ decomposes into orthogonal eigenspaces of L with eigenvalues that get increasingly farther to the left of the imaginary axis, one can deduce *improved convergence rates on "higher eigenspaces"*.
The first step in achieving the above is to recall the following result from [2], where we use the notation $\mathbb{N}_0 := \mathbb{N}\cup\{0\}$:

Theorem 1.3.1. *Denote by*

$$V_m := \operatorname{span}\left\{\partial_{x_1}^{\alpha_1}\ldots\partial_{x_d}^{\alpha_d} f_\infty(x) \;\middle|\; \alpha_1,\ldots,\alpha_d\in\mathbb{N}_0, \sum_{i=1}^d \alpha_i = m\right\}.$$

Then, $\{V_m\}_{m\in\mathbb{N}_0}$ *are mutually orthogonal in* $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$,

$$L^2\left(\mathbb{R}^d, f_\infty^{-1}\right) = \bigoplus_{m\in\mathbb{N}_0} V_m,$$

and V_m *are invariant under* L *and its adjoint (and thus under the flow of* (1.1.2)*).*
Moreover, the spectrum of L *satisfies*

$$\sigma(L) = \bigcup_{m\in\mathbb{N}_0} \sigma\left(L|_{V_m}\right),$$

$$\sigma\left(L|_{V_m}\right) = \left\{-\sum_{i=1}^d \alpha_i\lambda_i \;\middle|\; \alpha_1,\ldots,\alpha_d\in\mathbb{N}_0, \sum_{i=1}^d \alpha_i = m\right\},$$

where $\left\{\lambda_j\right\}_{j=1,\ldots,d}$ *are the eigenvalues (with possible multiplicity) of the matrix* $\boldsymbol{C}$*. The eigenfunctions of* L *(or eigenfunctions and generalized eigenfunctions in the case* $\boldsymbol{C}$ *is defective) form a basis to* $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$.

Let us note that this orthogonal decomposition is non-trivial since L is in general non-symmetric. The above theorem quantifies our previous statement about "higher eigenspaces": the minimal distance between the eigenvalues of L restricted to the "higher" L-invariant eigenspace V_m and the imaginary axis is $m\mu$. Thus, the decay we expect to find for initial datum from V_m is of order $e^{-2m\mu t}$ (in the quadratic entropy, e.g.). However, as the function we will use in our entropies are not necessarily contained in only finitely many V_m, we might need to pay a price in the rate of convergence.
This intuition is indeed true. Denoting by

$$H_k := \bigoplus_{m\geq k} V_m \tag{1.3.1}$$

for any $k \geq 0$, we have the following:

Theorem 1.3.2. *Let $f_k \in H_k$ for some $k \geq 1$ and let $f(t)$ be the solution to* (1.1.2) *with initial data $f_0 = f_\infty + f_k$. Then for any $0 < \varepsilon < \mu$ there exists a geometric constant $c_{k,\varepsilon} \geq 1$ that depends only on k and ε such that*

$$e_2\left(f(t)|f_\infty\right) \leq c_{k,\varepsilon} e_2(f_0|f_\infty) e^{-2(k\mu-\varepsilon)t}, \quad t \geq 0. \tag{1.3.2}$$

Remark 1.3.3. *The loss of an ε in the decay rate of* (1.3.2) *– compared to the decay rate solely on V_k – can have two causes:*

1. *For drift matrices $\mathbf{C}$ with a defective eigenvalue with real part μ, the larger decay rate $2k\mu$ would not hold in general. This is illustrated in* (1.1.7)*, which provides the best possible purely exponential decay result, as proven in [2].*

2. *For* non-defective matrices *$\mathbf{C}$, the improved decay rate $2k\mu$ actually holds, but our method of proof, that uses the Gearhart-Prüss Theorem, cannot yield this result. The decay estimate* (1.3.2) *will be improved in Theorem 1.3.11: There, the ε-reduction drops out in the non-defective case.*

Remark 1.3.4. *As we insinuated in the introduction to our work, an important observation to make here is that the initial data, f_0,* doesn't have to be non-negative *(and in many cases, is not). While this implies that $f(t)$ might also be non-negative, this poses no problems as e_2 is the squared (weighted) L^2 norm (up to a constant). Theorem 1.3.2* would not work in general *for e_p as the non-negativity of $f(t)$ is crucial there (in other words, f_0 would not be admissible).*

The main tool to prove Theorem 1.3.2 is the Gearhart–Prüss Theorem (see for instance Th. 1.11 Chap. V in [8]). In order to be able to do that, we will need more information about the dissipativity of L and its resolvents with respect to H_k.

Lemma 1.3.5. *Let V_m be as defined in Theorem 1.3.1. Consider the operator L with the domain $D(L) = \operatorname{span}\{V_m, m \in \mathbb{N}_0\}$. Then L is dissipative, and as such closable. Moreover, its closure, $\overline{L}$, generates a contraction semigroup on $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$.*

Proof. Given $f \in D(L)$, and denoting $g := \frac{f}{f_\infty}$, we notice that (1.2.2) with $\boldsymbol{K} = \boldsymbol{I}$ implies that

$$\left(Lf, f\right)_{L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)} = \int_{\mathbb{R}^d} \operatorname{div}\left(f_\infty(x)\boldsymbol{C}\nabla g(x)\right) g(x)dx = -\int_{\mathbb{R}^d} \nabla g(x)^T \boldsymbol{C}\nabla g(x) f_\infty(x)dx$$

$$= -\int_{\mathbb{R}^d} \nabla g(x)^T \boldsymbol{D}\nabla g(x) f_\infty(x)dx \le 0,$$

where we have used the fact that $\boldsymbol{C}_s = \boldsymbol{D}$. Thus, L is dissipative.
To show the second statement we use the Lumer-Phillips Theorem (see for instance Th. 3.15 Chap. II in [8]). Since $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right) = \bigoplus_{m\in\mathbb{N}_0} V_m$ it will be enough to show that for $\lambda > 0$ we have that $V_m \subset \operatorname{Range}(\lambda I - L)$ for any m. As $V_m \subset D(L)$, is finite dimensional, and is invariant under L (Theorem 1.3.1 again) we can consider the linear bounded operator $L|_{V_m} : V_m \to V_m$. Since we have shown that L is dissipative, we can conclude that the eigenvalues of $L|_{V_m}$ have non-positive real parts, implying that $(\lambda I - L)|_{V_m}$ is invertible. This in turn implies that

$$V_m = \operatorname{Range}\left((\lambda I - L)|_{V_m}\right) \subset \operatorname{Range}(\lambda I - L),$$

completing the proof. □

To study the resolvents of L we will need to use some information about its "dual": the Ornstein-Uhlenbeck operator.
For a given symmetric positive semidefinite matrix $\boldsymbol{Q} = (q_{ij})$ and a real, negatively stable matrix $\boldsymbol{B} = (b_{ij})$ on $\mathbb{R}^d$ we consider the Ornstein-Uhlenbeck operator

$$P_{\boldsymbol{Q},\boldsymbol{B}} := \frac{1}{2}\sum_{i,j} q_{ij}\partial^2_{x_i x_j} + \sum_{i,j} b_{ij}x_j\partial_{x_i} = \frac{1}{2}\operatorname{Tr}\left(\boldsymbol{Q}\nabla_x^2\right) + (\boldsymbol{B}x, \nabla_x), \quad x \in \mathbb{R}^d. \tag{1.3.3}$$

Similarly to our conditions on the diffusion and drift matrices, we will only be interested in Ornstein-Uhlenbeck operators that are *hypoelliptic*. In the above setting, this corresponds to the condition

$$\operatorname{rank}\left[\boldsymbol{Q}^{\frac{1}{2}}, \boldsymbol{B}\boldsymbol{Q}^{\frac{1}{2}}, \dots, \boldsymbol{B}^{d-1}\boldsymbol{Q}^{\frac{1}{2}}\right] = d.$$

The hypoellipticity condition guarantees the existence of an invariant measure, $d\mu$, to the process. This measure has a density w.r.t. the Lebesgue measure, which is given by

$$\frac{d\mu}{dx}(x) = c_{\boldsymbol{M}} e^{-\frac{1}{2}x^T\boldsymbol{M}^{-1}x}, \quad \text{with} \quad \boldsymbol{M} := \int_0^\infty e^{\boldsymbol{B}s}\boldsymbol{Q}e^{\boldsymbol{B}^Ts}ds$$

where $c_{\boldsymbol{M}} > 0$ is a normalization constant. It is well known that the above definition of $\boldsymbol{M}$ is equivalent to finding the unique solution to the continuous Lyapunov equation

$$\boldsymbol{Q} = -\boldsymbol{B}\boldsymbol{M} - \boldsymbol{M}\boldsymbol{B}^T. \tag{1.3.4}$$

(See for instance Theorem 2.2 in [20], §2.2 of [13].)
Hypoelliptic Ornstein-Uhlenbeck operators have been studied for many years, and more recently in [18] the authors considered them under the additional possibility of degeneracy in their diffusion matrix $\boldsymbol{Q}$. In [18], the authors described the domain of the closed operator $P_{\boldsymbol{Q},\boldsymbol{B}}$, and have found the following resolvent estimation:

Theorem 1.3.6. *Consider the hypoelliptic Ornstein-Uhlenbeck operator $P_{\boldsymbol{Q},\boldsymbol{B}}$, as in* (1.3.3)*, and its invariant measure $d\mu(x)$. Then there exist some positive constants $c, C > 0$ such that for any $z \in \Gamma_\kappa$, with*

$$\Gamma_\kappa := \left\{ z \in \mathbb{C} \,\middle|\, \operatorname{Re} z \le \tfrac{1}{2}\left(1 - \operatorname{Tr}(\boldsymbol{B})\right), \left|\operatorname{Re} z - \left(1 - \tfrac{1}{2}\operatorname{Tr}(\boldsymbol{B})\right)\right| \le c\left|z - \left(1 - \tfrac{1}{2}\operatorname{Tr}(\boldsymbol{B})\right)\right|^{\frac{1}{2\kappa+1}} \right\} \tag{1.3.5}$$

and where κ is the smallest integer $0 \le \kappa \le d-1$ such that

$$\operatorname{rank}\left[\boldsymbol{Q}^{\frac{1}{2}}, \boldsymbol{B}\boldsymbol{Q}^{\frac{1}{2}}, \dots, \boldsymbol{B}^{\kappa}\boldsymbol{Q}^{\frac{1}{2}}\right] = d\,, \tag{1.3.6}$$

one has that

$$\|\left(P_{\boldsymbol{Q},\boldsymbol{B}} - zI\right)^{-1}\|_{B(L^2(\mathbb{R}^d, d\mu))} \le C\left|z - \left(1 - \frac{1}{2}\operatorname{Tr}(\boldsymbol{B})\right)\right|^{-\frac{1}{2\kappa+1}}.$$

We illustrate the spectrum of $P_{\boldsymbol{Q},\boldsymbol{B}}$ and the domain Γ_κ in Figure 1.3.1.
In order to use the above theorem for our operator, L, we show the connection between it and P in the following lemma:

Lemma 1.3.7. *Assume that the associated diffusion and drift matrices for L, defined on $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$, and $P_{\boldsymbol{Q},\boldsymbol{B}}$, defined on $L^2\left(\mathbb{R}^d, d\mu(x)\right)$, satisfy*

$$\boldsymbol{Q} = 2\boldsymbol{D},\, \boldsymbol{B} = -\boldsymbol{C}.$$

Then $d\mu(x) = f_\infty(x)dx$ is the invariant measure for $P = P_{\boldsymbol{Q},\boldsymbol{B}}$ and its adjoint, and (up to the natural transformation $\frac{Lf}{f_\infty} = P^*(\frac{f}{f_\infty})$) we have $L = P^*$.

Proof. We start by recalling that we assume that $\boldsymbol{D} = \boldsymbol{C}_s$. Since (1.3.4) can be rewritten as

$$2\boldsymbol{D} = \boldsymbol{C}\boldsymbol{M} + \boldsymbol{M}\boldsymbol{C}^T$$

for our choice of $\boldsymbol{Q}$ and $\boldsymbol{B}$, we conclude that $\boldsymbol{M} = \boldsymbol{I}$ for $P_{2\boldsymbol{D},-\boldsymbol{C}}$ and that $\left(P_{2\boldsymbol{D},-\boldsymbol{C}}\right)^* = P_{2\boldsymbol{D},-\boldsymbol{C}^T}$ (the last equality can be shown in a similar way to (1.2.4)). Thus, the invariant measure corresponding to both these operators is $f_\infty(x)dx$.
Let $f \in D(L) \subset L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$ and define $g_f := \frac{f}{f_\infty} \in L^2\left(\mathbb{R}^d, f_\infty\right)$. Then

$$\begin{aligned}\frac{L_{\boldsymbol{D},\boldsymbol{C}} f(x)}{f_\infty(x)} &= \frac{\operatorname{div}\left(f_\infty(x)\boldsymbol{C}\nabla g_f(x)\right)}{f_\infty(x)} = \operatorname{div}\left(\boldsymbol{C}\nabla g_f(x)\right) + \frac{\nabla f_\infty(x)^T \boldsymbol{C}\nabla g_f(x)}{f_\infty(x)} \\ &= \operatorname{div}\left(\boldsymbol{D}\nabla g_f(x)\right) - x^T\boldsymbol{C}\nabla g_f(x) = P_{2\boldsymbol{D},-\boldsymbol{C}^T} g_f(x) = \left(P_{2\boldsymbol{D},-\boldsymbol{C}}\right)^* g_f(x),\end{aligned} \tag{1.3.7}$$

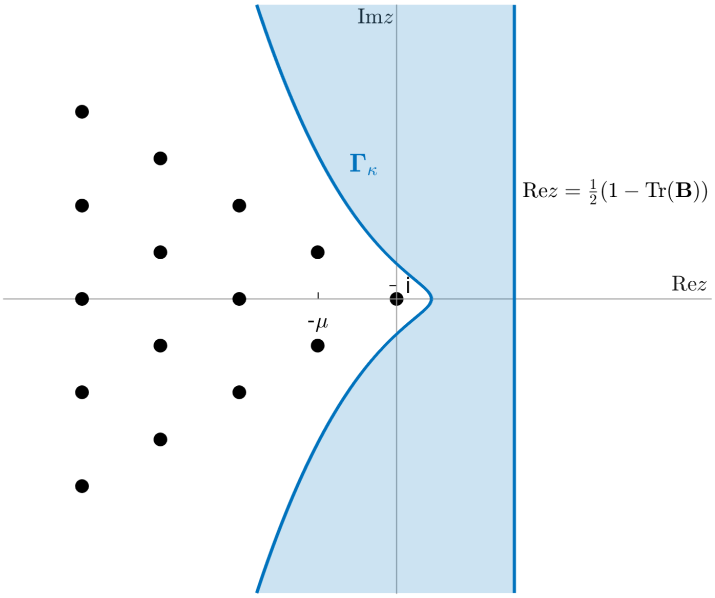


Figure 1.3.1: The black dots represent $\sigma(P_{\boldsymbol{Q},\boldsymbol{B}})$ with the eigenvalues of the 2×2 matrix $\boldsymbol{B}$ given as $\lambda_{1,2} = -1 \pm \frac{7}{2}i$. The shaded area represents the set Γ_κ of Theorem 1.3.6 with $\kappa = 1$.

where the adjoint is considered w.r.t. $L^2\left(\mathbb{R}^d, f_\infty\right)$. In particular, if $f(t,\cdot) \in L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$ solves (1.1.2) then $g_f(t,\cdot)$ satisfies the adjoint equation $\partial_t g_f = \left(P_{2\boldsymbol{D},-\boldsymbol{C}}\right)^* g_f$. □

With this at hand we can recast, and improve, Theorem 1.3.6 for the operator L and its closure.

Proposition 1.3.8. *Let any $k \in \mathbb{N}_0$ be fixed. Consider the set Γ_κ, defined by* (1.3.5), *associated to $\boldsymbol{Q} = 2\boldsymbol{D}$, $\boldsymbol{B} = -\boldsymbol{C}^T$ (Condition (C) guarantees the existence of such κ). Then we have that, for any $z \in \Gamma_\kappa$, the operator $(L - zI)|_{H_k} : H_k \to H_k$ is well defined, closable, and its closure is invertible with*

$$\left\| \left(\left(\overline{L} - zI \right) |_{H_k} \right)^{-1} \right\|_{B(H_k)} \le C \left| z - \left(1 + \frac{1}{2} \text{Tr}(\boldsymbol{C}) \right) \right|^{-\frac{1}{2\kappa+1}}, \tag{1.3.8}$$

where $C > 0$ is the same constant as in Theorem 1.3.6.

Proof. We consider the case $k = 0$ first. Due to Theorem 1.3.6 we know that for any $z \in \Gamma_\kappa$, $P_{2\boldsymbol{D},-\boldsymbol{C}^T} - zI$ is invertible on $L^2\left(\mathbb{R}^d, f_\infty\right)$. Hence, for any $f \in L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$ there exists a unique $\ell_f \in L^2\left(\mathbb{R}^d, f_\infty\right)$ such that

$$\left(P_{2\boldsymbol{D},-\boldsymbol{C}^T} - zI\right) \ell_f(x) = \frac{f(x)}{f_\infty(x)},$$

which can also be written differently due to (1.3.7), as

$$\left(\overline{L}-zI\right)\left(f_\infty(x)\ell_f(x)\right)=f(x).$$

This implies that $\overline{L}-zI$ is bijective on its appropriate space.
Next we notice that, with the notations from Lemma 1.3.7

$$\sup_{\|f\|=1}\|\left(\overline{L}-zI\right)^{-1}f\|_{L^2(\mathbb{R}^d,f_\infty^{-1})}=\sup_{\|f\|=1}\|f_\infty\ell_f\|_{L^2(\mathbb{R}^d,f_\infty^{-1})}$$

$$=\sup_{\|f\|=1}\|\ell_f\|_{L^2(\mathbb{R}^d,f_\infty)}=\sup_{\|g_f\|=1}\|\left(P_{2\boldsymbol{D},-\boldsymbol{C}^T}-zI\right)^{-1}g_f\|_{L^2(\mathbb{R}^d,f_\infty)},$$

from which we conclude that

$$\|\left(\overline{L}-zI\right)^{-1}\|_{B(L^2(\mathbb{R}^d,f_\infty^{-1}))}=\|\left(P_{2\boldsymbol{D},-\boldsymbol{C}^T}-zI\right)^{-1}\|_{B(L^2(\mathbb{R}^d,f_\infty))},$$

completing the proof for this case.
We now turn our attention to the restrictions $(L-zI)|_{H_k}$ with $k\geq 1$ and domain

$$D_k:=\operatorname{span}\{V_m,m\geq k\}=D(L)\cap H_k.$$

Since $L|_{V_m}:V_m\to V_m$ $\forall m\in\mathbb{N}_0$ we have that $(L-zI)|_{H_k}:D_k\to H_k$. Moreover, the dissipativity of L on $D(L)$ assures us that L is dissipative, and as such closable, on the Hilbert space H_k. Thus $(L-zI)|_{H_k}$ is closable too and

$$\overline{(L-zI)|_{H_k}}=\left(\overline{L}-zI\right)|_{H_k}.$$

Additionally, since the only part of $L^2\left(\mathbb{R}^d,f_\infty^{-1}\right)$ that is not in H_k is a finite dimensional subspace of $D(L)$, we can conclude that

$$D\left((\overline{L}-zI)|_{H_k}\right)=D(\overline{L})\cap H_k.$$

Given z in the resolvent set of $\overline{L}$ we know that $\overline{L}-zI|_{V_m}:V_m\to V_m$ is invertible for any m and as such

$$(\overline{L}-zI)|_{V_m}(V_m)=V_m.$$

Thus,

$$V_m\subset\operatorname{Range}\left((\overline{L}-zI)|_{H_k}\right),\qquad\forall m\geq k.$$

We conclude that $(\overline{L}-zI)|_{H_k}$ is injective with a dense range in H_k for any $z\in\Gamma_\kappa$, and hence invertible on its range. The validity of (1.3.8) for $k=0$ allows us to extend our inverse to H_k with the same *uniform bound* as is given in (1.3.8). The general case is now proved. □

From this point onward, we will assume that we are dealing with the closed operator $\overline{L}$ and with its appropriate domain (that includes $\bigcup_{m\in\mathbb{N}_0} V_m$) when we consider our equation. We will also write L instead of $\overline{L}$ in what is to follow.
Lemma 1.3.5 and Proposition 1.3.8 are all the tools we need to estimate the uniform exponential stability of our evolution semigroup on each H_k, an estimation that is crucial to show Theorem 1.3.2.

Proposition 1.3.9. *Consider the Fokker–Planck operator L, defined on $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$, and the spaces $\{H_k\}_{k\geq 1}$ defined in* (1.3.1)*. Then, for any $0<\varepsilon<\mu$, the semigroup generated by the operator $L+\left(k\mu-\varepsilon\right)I|_{H_k}$, with domain $D(L)\cap H_k$, is uniformly exponentially stable. I.e., there exists some geometric constant $C_{k,\varepsilon}>0$ such that*

$$\|e^{Lt}\|_{B(H_k)} \leq C_{k,\varepsilon} e^{-(k\mu-\varepsilon)t}, \quad t\geq 0. \tag{1.3.9}$$

Proof. We will show that

$$M_{k,\varepsilon} := \sup_{\operatorname{Re} z>0} \left\| \left(\left(L+[k\mu-\varepsilon]I\right) - zI\right)^{-1} \right\|_{B(H_k)} < \infty,$$

and conclude the result from the fact that L generates a contraction semigroup according to Lemma 1.3.5 and the Gearhart-Prüss Theorem.
The study of upper bounds for the resolvents of $L+[k\mu-\varepsilon]I$ in the right-hand complex plane relies on subdividing this domain into several pieces. This is illustrated in Figure 1.3.2, which we will refer to during the proof to help visualise this division.
Since L generates a contraction semigroup, for any $\varepsilon>0$, $L-\varepsilon I$ generates a semigroup that is uniformly exponentially stable on $L^2(\mathbb{R}^d, f_\infty^{-1})$. The Gearhart-Prüss Theorem applied to $L-\varepsilon I$ implies that

$$\widetilde{M}_{k,\varepsilon} := \sup_{\operatorname{Re} z>0} \|(L-(\varepsilon+z)I)^{-1}\|_{B(H_k)} \leq \sup_{\operatorname{Re} z>0} \|(L-(\varepsilon+z)I)^{-1}\|_{B(L^2(\mathbb{R}^d, f_\infty^{-1}))} < \infty,$$

where we removed the subscript H_k from the operator on the left-hand side to simplify notations.
Since

$$L-(\varepsilon+z)I = L+[k\mu-\varepsilon]I - \left(z+k\mu\right)I,$$

we see that

$$\begin{aligned}\widetilde{M}_{k,\varepsilon} &= \sup_{\operatorname{Re} z_1>0} \left\| \left(\left(L+[k\mu-\varepsilon]I\right) - \left(z_1+k\mu\right)I\right)^{-1} \right\|_{B(H_k)} \\ &= \sup_{\operatorname{Re} z>k\mu} \left\| \left(\left(L+[k\mu-\varepsilon]I\right) - zI\right)^{-1} \right\|_{B(H_k)}\end{aligned}$$

(this term corresponds to the right-hand side of the dashed line in Figure 1.3.2).
From the above we conclude that

$$M_{k,\varepsilon} = \max\left(\widetilde{M}_{k,\varepsilon}, \sup_{0<\operatorname{Re} z\leq k\mu} \left\| \left(L-[z-k\mu+\varepsilon]I\right)^{-1} \right\|_{B(H_k)} \right),$$

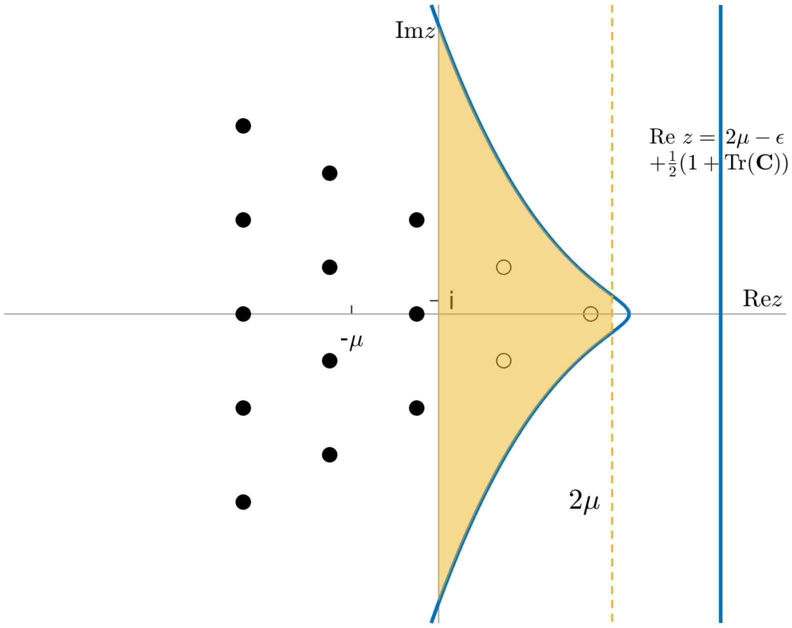


Figure 1.3.2: choosing $k = 2$, the solid dots represent $\sigma((L+[2\mu-\varepsilon]I)|_{H_2})$ where the eigenvalues of the 2 × 2 matrix C are given by $\lambda_{1,2} = 1 \pm \frac{7}{2}i$. The empty dots are the eigenvalues of the operator $L+[2\mu-\varepsilon]I$ that disappear due to the restriction to H_2, and the shaded area represents the compact set $\{z\in\mathbb{C} \mid 0\leq \mathrm{Re}\, z\leq 2\mu\}\cap\{z\not\in\Gamma_\kappa+2\mu-\varepsilon\}$ where $\kappa=1$.

which implies that we only need to show that the second term in the parenthesis is finite (this term corresponds to the area between the dashed line and the imaginary axis in Figure 1.3.2).

Using Proposition 1.3.8 we conclude that

$$\sup_{z-k\mu+\varepsilon\in\Gamma_\kappa}\left\|\left(L-\left[z-k\mu+\varepsilon\right]I\right)^{-1}\right\|<\infty$$

(represented in Figure 1.3.2 by the domain between the two solid blue curves). We conclude that $M_{k,\varepsilon}<\infty$ if and only if

$$\sup_{\{0<\mathrm{Re}\, z\leq k\mu\}\cap\{z\not\in\Gamma_\kappa+k\mu-\varepsilon\}}\left\|\left(L-[z-k\mu+\varepsilon]I\right)^{-1}\right\|_{B(H_k)}<\infty.$$

Since $\mathrm{Re}\, z = -\varepsilon$ is the closest vertical line to $\mathrm{Re}\, z = 0$ which intersects $\sigma\left(\left(L+[k\mu-\varepsilon]I\right)|_{H_k}\right)$, we notice that $\left\{0<\mathrm{Re}\, z\leq k\mu\right\}\cap\left\{z\not\in\Gamma_\kappa+k\mu-\varepsilon\right\}$ (represented by the shaded area in Figure 1.3.2) is a compact set in the resolvent set of

$\left(L+[k\mu-\varepsilon]I\right)|_{H_k}$. As the resolvent map is analytic on the resolvent set, we conclude that $M_{k,\varepsilon}<\infty$, completing the proof. □

Remark 1.3.10. *While the constant mentioned in* (1.3.9) *is a fixed geometric one, the original Gearhart-Prüss theorem doesn't give an estimation for it. However, recent studies have improved the original theorem and have managed to find explicit expression for this constant by paying a small price in the exponential power. As we can afford to "lose" another small ε, we could use references such as [11, 14] to have a more concrete expression for $C_{k,\varepsilon}$. We will avoid giving such an expression in this work to simplify its presentation.*

We finally have all the tools to show Theorem 1.3.2:

Proof of Theorem 1.3.2. Using the invariance of V_0 and H_k under L and Proposition 1.3.9 we find that for any $f_k\in H_k$

$$\begin{aligned}e_2\left(e^{Lt}\left(f_k+f_\infty\right)|f_\infty\right)&=e_2\left(e^{Lt}\left(f_k\right)+f_\infty|f_\infty\right)=\frac{1}{2}\left\|e^{Lt}f_k\right\|^2_{H_k}\\&\leq\frac{1}{2}C^2_{k,\varepsilon}e^{-2(k\mu-\varepsilon)t}\left\|f_k\right\|^2_{H_k}=C^2_{k,\varepsilon}e^{-2(k\mu-\varepsilon)t}e_2\left(f_k+f_\infty|f_\infty\right),\end{aligned}$$

showing the desired result. □

Theorem 1.3.2 has given us the ability to control the rate of convergence to equilibrium of functions with initial data that, up to f_∞, live on a "higher eigenspace". Can we use this information to understand what happens to the solution of an arbitrary initial datum $f_0\in L^2\left(\mathbb{R}^d,f_\infty^{-1}\right)$ with unit mass?

The answer to this question is *Yes.*

Since for any $k\geq 1$

$$L^2\left(\mathbb{R}^d,f_\infty^{-1}\right)=V_0\oplus\left(\bigoplus_{m=1}^{k}V_m\right)\oplus H_{k+1}$$

and the Fokker–Planck semigroup is invariant under all the above spaces, we are motivated to *split* the solution of our equation into a part in $V_0\oplus H_{k+1}$ and a part in $\bigoplus_{m=1}^{k}V_m$ - which is a *finite dimensional subset of* $D(L)$. As we now know that decay in $\bigoplus_{m=1}^{k}V_m$ is slower than that for H_{k+1} we will obtain a *sharp* rate of convergence to equilibrium. We summarise the above intuition in the following theorem:

Theorem 1.3.11. *Consider the Fokker–Planck equation* (1.1.2) *with diffusion and drift matrices satisfying Conditions (A)–(C). Let $f_0\in L^1_+\left(\mathbb{R}^d\right)\cap L^2\left(\mathbb{R}^d,f_\infty^{-1}\right)$ be a given function with unit mass such that*

$$f_0=f_\infty+f_{k_0}+\tilde{f}_{k_0},$$

where $f_{k_0}\in V_{k_0}$ is non-zero and $\tilde{f}_{k_0}\in H_{k_0+1}$. Denote by $[L]_{k_0}$ the matrix representation of L with respect to an orthonormal basis of V_{k_0} and let

$$n_{k_0}:=\max\left\{\textit{defect of }\lambda\mid\lambda\textit{ is an eigenvalue of }[L]_{k_0}\textit{ and }\operatorname{Re}\lambda=-k_0\mu\right\},$$

where μ is defined in (1.1.5). *Then, there exists a geometric constant c_{k_0}, which is independent of f_0, such that*

$$e_2\left(f(t)|f_\infty\right) \le c_{k_0} e_2\left(f_0|f_\infty\right)\left(1+t^{2n_{k_0}}\right)e^{-2k_0\mu t}. \tag{1.3.10}$$

Remark 1.3.12. *As can be seen in the proof of the theorem, the sign of f_0 plays no role. As such, the theorem could have been stated for $f_0 \in L^1\left(\mathbb{R}^d\right) \cap L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$. We decided to state it as is since it is the form we will use later on, and we wished to avoid possible confusion.*

Proof of Theorem 1.3.11. Due to the invariance of all V_m under L we see that

$$f(t) = f_\infty + e^{Lt} f_{k_0} + e^{Lt}\tilde{f}_{k_0},$$

with $e^{Lt} f_{k_0} \in V_{k_0}$ and $e^{Lt}\tilde{f}_{k_0} \in H_{k_0+1}$. From Theorem 1.3.2 we conclude that

$$e_2\left(f_\infty + e^{Lt}\left(\tilde{f}_{k_0}\right)|f_\infty\right) \le c_{k_0,\varepsilon} e_2\left(f_\infty + \tilde{f}_{k_0}|f_\infty\right)e^{-2((k_0+1)\mu-\varepsilon)t},$$

for any $0<\varepsilon<\mu$.

Next, we denote by $d_k := \dim(V_k)$ and let $\{\xi_i\}_{i=1,\dots,d_{k_0}}$ be an orthonormal basis for V_{k_0}. The invariance of V_m under L implies that we can write

$$e^{Lt} f_{k_0} = \sum_{i=1}^{d_{k_0}} a_i(t)\xi_i$$

with $\boldsymbol{a}(t) := \left(a_1(t),\dots,a_{d_{k_0}}(t)\right)$ satisfying the simple ODE

$$\dot{\boldsymbol{a}}(t) = [L]_{k_0}^T \boldsymbol{a}(t).$$

This, together with the definition of n_{k_0} and the fact that a matrix and its transpose share eigenvalues and defect numbers, implies that we can find a geometric constant that depends only on k_0 such that

$$\sum_{i=1}^{d_{k_0}} a_i^2(t) \le c_{k_0}\left(1+t^{2n_{k_0}}\right)e^{-2k_0\mu t}\sum_{i=1}^{d_{k_0}} a_i^2(0). \tag{1.3.11}$$

Since

$$e_2\left(f(t)|f_\infty\right) = e_2\left(f_\infty + e^{Lt}(\tilde{f}_{k_0}) + e^{Lt}(f_{k_0})|f_\infty\right) = \frac{1}{2}\left\|e^{Lt}(\tilde{f}_{k_0}) + e^{Lt}(f_{k_0})\right\|^2_{L^2(\mathbb{R}^d, f_\infty^{-1})}$$

$$= \frac{1}{2}\left\|e^{Lt}(\tilde{f}_{k_0})\right\|^2_{L^2(\mathbb{R}^d, f_\infty^{-1})} + \frac{1}{2}\left\|\sum_{i=1}^{d_{k_0}} a_i(t)\xi_i\right\|^2_{L^2(\mathbb{R}^d, f_\infty^{-1})}$$

$$= e_2\left(f_\infty + e^{Lt}(\tilde{f}_{k_0})|f_\infty\right) + \frac{1}{2}\sum_{i=1}^{d_{k_0}} a_i(t)^2,$$

we see, by combining Theorem 1.3.2 and (1.3.11) that

$$\begin{aligned} e_2\left(f(t)|f_\infty\right) \leq & c_{k_0,\varepsilon} e_2\left(f_\infty + \tilde{f}_{k_0}|f_\infty\right) e^{-2((k_0+1)\mu-\varepsilon)t} \\ & + \frac{c_{k_0}}{2}\sum_{i=1}^{d_{k_0}} a_i^2(0)\left(1+t^{2n_{k_0}}\right)e^{-2k_0\mu t}. \end{aligned}$$

Hence

$$e_2\left(f(t)|f_\infty\right) \leq \max\left(c_{k_0,\varepsilon}, c_{k_0}\right)\left(e_2\left(f_\infty + \tilde{f}_{k_0}|f_\infty\right) + \frac{1}{2}\left\|f_{k_0}\right\|^2_{L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)}\right)\left(1+t^{2n_{k_0}}\right)e^{-2k_0\mu t}.$$

This completes the proof, as we have seen that

$$e_2(f_0|f_\infty) = e_2\left(f_\infty + \tilde{f}_{k_0}|f_\infty\right) + \frac{1}{2}\left\|f_{k_0}\right\|^2_{L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)}.$$

□

Remark 1.3.13. *The idea to* split *a solution into a few parts is viable* only *for the* 2-*entropy. The reason behind it is that such splitting, regardless of whether or not it can be done to functions outside of* $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$, *will most likely create functions without a definite sign. These functions can not be explored using the* p*-entropy with* $1 < p < 2$.

Theorem 1.3.11 gives an optimal rate of decay for the 2−entropy. However, one can underestimate the rate of decay by using Theorem 1.3.2 and remove the condition $f_{k_0} \neq 0$ to obtain the following:

Corollary 1.3.14. The statement of Theorem 1.3.11 remains valid when replacing k_0 by any $1 \leq k_1 \leq k_0$. However, the decay estimate (1.3.10) will not be sharp when $k_1 < k_0$.

Proof of Theorem 1.1.4 for $p = 2$. The proof follows immediately from Corollary 1.3.14 for $k_1 = 1$. □

Now that we have learned everything we can on the convergence to equilibrium for e_2, we can proceed to understand the convergence to equilibrium of e_p.

1.4 Non-symmetric Hypercontractivity and the p-Entropy

In this section we will show how to deduce the rate of convergence to equilibrium for the family of p-entropies, with $1 < p < 2$, from e_2. The main thing that will make the above possible is *a non-symmetric hypercontractivity* property of our Fokker–Planck equation

- namely, that any solution to the equation with (initially only) a finite p-entropy will eventually be "pushed" into $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$, at which point we can use the information we gained on e_2.
Before we show this result, and see how it implies our main theorem, we explain why and how this non-symmetric hypercontractivity helps.

Lemma 1.4.1. *Let $f \in L^1_+\left(\mathbb{R}^d\right)$ with unit mass. Then*

(i)

$$e_p(f|f_\infty) = \frac{1}{p(p-1)}\left(\|f\|^p_{L^p\left(\mathbb{R}^d, f_\infty^{1-p}\right)} - 1\right).$$

(ii) for any $1 < p_1 < p_2 \leq 2$ there exists a constant $C_{p_1,p_2} > 0$ such that

$$e_{p_1}(f|f_\infty) \leq C_{p_1,p_2} e_{p_2}(f|f_\infty).$$

In particular, for any $1 < p < 2$

$$e_p(f|f_\infty) \leq C_p e_2(f|f_\infty),$$

for a fixed geometric constant.

Proof. (i) is trivial. To prove (ii) we consider the function

$$g(y) := \begin{cases} \frac{p_2(p_2-1)}{p_1(p_1-1)} \frac{y^{p_1} - p_1(y-1) - 1}{y^{p_2} - p_2(y-1) - 1}, & y \geq 0, y \neq 1 \\ 1, & y = 1. \end{cases}$$

Clearly $g \geq 0$ on $\mathbb{R}^+$, and it is easy to check that it is continuous. Since we have $\lim_{y\to\infty} g(y) = 0$, we can conclude the result using (1.1.4). □

It is worth to note that the second point of part (ii) of Lemma 1.4.1 can be extended to general generating function for an admissible relative entropy. The following is taken from [3]:

Lemma 1.4.2. *Let ψ be a generating function for an admissible relative entropy. Then one has that*

$$\psi(y) \leq 2\psi''(1)\psi_2(y), \quad y \geq 0.$$

In particular $e_p \leq 2e_2$ for any $1 < p < 2$ whenever e_2 is finite.

Lemma 1.4.1 assures us that, if we start with initial data in $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$, then e_p will be finite. Moreover, due to Theorem 1.1.4 for $p = 2$, and the fact that the solution to (1.1.2) remains in $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$, we have that

$$e_p(f(t)|f_\infty) \leq 2e_2(f(t)|f_\infty) \leq Ce_2(f_0|f_\infty)\left(1 + t^{2n}\right)e^{-2\mu t}.$$

However, one can easily find initial data $f_0 \notin L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$ with finite p-entropies. If one can show that the flow of the Fokker–Planck equation eventually forces the solution to enter $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$, we would be able to utilise the idea we just presented, at least from that time on.
This *explicit non-symmetric* hypercontractivity result we desire, is the main new theorem we present in this section.

Theorem 1.4.3. *Consider the Fokker–Planck equation* (1.1.2) *with diffusion and drift matrices* $\boldsymbol{D}$ *and* $\boldsymbol{C}$ *satisfying Conditions (A)–(C). Let* $f_0 \in L^1_+\left(\mathbb{R}^d\right)$ *be a function with unit mass and assume there exists* $\varepsilon > 0$ *such that*

$$\int_{\mathbb{R}^d} e^{\varepsilon|x|^2} f_0(x)dx < \infty. \tag{1.4.1}$$

(i) Then, for any $q > 1$*, there exists an explicit* $t_0 > 0$ *that depends only on geometric constants of the problem such that the solution to* (1.1.2) *satisfies*

$$\int_{\mathbb{R}^d} f(t,x)^q f_\infty^{-1}(x)dx \le \left(\frac{q}{\pi(q+1)}\right)^{\frac{qd}{2}} \left(\frac{8\pi^2}{q-1}\right)^{\frac{d}{2}} \left(\int_{\mathbb{R}^d} e^{\varepsilon|x|^2} f_0(x)dx\right)^q \tag{1.4.2}$$

for all $t \ge t_0$.

(ii) In particular, if f_0 *satisfies* $e_p(f_0|f_\infty) < \infty$ *for some* $1 < p < 2$ *we have that*

$$e_2(f(t)|f_\infty) \le \frac{1}{2}\left(\left(\frac{8\sqrt{2}}{3\cdot 2^{\frac{1}{p}}}\right)^d \left(p(p-1)e_p(f_0|f_\infty)+1\right)^{\frac{2}{p}} - 1\right), \tag{1.4.3}$$

for $t \ge \tilde{t}_0(p) > 0$*, which can be given explicitly.*

Remark 1.4.4. *As we consider* e_p *in our hypercontractivity, which is, up to a constant, the* L^p *norm of* $g := \frac{f}{f_\infty}$ *with the measure* $f_\infty(x)dx$*, one can view our result as a hypercontractivity property of the Ornstein-Uhlenbeck operator, P (for an appropriate choice of the diffusion matrix* $\boldsymbol{Q}$ *and drift matrix* $\boldsymbol{B}$*), discussed in §1.3. With this notation,* (1.4.3) *is equivalent to*

$$\|g(t)\|_{L^2(f_\infty)} \le C_{p,d}\|g_0\|_{L^p(f_\infty)}, \quad t \ge \tilde{t}_0(p) \tag{1.4.4}$$

for $1 < p < 2$*, where* $C_{p,d} := \left(\frac{8\sqrt{2}}{3\cdot 2^{\frac{1}{p}}}\right)^{\frac{d}{2}}$*. Since* e_2 *decreases along the flow of our equation,* (1.4.4) *is valid for* $p = 2$ *with* $C_{2,d} = 1$*. Thus, by using the Riesz-Thorin theorem one can improve inequality* (1.4.4) *to the same inequality with the constant* $C_{p,d}^{\frac{2}{p}-1}$*. We would like to point out at this point that a simple limit process shows that* (1.4.4) *is also valid for* $p = 1$*, but there is no connection between the* L^1 *norm of* g *and the Boltzmann entropy,* e_1*, of* f_0.

Remark 1.4.5. *Since its original definition for the Ornstein-Uhlenbeck semigroup in the work of Nelson, [16], the notion of* hypercontractivity *has been studied extensively for Markov diffusive operators (implying selfadjointness). A contemporary review of this topic can be found in [4]. For such selfadjoint generators, hypercontractivity is equivalent to the validity of a logarithmic Sobolev inequality, as proved by Gross [10]. For non-symmetric generators, however, this equivalence does not hold: While a log Sobolev inequality still implies hypercontractvity of related semigroups (cf. the proof of Theorem 5.2.3 in [4]), the reverse implication is not true in general (cf. Remark 5.1.1 in [22]). In particular, hypocoercive degenerate parabolic equations cannot give rise to a log Sobolev inequality, but they may exhibit hypercontractivity (as just stated above).*
The last 20 years have seen the emergence of the, more delicate, study of hypercontractivity for non-symmetric and even degenerate semigroups. Notable works in the field are the paper of Fuhrman, [9], and more recently the work of Wang et al., [6, 7, 21]. Most of these works consider an abstract Hilbert space as an underlying domain for the semigroup, and to our knowledge none of them give an explicit time after which one can observe the hypercontractivity phenomena (Fuhrman gives a condition on the time in [9]).
Our hypercontractivity theorem, which we will prove shortly, gives not only an explicit and quantitative inequality, but also provides an estimation on the time one needs to wait before the hypercontractivity occurs. To keep the formulation of Theorem 1.4.3 simple we did not include this "waiting time" there, but we emphasised it in its proof. Moreover, the hypercontractivity estimate from Theorem 1.4.3(i) only requires (1.4.1)*, a weighted L^1 norm of f_0. This is weaker than in usual hypercontractivity estimates, which use L^p norms as on the r.h.s. of* (1.4.4)*.*

It is worth to note that we prove our theorem under the setting of the e_p entropies, which can be thought of as L^p spaces with a weight function that depends on p.

In order to be able to prove Theorem 1.4.3 we will need a few technical lemmas.

Lemma 1.4.6. *Given $f_0 \in L^1_+\left(\mathbb{R}^d\right)$ with unit mass, the solution to the Fokker–Planck equation* (1.1.2) *with diffusion and drift matrices $\boldsymbol{D}$ and $\boldsymbol{C}$ that satisfy Conditions (A)–(C) is given by*

$$f(t,x) = \frac{1}{(2\pi)^{\frac{d}{2}}\sqrt{\det \boldsymbol{W}(t)}}\int_{\mathbb{R}^d} e^{-\frac{1}{2}\left(x-e^{-\boldsymbol{C}t}y\right)^T \boldsymbol{W}(t)^{-1}\left(x-e^{\boldsymbol{C}t}y\right)} f_0(y)dy, \tag{1.4.5}$$

where

$$\boldsymbol{W}(t) := 2\int_0^t e^{-\boldsymbol{C}s}\boldsymbol{D}e^{-\boldsymbol{C}^T s}ds.$$

This is a well known result, see for instance §1 in [12] or §6.5 in [19].

Lemma 1.4.7. *Assume that the diffusion and drift matrices, $\boldsymbol{D}$ and $\boldsymbol{C}$, satisfy Conditions (A)–(C), and let $\boldsymbol{K}$ be the unique positive definite matrix that satisfies*

$$2\boldsymbol{D} = \boldsymbol{C}\boldsymbol{K} + \boldsymbol{K}\boldsymbol{C}^T.$$

Then (in any matrix norm)

$$\|\boldsymbol{W}(t)-\boldsymbol{K}\| \leq c(1+t^{2n})e^{-2\mu t}, \quad t \geq 0,$$

where $c>0$ *is a geometric constant depending on* n *and* μ*, with* n *being the maximal defect of the eigenvalues of* $\boldsymbol{C}$ *with real part* μ*, defined in* (1.1.5).

Proof. We start the proof by noticing that $\boldsymbol{K}$ is given by

$$\boldsymbol{K} = 2\int_0^\infty e^{-\boldsymbol{C}s}\boldsymbol{D}e^{-\boldsymbol{C}^T s}ds$$

(see for instance [18]). As such

$$\|\boldsymbol{W}(t)-\boldsymbol{K}\| \leq 2\int_t^\infty \|e^{-\boldsymbol{C}s}\boldsymbol{D}e^{-\boldsymbol{C}^T s}\|ds \leq 2\|\boldsymbol{D}\|\int_t^\infty \|e^{-\boldsymbol{C}s}\|\|e^{-\boldsymbol{C}^T s}\|ds.$$

Using the fact that

$$\boldsymbol{A}e^{-\boldsymbol{C}t}\boldsymbol{A}^{-1} = e^{-\boldsymbol{A}\boldsymbol{C}\boldsymbol{A}^{-1}t}$$

for any regular matrix $\boldsymbol{A}$, we conclude that, if $\boldsymbol{J}$ is the Jordan form of $\boldsymbol{C}$, then

$$\|e^{-\boldsymbol{C}t}\| \leq \|\boldsymbol{A}_{\boldsymbol{J}}\|\|\boldsymbol{A}_{\boldsymbol{J}}^{-1}\|\|e^{-\boldsymbol{J}t}\|, \tag{1.4.6}$$

where $\boldsymbol{A}_{\boldsymbol{J}}$ is the similarity matrix between $\boldsymbol{C}$ and its Jordan form.
For a single Jordan block of size $n+1$ (corresponding to a defect of n in the eigenvalue λ), $\widetilde{\boldsymbol{J}}$, we find that

$$e^{\widetilde{\boldsymbol{J}}t} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \dots & \frac{t^n}{n!}e^{\lambda t} \\ & e^{\lambda t} & \ddots & \frac{t^{n-1}}{(n-1)!}e^{\lambda t} \\ & & \ddots & \vdots \\ 0 & & & e^{\lambda t} \end{pmatrix} \quad \text{where} \quad \widetilde{\boldsymbol{J}} = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & & 1 \\ 0 & & & \lambda \end{pmatrix}.$$

Thus, we conclude that

$$\|e^{\widetilde{\boldsymbol{J}}t}x\|_1 \leq \sum_{i=1}^{n+1}\sum_{j=i}^{n+1}\frac{t^{j-i}}{(j-i)!}e^{\mathrm{Re}(\lambda)t}\left|x_j\right| \leq \left(\sum_{i=1}^{n+1}\left(1+t^n\right)e^{\mathrm{Re}(\lambda)t}\right)\|x\|_1$$

$$= (n+1)\left(1+t^n\right)e^{\mathrm{Re}(\lambda)t}\|x\|_1, \quad t \geq 0.$$

Due to the equivalence of norms on finite dimensional spaces, there exists a geometric constant $c_1>0$, that depends on n, such that

$$\|e^{\widetilde{\boldsymbol{J}}t}\| \leq c_1\left(1+t^n\right)e^{\mathrm{Re}(\lambda)t}. \tag{1.4.7}$$

Coming back to $\boldsymbol{C}$, we see that the above inequality together with (1.4.6) imply that $\|e^{-\boldsymbol{C}t}\|$ is controlled by the norm of $\boldsymbol{C}$'s largest (measured by the defect number) Jordan block of the eigenvalue with smallest real part. From this, and (1.4.7), we conclude that

$$\|e^{-\boldsymbol{C}t}\| \leq c_2(1+t^n)e^{-\mu t}, \quad t \geq 0. \tag{1.4.8}$$

The same estimation for $\|e^{-\boldsymbol{C}^T t}\|$ implies that

$$\|\boldsymbol{W}(t) - \boldsymbol{K}\| \leq c_3 \int_t^\infty \left(1+s^{2n}\right) e^{-2\mu s} ds,$$

for some geometric constant $c_3 > 0$ that depends on n. Since

$$\int_t^\infty s^{2n} e^{-2\mu s} ds = \left[\frac{1}{2\mu} t^{2n} + \frac{2n}{(2\mu)^2} t^{2n-1} + \frac{2n(2n-1)}{(2\mu)^3} t^{2n-2} + \ldots + \frac{(2n)!}{(2\mu)^{2n+1}}\right] e^{-2\mu t}$$

we conclude the desired result. □

While we can continue with a general matrix $\boldsymbol{K}$, it will simplify our computations greatly if $\boldsymbol{K}$ would have been $\boldsymbol{I}$. Since we are working under the assumption that $\boldsymbol{D} = \boldsymbol{C}_S$, the normalization from Theorem 1.2.5 implies exactly that. Thus, from this point onwards we will assume that $\boldsymbol{K}$ is $\boldsymbol{I}$.

Lemma 1.4.8. *For any* $\varepsilon > 0$ *there exists an explicit* $t_1 > 0$ *such that for all* $t \geq t_1$

$$\|\boldsymbol{W}^{-1}(t) - \boldsymbol{I}\| \leq \varepsilon,$$

where $\boldsymbol{W}(t)$ *is as in Lemma 1.4.7. An explicit, but not optimal choice for* t_1 *is given by*

$$t_1(\varepsilon) := \frac{1}{2(\mu-\alpha)} \log\left(\frac{c(1+\varepsilon)\left(1+\left(\frac{n}{\alpha e}\right)^{2n}\right)}{\varepsilon}\right), \tag{1.4.9}$$

where $0 < \alpha < \mu$ *is arbitrary and* $c > 0$ *is given by Lemma 1.4.7.*

Proof. We have that for any invertible matrix $\boldsymbol{A}$

$$\|\boldsymbol{A}^{-1} - \boldsymbol{I}\| = \|(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{A}^{-1}\| \leq \|\boldsymbol{A}-\boldsymbol{I}\|\|\boldsymbol{A}^{-1}\|.$$

In addition, if $\|\boldsymbol{A}-\boldsymbol{I}\| < 1$, then

$$\|\boldsymbol{A}^{-1}\| = \|(\boldsymbol{I}-(\boldsymbol{I}-\boldsymbol{A}))^{-1}\| \leq \frac{1}{1-\|\boldsymbol{A}-\boldsymbol{I}\|}.$$

Thus, for any $t > 0$ such that $\|\boldsymbol{W}(t) - \boldsymbol{I}\| < 1$ we have that

$$\|\boldsymbol{W}^{-1}(t) - \boldsymbol{I}\| \leq \frac{\|\boldsymbol{W}(t)-\boldsymbol{I}\|}{1-\|\boldsymbol{W}(t)-\boldsymbol{I}\|}. \tag{1.4.10}$$

Defining $\tilde{t}_1(\varepsilon)$ as

$$\tilde{t}_1(\varepsilon) := \min\left\{ s \geq 0 \,\middle|\, \left(1+t^{2n}\right)e^{-2\mu t} \leq \frac{\varepsilon}{c(1+\varepsilon)}, \quad \forall t \geq s \right\}, \tag{1.4.11}$$

with the constant c given by Lemma 1.4.7, we see from Lemma 1.4.7 that for any $t \geq \tilde{t}_1(\varepsilon)$

$$\|\boldsymbol{W}(t) - \boldsymbol{I}\| \leq \frac{\varepsilon}{1+\varepsilon}.$$

Combining the above with (1.4.10), shows the first result for $t_1 = \tilde{t}_1(\varepsilon)$.

To prove the second claim we will show that

$$t_1(\varepsilon) \geq \tilde{t}_1(\varepsilon).$$

For this elementary proof we use the fact that

$$\max_{t\geq 0} e^{-at} t^b = \left(\frac{b}{ae}\right)^b$$

for any $a, b > 0$. Thus, choosing $a = 2\alpha$, where $0 < \alpha < \mu$ is arbitrary, and $b = 2n$ we have that

$$\left(1+t^{2n}\right)e^{-2\mu t} \leq \left(1+\left(\frac{n}{\alpha e}\right)^{2n}\right)e^{-2(\mu-\alpha)t}, \quad t \geq 0.$$

As a consequence, if

$$\left(1+\left(\frac{n}{\alpha e}\right)^{2n}\right)e^{-2(\mu-\alpha)t} \leq \frac{\varepsilon}{c(1+\varepsilon)}, \quad \forall t \geq s, \tag{1.4.12}$$

then $s \geq \tilde{t}_1(\varepsilon)$ due to (1.4.11). The smallest possible s in (1.4.12) is obtained by solving the corresponding equality for t, and yields (1.4.9), concluding the proof. □

We now have all the tools to prove Theorem 1.4.3

Proof of Theorem 1.4.3. To show (i) we recall Minkowski's integral inequality, which will play an important role in estimating the L^p norms of $f(t)$.

Minkowski's Integral Inequality: *For any non-negative measurable function F on $(X_1 \times X_2, \mu_1 \times \mu_2)$, and any $q \geq 1$ one has that*

$$\begin{aligned}\left(\int_{X_2}\left|\int_{X_1} F(x_1,x_2)\,d\mu_1(x_1)\right|^q d\mu_2(x_2)\right)^{\frac{1}{q}} \\ \leq \int_{X_1}\left(\int_{X_2}|F(x_1,x_2)|^q\,d\mu_2(x_2)\right)^{\frac{1}{q}} d\mu_1(x_1).\end{aligned} \tag{1.4.13}$$

Next, we fix an $\varepsilon_1 = \varepsilon_1(\varepsilon, q) \in (0,1)$, to be chosen later. From Lemma 1.4.7 and 1.4.8 we see that, for $t \geq t_1(\varepsilon_1)$ with

$$t_1(\varepsilon_1) := \frac{1}{2(\mu - \alpha)} \log \left(\frac{c(1+\varepsilon_1)\left(1 + \left(\frac{n}{\alpha e}\right)^{2n}\right)}{\varepsilon_1} \right)$$

for some fixed $0 < \alpha < \mu$, we have that

$$\|\boldsymbol{W}(t) - \boldsymbol{I}\| \leq \frac{\varepsilon_1}{1+\varepsilon_1} < \varepsilon_1, \qquad \|\boldsymbol{W}^{-1}(t) - \boldsymbol{I}\| \leq \varepsilon_1,$$

and hence

$$\boldsymbol{W}(t) > (1-\varepsilon_1)\boldsymbol{I}, \qquad \boldsymbol{W}(t)^{-1} \geq (1-\varepsilon_1)\boldsymbol{I}.$$

As such, for $t \geq t_1(\varepsilon_1)$

$$\left| e^{-\frac{1}{2}\left(x - e^{-\boldsymbol{C}t}y\right)^T \boldsymbol{W}(t)^{-1}\left(x - e^{\boldsymbol{C}t}y\right)} f_0(y) \right|^q \leq e^{-\frac{q}{2}(1-\varepsilon_1)\left|x - e^{-\boldsymbol{C}t}y\right|^2} \left|f_0(y)\right|^q \tag{1.4.14}$$

and

$$\det \boldsymbol{W}(t) \geq (1-\varepsilon_1)^d. \tag{1.4.15}$$

We conclude, using (1.4.13), the exact solution formula (1.4.5), (1.4.14) and (1.4.15) that for $t \geq t_1(\varepsilon_1)$ it holds:

$$\begin{aligned}
&\int_{\mathbb{R}^d} \left|f(t,x)\right|^q f_\infty^{-1}(x)\,dx \\
\leq& \frac{(2\pi)^{\frac{d}{2}}}{(2\pi(1-\varepsilon_1))^{\frac{qd}{2}}} \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} e^{-\frac{q}{2}(1-\varepsilon_1)\left|x - e^{-\boldsymbol{C}t}y\right|^2} \left|f_0(y)\right|^q e^{\frac{|x|^2}{2}}\,dx \right)^{\frac{1}{q}} dy \right)^q \\
=& \frac{(2\pi)^{\frac{d}{2}}}{(2\pi(1-\varepsilon_1))^{\frac{qd}{2}}} \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} e^{-\frac{q}{2}(1-\varepsilon_1)\left|x - e^{-\boldsymbol{C}t}y\right|^2} e^{\frac{|x|^2}{2}}\,dx \right)^{\frac{1}{q}} \left|f_0(y)\right| dy \right)^q .
\end{aligned} \tag{1.4.16}$$

We proceed by choosing $\varepsilon_1 > 0$ such that $q(1-\varepsilon_1) > 1$ (or equivalently $\varepsilon_1 < \frac{q-1}{q}$) and denoting

$$\eta := q(1-\varepsilon_1) - 1 > 0.$$

Shifting the x variable by $\frac{1}{2}e^{-\boldsymbol{C}t}y$ and completing the square, we find that

$$\begin{aligned}
\int_{\mathbb{R}^d} e^{-\frac{q}{2}(1-\varepsilon_1)\left|x - e^{-\boldsymbol{C}t}y\right|^2} e^{\frac{|x|^2}{2}}\,dx &= \int_{\mathbb{R}^d} e^{-\frac{\eta+1}{2}\left|x - \frac{1}{2}e^{-\boldsymbol{C}t}y\right|^2} e^{\frac{\left|x + \frac{1}{2}e^{-\boldsymbol{C}t}y\right|^2}{2}}\,dx \\
= \int_{\mathbb{R}^d} e^{xe^{-\boldsymbol{C}t}y} e^{-\frac{\eta}{2}\left|x - \frac{1}{2}e^{-\boldsymbol{C}t}y\right|^2}\,dx &= \int_{\mathbb{R}^d} e^{-\frac{\eta}{2}\left|x - \frac{1}{2}\left(1+\frac{2}{\eta}\right)e^{-\boldsymbol{C}t}y\right|^2} e^{\left(\frac{1}{2}+\frac{1}{2\eta}\right)\left|e^{-\boldsymbol{C}t}y\right|^2}\,dx \\
&= \left(\frac{2\pi}{\eta}\right)^{\frac{d}{2}} e^{\left(\frac{1}{2}+\frac{1}{2\eta}\right)\left|e^{-\boldsymbol{C}t}y\right|^2}.
\end{aligned} \tag{1.4.17}$$

Using (1.4.8) we can find a uniform geometric constant c_2 such that

$$\|e^{-\boldsymbol{C}t}\|^2 \leq c_2^2 \left(1+t^n\right)^2 e^{-2\mu t} \leq 2c_2^2 \left(1+t^{2n}\right) e^{-2\mu t}.$$

Following the proof of Lemma 1.4.8 we recall that if

$$t \geq \frac{1}{2(\mu-\alpha)} \log\left(\frac{\tilde{c}(1+\varepsilon_2)\left(1+\frac{n}{\alpha e}\right)^{2n}}{\varepsilon_2}\right),$$

where $0<\alpha<\mu$ is arbitrary and for any $\tilde{c}, \varepsilon_2 > 0$, then

$$\left(1+t^{2n}\right) e^{-2\mu t} \leq \frac{\varepsilon_2}{\tilde{c}(1+\varepsilon_2)}.$$

Thus, choosing

$$\tilde{c} = \frac{c_2^2(1+\eta)}{q\eta} = \frac{c_2^2(1-\varepsilon_1)}{q(1-\varepsilon_1)-1} \quad \text{and} \quad \varepsilon_2 = \frac{\varepsilon_1}{1-\varepsilon_1}$$

we get that if

$$t \geq t_2(\varepsilon_1) := \frac{1}{2(\mu-\alpha)} \log\left(\frac{c_2^2(1-\varepsilon_1)\left(1+\frac{n}{\alpha e}\right)^{2n}}{\left(q(1-\varepsilon_1)-1\right)\varepsilon_1}\right),$$

where $0<\alpha<\mu$ is arbitrary and for any $\tilde{c}, \varepsilon_2 > 0$, then

$$\left(\frac{1}{2}+\frac{1}{2\eta}\right)\|e^{-\boldsymbol{C}t}\|^2 \leq \frac{c_2^2(1+\eta)}{q\eta} q \left(1+t^{2n}\right) e^{-2\mu t} \leq q\varepsilon_1.$$

Combining this with our previous computations ((1.4.16) and (1.4.17)), we find that for any $t \geq t_0(\varepsilon_1) := \max(t_1(\varepsilon_1), t_2(\varepsilon_1))$

$$\int_{\mathbb{R}^d} \left|f(t,x)\right|^q f_\infty^{-1}(x)dx \leq \frac{(2\pi)^{d(1-\frac{q}{2})}}{(1-\varepsilon_1)^{\frac{qd}{2}}\eta^{\frac{d}{2}}} \left(\int_{\mathbb{R}^d} e^{\varepsilon_1|y|^2} f_0(y)dy\right)^q.$$

If ε_1 is chosen more restrictively than before, namely $\varepsilon_1 \leq \frac{q-1}{2q}$, then we have

$$\frac{q-1}{2} \leq \eta < q-1 \qquad \text{and} \qquad 1-\varepsilon_1 \geq \frac{q+1}{2q},$$

which implies the first statement of the theorem by choosing $\varepsilon_1 := \min\left(\varepsilon, \frac{q-1}{2q}\right)$.

For the proof of (ii) we note that (1.4.3) is equivalent to

$$\|f(t)\|^2_{L^2(\mathbb{R}^d, f_\infty^{-1})} \leq \left(\frac{8\sqrt{2}}{3\cdot 2^{\frac{1}{p}}}\right)^d \|f_0\|^2_{L^p\left(\mathbb{R}^d, f_\infty^{1-p}\right)}. \tag{1.4.18}$$

With the Hölder inequality we obtain

$$\begin{aligned}\int_{\mathbb{R}^d} e^{\frac{p-1}{4p}|x|^2} f_0(x)dx &\leq \left(\int_{\mathbb{R}^d} e^{-\frac{|x|^2}{4}}dx\right)^{\frac{p-1}{p}} \left(\int_{\mathbb{R}^d} e^{\frac{p-1}{2}|x|^2} f_0^p(x)dx\right)^{\frac{1}{p}} \\ &= 2^{\frac{d}{2}\frac{p-1}{p}} \|f_0\|_{L^p\left(\mathbb{R}^d, f_\infty^{1-p}\right)}.\end{aligned}$$

Hence, $e_p(f_0|f_\infty) < \infty$ implies (1.4.1) with $\varepsilon = \frac{p-1}{4p}$, and (1.4.18) follows from (1.4.2) with $q = 2$ and $\tilde{t}_0(p) = t_0\left(\frac{p-1}{4p}\right)$. □

Remark 1.4.9. *If the condition* (1.4.1) *holds for* $\varepsilon = \frac{1}{2}$ *we can give an explicit upper bound for the "waiting time" in the hypercontractivity estimate* (1.4.2). *For such* ε *we have* $\varepsilon_1 := \min\left(\varepsilon, \frac{q-1}{2q}\right) = \frac{q-1}{2q}$, *and by choosing* $\alpha = \frac{\mu}{2}$ *we can see that* $t_0(\varepsilon_1)$ *from the proof of Theorem 1.4.3 is*

$$\overline{t_0}(q) := \frac{1}{\mu}\log\left(\frac{\max\left(c(3q-1), 2c_2^2\frac{q+1}{q-1}\right)\left(1+\left(\frac{2n}{\mu e}\right)^{2n}\right)}{q-1}\right),$$

where c, c_2 *are geometric constants found in the proof of Lemma 1.4.7.*

With the non-symmetric hypercontractivity result at hand, we can finally complete the proof of our main theorem for $1 < p < 2$.

Proof of Theorem 1.1.4 for $1 < p < 2$. Using Theorem 1.4.3 (ii) we find an explicit $T_0(p)$ such that for any $t \geq T_0(p)$ the solution to the Fokker–Planck equation, $f(t)$, is in $L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$. Proceeding similarly to the previous remark (but now with $q = 2$ and $\varepsilon = \frac{p-1}{4p}$) we have $\varepsilon_1 := \min\left(\frac{p-1}{4p}, \frac{1}{4}\right) = \frac{p-1}{4p}$. This yields the following upper bound for the "waiting time" in the hypercontractivity estimate (1.4.3):

$$T_0(p) := \frac{1}{\mu}\log\left(\frac{\max\left(c(5p-1), 2c_2^2\frac{3p^2+p}{p+1}\right)\left(1+\left(\frac{2n}{\mu e}\right)^{2n}\right)}{p-1}\right).$$

Using Lemma 1.4.2, Theorem 1.1.4 for $p = 2$ (which was already proven in §1.3), and inequality (1.4.3) we conclude that for any $t \geq T_0(p)$

$$\begin{aligned}e_p(f(t)|f_\infty) &\leq 2e_2(f(t)|f_\infty) \leq 2\tilde{c}_2 e_2\left(f(T_0(p))|f_\infty\right)\left(1+\left(t-T_0(p)\right)^{2n}\right)e^{-2\mu(t-T_0(p))} \\ &\leq 2\tilde{c_p} e^{2\mu T_0(p)}\left(p(p-1)e_p(f_0|f_\infty)+1\right)^{\frac{2}{p}}\left(1+t^{2n}\right)e^{-2\mu t}.\end{aligned} \tag{1.4.19}$$

To complete the proof we recall that any admissible relative entropy decreases along the flow of the Fokker–Planck equation (see [2] for instance). Thus, for any $t \leq T_0(p)$ we have that

$$e_p(f(t)|f_\infty) \leq e_p(f_0|f_\infty) \leq e_p(f_0|f_\infty)e^{2\mu T_0(p)}\left(1+t^{2n}\right)e^{-2\mu t}. \tag{1.4.20}$$

The theorem now follows from (1.4.19) and (1.4.20), together with the fact that for a $1<p<2$

$$e_p(f_0|f_\infty) \leq \mathscr{C}_p\left(p(p-1)e_p(f_0|f_\infty)+1\right)^{\frac{2}{p}},$$

where $\mathscr{C}_p := \sup_{x\geq 0}\frac{x}{(p(p-1)x+1)^{\frac{2}{p}}} < \infty$. □

We end this section with a slight generalization of our main theorem:

Theorem 1.4.10. *Let ψ be a generating function for an admissible relative entropy. Assume in addition that there exists $C_\psi > 0$ such that*

$$\psi_p(y) \leq C_\psi \psi(y) \tag{1.4.21}$$

for some $1 < p < 2$ and all $y \in \mathbb{R}^+$. Then, under the same setting of Theorem 1.1.4 (but now with the assumption $e_\psi(f_0|f_\infty) < \infty$) we have that

$$e_\psi(f(t)|f_\infty) \leq c_{p,\psi}\left(e_\psi(f_0|f_\infty)+1\right)^{\frac{2}{p}}\left(1+t^{2n}\right)e^{-2\mu t}, \quad t\geq 0,$$

where $c_{p,\psi} > 0$ is a fixed geometric constant.

Proof. The proof is almost identical to the proof of Theorem 1.1.4. Due to (1.4.21) we know that $e_p(f_0|f_\infty) < \infty$. As such, according to Theorem 1.4.3 (ii) there exists an explicit $T_0(p)$ such that for all $t \geq T_0(p)$ we have that $f(t) \in L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$ and

$$e_2(f(t)|f_\infty) \leq \frac{1}{2}\left(\left(\frac{8\sqrt{2}}{3\cdot 2^{\frac{1}{p}}}\right)^d \left(C_\psi p(p-1)e_\psi(f_0|f_\infty)+1)\right)^{\frac{2}{p}} - 1\right).$$

The above, together with Lemma 1.4.2 gives the appropriate decay estimate on e_ψ for $t \geq T_0(p)$. Since e_ψ decreases along the flow of our equation, we can deal with the interval $t \leq T_0(p)$ like in the previous proof, yielding the desired result. □

In the next, and last, section of this chapter we will mention another natural quantity in the theory of the Fokker–Planck equations - the Fisher information. We will briefly explain how the method we presented here is different to the usual technique one considers when dealing with the entropy. Moreover we describe how to infer from our main theorem an improved rate of convergence to equilibrium - in relative Fisher information.

1.5 Decay of the Fisher Information

The study of convergence to equilibrium for the Fokker–Planck equations via relative entropies has a long history. Unlike the study we presented here, which relies on detailed spectral investigation of the Fokker–Planck operator together with a non-symmetric hypercontractivity result, the common method to approach this problem - even in the degenerate case - is the so called *entropy method*.
The idea behind the entropy method is fairly simple: once an entropy has been chosen and shown to be a Lyapunov functional to the equation, one attempts to find a linear relation between it and the absolute value of its dissipation. In the setting of the our equation, the latter quantity is referred to as *the Fisher information*.
More precisely, it has been shown in [2] that:

Lemma 1.5.1. *Let ψ be a generating function for an admissible relative entropy and let $f(t,x)$ be a solution to the Fokker–Planck equation* (1.1.2) *with initial datum $f_0 \in L^1_+\left(\mathbb{R}^d\right)$. Then, for any $t>0$ we have that*

$$\frac{d}{dt}e_\psi\left(f(t)|f_\infty\right) = \\ -\int_{\mathbb{R}^d}\psi''\left(\frac{f(t,x)}{f_\infty(x)}\right)\nabla\left(\frac{f(t,x)}{f_\infty(x)}\right)^T \boldsymbol{C}_s\nabla\left(\frac{f(t,x)}{f_\infty(x)}\right)f_\infty(x)dx \le 0.$$

Definition 1.5.2. For a given positive semidefinite matrix $\boldsymbol{P}$ the expression

$$I_\psi^{\boldsymbol{P}}(f|f_\infty) := \int_{\mathbb{R}^d}\psi''\left(\frac{f(x)}{f_\infty(x)}\right)\nabla\left(\frac{f(x)}{f_\infty(x)}\right)^T \boldsymbol{P}\nabla\left(\frac{f(x)}{f_\infty(x)}\right)f_\infty(x)dx \ge 0.$$

is called *the relative Fisher Information generated by ψ*.

The entropy method boils down to proving that there exists a constant $\lambda>0$ such that

$$I_\psi^{\boldsymbol{P}}(f|f_\infty) \ge \lambda e_\psi(f|f_\infty). \tag{1.5.1}$$

When $\boldsymbol{D}$ is positive definite, the above (with the choice $\boldsymbol{P} := \boldsymbol{D}$) is a Sobolev inequality (and a log-Sobolev inequality for $\psi = \psi_1$), and a standard way to prove it is by using the Bakry-Émery technique (see [3, 5] for instance). This technique involves differentiating the Fisher information along the flow of the Fokker–Planck equation and finding a closed functional inequality for it. By an appropriate integration in time, one can then obtain (1.5.1).
Problems start arising with the above method when $\boldsymbol{D}$ is not invertible. As can be seen from the expression of $I_\psi^{\boldsymbol{D}}$ - there are some functions that are not identically f_∞ yet yield a zero Fisher information. In recent work of Arnold and Erb ([2]), the authors managed to circumvent this difficulty by defining a new positive definite matrix $\boldsymbol{P_0}$ that is strongly connected to the drift matrix $\boldsymbol{C}$, and for which (1.5.1) is valid as a functional inequality.

They proceeded to successfully use the Bakry-Émery method on $I_\psi^{\boldsymbol{P}_0}$ and conclude from it, and the log-Sobolev inequality, rates of decay for $I_\psi^{\boldsymbol{D}}$ (which is controlled by $I_\psi^{\boldsymbol{P}_0}$) and e_ψ. This is essentially what is behind the exponential decay in Theorem 1.1.3. Moreover, in the defective case (ii), it led to an ε-reduced exponential decay rate.

As we have managed to obtain better convergence rates to equilibrium (in relative entropy) for the case of defective drift matrices $\boldsymbol{C}$, one might ask whether or not the same rates will be valid for the associated Fisher information $I_p^{\boldsymbol{D}} := I_{\psi_p}^{\boldsymbol{D}}$. The answer to that question is *Yes*, and we summarise this in the next theorem:

Theorem 1.5.3. *Consider the Fokker–Planck equation* (1.1.2) *with diffusion and drift matrices* $\boldsymbol{D}$ *and* $\boldsymbol{C}$ *which satisfy Conditions (A)–(C). Let* μ *be defined as in* (1.1.5) *and assume that one, or more, of the eigenvalues of* $\boldsymbol{C}$ *with real part* μ *are defective. Denote by* $n > 0$ *the maximal defect of these eigenvalues. Then, for any* $1 < p \le 2$, *the solution* $f(t)$ *to* (1.1.2) *with initial datum* $f_0 \in L_+^1\left(\mathbb{R}^d\right)$ *that has unit mass and* $I_p^{\boldsymbol{P}_0}(f_0|f_\infty) < \infty$ *satisfies:*

$$I_p^{\boldsymbol{D}}\left(f(t)|f_\infty\right) \le c I_p^{\boldsymbol{P}_0}\left(f(t)|f_\infty\right) \le c_p(f_0)\left(1+t^{2n}\right)e^{-2\mu t}, \quad t \ge 0,$$

where $c_p(f_0)$ *depends on* $I_p^{\boldsymbol{P}_0}(f_0|f_\infty)$.

Proof. We first note that Proposition 4.4 from [2] implies the estimate $e_p\left(f_0|f_\infty\right) \le c I_p^{\boldsymbol{P}_0}(f_0|f_\infty) < \infty$, and hence Theorem 1.1.4 applies. This decay of e_p carries over to $I_p^{\boldsymbol{P}_0}$ due to the following two ingredients: For small t we can use the purely exponential decay of $I_p^{\boldsymbol{P}_0}$ as established in Proposition 4.5 of [2] (with the rate $2(\mu-\varepsilon)$). And for large time we use the (degenerate) parabolic regularisation of the Fokker–Planck equation (1.1.2): As proven in Theorem 4.8 of [2] we have for all $\tau \in (0,1]$ that

$$I_\psi^{\boldsymbol{P}_0}(f(\tau)|f_\infty) \le \frac{c_{k_0}}{\tau^{2\kappa+1}} e_\psi\left(f_0|f_\infty\right),$$

where ψ is the generating function for an admissible relative entropy. And $\kappa > 0$ is the minimal number such that there exists $\tilde{\lambda} > 0$ with

$$\sum_{j=0}^{\kappa} \boldsymbol{C}^j \boldsymbol{D}\left(\boldsymbol{C}^T\right)^j \ge \tilde{\lambda}\boldsymbol{I}.$$

The existence of such κ and $\tilde{\lambda}$ is guaranteed by Condition (C) and equivalent to the rank condition (1.3.6)- cf. Lemma 2.3 in [1]. □

Bibliography

[1] F. Achleitner, A. Arnold, D. Stürzer, *Large-Time Behavior in Non-Symmetric Fokker–Planck Equations.* Rivista di Matematica della Università di Parma **6** (2015), 1–68.

[2] A. Arnold, J. Erb, *Sharp Entropy Decay for Hypocoercive and Non-Symmetric Fokker–Planck Equations with Linear Drift.* Preprint. https://arxiv.org/abs/1409.5425 .

[3] A. Arnold, P. Markowich, G. Toscani, A. Unterreiter, *On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker–Planck type equations,* Communications in Partial Differential Equations **26** (2001), 43–100.

[4] D. Bakry, I. Gentil, M. Ledoux, *Analysis and Geometry of Markov Diffusion Operators,* Springer (2014).

[5] D. Bakry, M. Émery, *Diffusions hypercontractives,* Séminaire de probabiltés de Strasbourg **19** (1985), 177–206.

[6] J. Bao, F.-Y. Wang, C. Yuan, *Hypercontractivity for functional stochastic differential equations*, Stoch. Proc. Appl.125 (2015), 3636–3656.

[7] J. Bao, F.-Y. Wang, C. Yuan, *Hypercontractivity for Functional Stochastic Partial Differential Equations*, Electron. J. Probab. **20** (2015), no. 93, 15 pp.

[8] K.-J. Engel, R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations,* Springer 2000.

[9] M. Fuhrman, *Hypercontractivity properties of nonsymmetric Ornstein-Uhlenbeck semigroups in Hilbert spaces,* Stochastic Anal. Appl. **16** (1998), no. 2, 241-260.

[10] L. Gross, *Logarithmic Sobolev inequalities,* Amer. J. Math. **97** (1975), no. 4, 1061–1083.

[11] B. Helffer and J. Sjöstrand. *From resolvent bounds to semigroup bounds.* Preprint. ArXiv: 1001.4171v1.

[12] L. Hörmander, *Hypoelliptic second order differential equations,* Acta Math. **119** (1969), 147–171.

[13] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis,* Cambridge University Press (1991).

[14] Y. Latuskhin, Y. Valerian, *Stability estimates for semigroups on Banach spaces,* Discrete Contin. Dyn. Syst. **33**, no. 11-12 (2013), 5203–5216.

[15] P. Monmarché, *Generalized* Γ *calculus and application to interacting particles on a graph,* Preprint. https://arxiv.org/abs/1510.05936

[16] E. Nelson, *The free Markov field,* J. Funct. Anal., **12** (1973), 211–227.

[17] M. Ottobre, G.A. Pavliotis, K. Pravda-Starov, *Exponential return to equilibrium for hypoelliptic quadratic systems,* J. Funct. Anal. **262** (2012), 4000–4039.

[18] M. Ottobre, G.A. Pavliotis, K. Pravda-Starov, *Some remarks on degenerate hypoelliptic Ornstein-Uhlenbeck operators,* J. Math. Anal. Appl. **429** (2015), 676–712.

[19] H. Risken, *The Fokker–Planck equation. Methods of solution and applications.*, Springer-Verlag (1989).

[20] J. Snyders, M. Zakai, *On nonnegative solutions of the equation* $AD + DA' = -C$, SIAM J. Appl. Math. **18** (1970), 704–715.

[21] F-Y. Wang. *Hypercontractivity and applications for stochastic Hamiltonian systems,* J. Funct. Anal. **272**, no. 12 (2017), 5360–5383.

[22] F-Y. Wang. *Functional Inequalities, Markov Semigroups and Spectral Theory,* Science-Press (2005).

2 Sharp Decay Estimates in Local Sensitivity Analysis for Evolution Equations with Uncertainties: from ODEs to Linear Kinetic Equations

2.1 Introduction

Kinetic models arise from mesoscopic approximations of particle systems, as such, they are not first principle equations, thus contain empirical coefficients such as collision kernels in the Boltzmann equation, scattering coefficients in transport equations, forcing or source terms, and measurement errors in initial and boundary data, etc. Such errors can be modeled by uncertainties, or random inputs. Quantifying these uncertainties have important industrial and practical applications, in order to identify the sensitivities of input parameters, validate the models, conduct risk management, and ultimately improve these models.

In recent years one has seen activities in conducting uncertainty quantifications (UQ) for kinetic equations, see [14] for a recent review. One of the important analysis in UQ is the so-called local sensitivity analysis, in which one aims to understand how sensitive the solution depends on the input parameters [26]. For kinetic equations, a major tool to conduct sensitivity analysis for random kinetic equations has been the coercivity, or more generally, hypocoercivity, which originated in the study of long-time behavior of kinetic equations (see [28, 11, 9, 23]). In such analysis, by using the hypocoercivity of the kinetic operator, in a perturbative setting, namely, considering solutions near the global equilibrium (see [13]), one can establish the long time convergence toward the local equilibrium with an exponential time decay rate. Such analysis has been extended to kinetic equations with random inputs, in both linear (see [15, 20]) and nonlinear (see [17, 21, 25]) settings. For stochastic Galerkin methods, hypocoercivity analysis even leads to exponential decay of numerical errors [21, 25], while in classical numerical analysis one often obtains errors that grow exponentially in time. In these works, however, the decay rates were not sharp.

Over the last two decades, entropy methods have become an important and robust tool to prove exponential convergence to equilibrium in kinetic and parabolic equations (see [27, 6, 10, 19, 11]). But sharpness of the decay rate is only known in few cases (see [6] for the situation in Fokker–Planck equations). For linear finite dimensional ODEs, however, a method of constructing Lyapunov functionals to reveal optimal decay rates has been known for a long time, see [7], §22.4.

More recently in [5], such strategies were transferred to Fokker–Planck (FP) equations on $\mathbb{R}^d$ and used to estimate the decay behavior of their solutions. For both, the ODE and the FP setting, one obtains the sharp *exponential* decay rate, as long as none of the eigenvalues determining the spectral gap of the generator is defective[1]. Recently this method was applied to PDEs that allow for a modal decomposition, like kinetic BGK models on the torus (see [1, 2]). They are relaxation-type models for collisional gases, introduced by the physicists Bhatnagar, Gross and Krook in [8].

In the defective case, however, the sharp decay behavior is of the form of a *polynomial times an exponential*, and different strategies have to be applied.

In order to catch the sharp decay behavior in the case of a defective FP equation, one can use the spectral properties of the FP operator to split the solution into two subspace-invariant parts: The first one corresponds to the spectral gap and is finite dimensional; there the sharp (defective) decay behavior can be computed explicitly. The second part of the solutions corresponds to a subspace “away” from the spectral gap, and it has a faster exponential decay. This approach gives sharp decay functions for defective FP equations for various entropies as shown in [4].

Alternatively, one can extend the Lyapunov functional by allowing it to be time dependent, see [22]. In §2.2 of this chapter we will translate that strategy from linear Fokker–Planck equations in $\mathbb{R}^d$ (as in Corollary 12 of [22]) to the ODE setting. We shall also refine the method such that it can yield uniform decay bounds in the non-defective limit. This extension will be crucial for our PDE-applications presented in §3–§5: a convection-diffusion equation, a BGK model and a linear Fokker–Planck equation on $\mathbb{R}$ respectively. There we shall allow for uncertainty in the model coefficients and carry out a first (and for the convection-diffusion equations also second) order sensitivity analysis. In a local sensitivity analysis one tries to estimate the behavior of the (higher order) derivatives of the solution with respect to the input variables [26]. Estimates of such derivatives are not only important to assess the sensitivity of the solution on the input parameters, they also provide regularity of the solution in the parameter space which is important to determine the convergence order of numerical approximations in the random space [12, 15, 16]. In the Fourier space, the resulting evolution equations for the parametric derivatives are mostly defective systems for which the sharp decay estimates can be obtained by using the Lyapunov functional approaches for defective deterministic systems. We would also like to point out that for the case of linear Fokker–Planck equation with a random drift, studied in §5, the global equilibrium is also random, while in previous sensitivity analysis for uncertain kinetic equations the global equilibria were all

[1] An eigenvalue is *defective*, if its algebraic multiplicity is strictly greater than its geometric multiplicity.

deterministic [17, 21].

2.2 Lyapunov Functionals for Defective ODEs

In this section we first review (from [1]) the Lyapunov functional method for non-defective ODEs and then extend it to the defective case. This is based on constructing a norm *adapted to the problem* that allows to recover the sharp decay behavior.

2.2.1 Construction of Lyapunov Functionals

Let the matrix[2] $C \in \mathbb{C}^{d\times d}$ be positive stable, i.e. its eigenvalues satisfy $\mathrm{Re}(\lambda_i) > 0$ for $i = 1,\dots,d$, and let $\mu := \min_{i=1,\dots,d} \mathrm{Re}(\lambda_i) > 0$. We want to find a Lyapunov functional for the equation

$$\frac{d}{dt}x(t) = -Cx(t), \quad x \in \mathbb{C}^d, t \geq 0 \tag{2.2.1}$$

that allows to deduce the sharp decay rate of solutions with energy-type estimates. For the construction of this functionals we consider the Jordan transformation of the matrix C^H, denoting the Hermitian transpose of the matrix C (with eigenvalues $\overline{\lambda_i}$). We shall distinguish different cases of eigenvalue defectiveness:

$$C^H = V\,\mathrm{diag}(J_1,\dots,J_N)V^{-1}, \tag{2.2.2}$$

where J_n for $n \in \{1,\dots,N\}$ are the Jordan blocks of C^H with length $l_n \in \{1,\dots,d\}$. A Jordan block of length one is an eigenvalue as a diagonal element, and a Jordan block J_n of length $l_n > 1$ corresponds to a chain of generalized eigenvectors of C^H of order k satisfying

$$C^H v_n^{(k)} = \overline{\lambda}_n v_n^{(k)} + v_n^{(k-1)}, \quad k \in \{1,\dots,l_n - 1\}, \tag{2.2.3}$$

where $v_n^{(0)}$ is an eigenvector of C^H, corresponding to $\overline{\lambda}_n$. We denote the (semi-) norm

$$|x|_P^2 := x^H P x,$$

for a Hermitian positive (semi-)definite matrix $P \in \mathbb{C}^{d\times d}$ to be defined.

Case 1: J_n is a Jordan block of length $l_n = 1$ with $\mathrm{Re}(\lambda_n) \geq \mu$.

We define the rank 1 matrix[3] $P_n := v_n^{(0)} \otimes v_n^{(0)}$ and get (cf. (2.51) in [5])

$$\frac{d}{dt}|x(t)|_{P_n}^2 = -x^H(C^H P_n + P_n C)x \leq -2\mu x^H P_n x = -2\mu |x(t)|_{P_n}^2. \tag{2.2.4}$$

[2] Due to the large amount of matrices appearing, we will refrain from denoting them bold in this chapter.

[3] For $v, w \in \mathbb{C}^d$ we denote $v \otimes w := v \cdot w^H$ where $\cdot$ is the matrix-matrix multiplication.

Case 2: J_n is a Jordan block of length $l_n > 1$ with $\mathrm{Re}(\lambda_n) > \mu$.

As in the proof of Lemma 4.3 in [5], we choose the coefficients $b_n^i > 0$ as

$$b_n^1 := 1; \qquad\qquad b_n^j := c_j(\tau_n)^{2(1-j)}, \quad j \in \{2, \dots, l_n\},$$

where $c_1 := 1$, $c_j := 1 + (c_{j-1})^2$ for $j \in \{2, \dots, l_n\}$ and $\tau_n := 2(\mathrm{Re}(\lambda_n) - \mu) > 0$. Then the matrix

$$P_n := \sum_{i=1}^{l_n} b_n^i v_n^{(i-1)} \otimes v_n^{(i-1)}$$

satisfies

$$C^H P_n + P_n C \geq 2\mu P_n \tag{2.2.5}$$

and, as in (2.2.4), one gets

$$\frac{d}{dt} |x(t)|_{P_n}^2 \leq -2\mu |x(t)|_{P_n}^2. \tag{2.2.6}$$

Case 3: J_n is a Jordan block of length $l_n > 1$ with $\mathrm{Re}(\lambda_n) = \mu$.

A translation of the strategy of Corollary 12 in [22] to the ODE setting leads to the following construction. For each $m \in \{1, \dots, l_n\}$, define the vector function

$$w_n^m(t) := \sum_{k=1}^{m} \frac{t^{m-k}}{(m-k)!} v_n^{(k-1)}, \quad t \geq 0. \tag{2.2.7}$$

For $m \in \{2, \dots, l_n\}$ we have

$$\frac{d}{dt} w_n^m(t) = \sum_{k=1}^{m-1} \frac{t^{m-k-1}}{(m-k-1)!} v_n^{(k-1)},$$

from which it follows (using (2.2.3)) that

$$C^H w_n^m(t) - \overline{\lambda_n} w_n^m(t) = \sum_{k=2}^{m} \frac{t^{m-k}}{(m-k)!} v_n^{(k-2)} = \frac{d}{dt} w_n^m(t). \tag{2.2.8}$$

Next we define the time dependent Hermitian positive semi-definite matrix

$$P_n^m(t) := w_n^m(t) \otimes w_n^m(t), \quad m \in \{1, \dots, l_n\}. \tag{2.2.9}$$

Notice that $w_n^1(t) = v_n^{(0)}$. So for $m = 1$ the computation for the estimate of $\frac{d}{dt}|x(t)|^2_{P_n^1(t)}$ is the same as in (2.2.4) (with equality since $\mathrm{Re}(\lambda_n) = \mu$). For $m \in \{2, \dots, l_n\}$ we use the identity (2.2.8), and compute

$$\begin{aligned}\frac{d}{dt}|x(t)|^2_{P_n^m(t)} &= \dot{x}^H(t)P_n^m(t)x(t) + x^H(t)P_n^m(t)\dot{x}(t) + x^H(t)\dot{P}_n^m(t)x(t)\\ &= -x^H(t)[C^H w_n^m(t) \otimes w_n^m(t) + w_n^m(t) \otimes w_n^m(t)C]x(t)\\ &\quad + x^H(t)\left[(C^H w_n^m(t) - \overline{\lambda_n} w_n^m(t)) \otimes w_n^m(t)\right]x(t)\\ &\quad + x^H(t)\left[w_n^m(t) \otimes (C^H w_n^m(t) - \overline{\lambda_n} w_n^m(t))\right]x(t)\\ &= -2\mu x^H(t)w_n^m(t) \otimes w_n^m(t)x(t) = -2\mu|x(t)|^2_{P_n^m(t)}\end{aligned}$$

and directly obtain

$$|x(t)|^2_{P_n^m(t)} = e^{-2\mu t}|x(0)|^2_{P_n^m(0)}, \quad t \geq 0. \tag{2.2.10}$$

For arbitrary $\beta_n^m > 0$, define

$$P_n(t) := \sum_{m=1}^{l_n} \beta_n^m P_n^m(t). \tag{2.2.11}$$

We have $\mathrm{span}\{v_n^{(0)}, \dots, v_n^{(l_n-1)}\} = \mathrm{span}\{w_n^1(t), \dots, w_n^{l_n}(t)\}$ for all $t \geq 0$, since the transformation matrix between these two sets is given by $e^{-\overline{\lambda_n}t}e^{J_n t}$. Hence, the matrix $P_n(t)$ is positive definite on the subspace $\mathrm{span}\{v_n^{(0)}, \dots, v_n^{(l_n-1)}\}$. In the corresponding seminorm $|\cdot|^2_{P_n(t)}$ the solution $x(t)$ satisfies

$$|x(t)|^2_{P_n(t)} = e^{-2\mu t}|x(0)|^2_{P_n(0)}, \quad t \geq 0.$$

For later convenience, we denote

$$I_\mu := \{n \in \{1, \dots, N\} \mid l_n > 1, \mathrm{Re}(\lambda_n) = \mu\}, \tag{2.2.12}$$

to collect all indices with non-trivial Jordan blocks corresponding to μ, i.e. corresponding to the above Case 3.

Combining the three cases:

Now let us define

$$P(t) := \sum_{n \notin I_\mu} \beta_n P_n + \sum_{n \in I_\mu} P_n(t) = \sum_{n \notin I_\mu} \beta_n P_n + \sum_{n \in I_\mu} \sum_{m=1}^{l_n} \beta_n^m P_n^m(t), \tag{2.2.13}$$

where P_n and $P_n^m(t)$ are chosen, depending on the above Cases 1–3 of the corresponding Jordan block J_n. The weights $\beta_n > 0$ are arbitrary in Cases 1–2, and the (arbitrary) $\beta_n^m > 0$

pertain to Case 3. The matrix $P(t)$ is positive definite for every $t \geq 0$, since it has full rank by construction and it is the sum of positive semi-definite matrices. It satisfies

$$\frac{d}{dt}|x(t)|^2_{P(t)} \leq -2\mu |x(t)|^2_{P(t)}, \tag{2.2.14}$$

and by applying Gronwall's lemma, we conclude:

Lemma 2.2.1. *Let $C \in \mathbb{C}^{d\times d}$ be positive stable and let $\mu > 0$ be the smallest real part of all eigenvalues. Let $\operatorname{diag}(J_1, \dots, J_N)$ be the Jordan normal form of C^H, where J_n for $n \in \{1, \dots, N\}$ is a Jordan block of length l_n with eigenvalue $\overline{\lambda_n}$.*

1. *If all eigenvalues with real part equal to μ are non-defective, i.e. $I_\mu = \emptyset$, then there exists a time-independent Hermitian positive definite matrix $P \in \mathbb{C}^{d\times d}$, such that the solutions to* (2.2.1) *satisfy*

$$|x(t)|^2_P \leq e^{-2\mu t} |x(0)|^2_P. \tag{2.2.15}$$

2. *If at least one eigenvalue with real part equal to μ is defective, i.e. $I_\mu \neq \emptyset$, then there exists a time-dependent matrix $P(t) \in \mathbb{C}^{d\times d}$, which is Hermitian positive definite for all $t \geq 0$ such that the solutions to* (2.2.1) *satisfy*

$$|x(t)|^2_{P(t)} \leq e^{-2\mu t} |x(0)|^2_{P(0)}. \tag{2.2.16}$$

For further details on the algebraic interpretation of the time-independent matrix P, we refer to the remarks following Lemma 4.3 in [5]. See Example 2.2 in [3] (with $\omega \neq 0$) for an ODE example and the relevance of the modified (time-independent) P-norm for the trajectories of the ODE.

Remark 2.2.2. *The matrix $P(t)$ is — with the construction described above — not unique. For one, arbitrary coefficients $\beta_n, \beta_n^m > 0$ in the definition of $P(t)$ in* (2.2.13) *are admissible. Secondly, the construction depends on the specific choice of (generalized) eigenvectors fixed in* (2.2.3).

Remark 2.2.3. *The matrix $P(t)$ can also be written as a matrix product:*

$$P(t) = V e^{Jt} \Sigma(t) B (V e^{Jt})^H,$$

with V and J from (2.2.2),

$$B := \operatorname{diag}(\underbrace{\beta_1^1, \dots, \beta_1^{l_1}}_{l_1 \text{ entries}}, \dots, \underbrace{\beta_N^1, \dots, \beta_N^{l_N}}_{l_N \text{ entries}}) \in \mathbb{R}^{d\times d},$$

with notation $\beta_n^m := \beta_n$ for each $n \notin I_\mu$ and corresponding $m \in \{1, \dots, l_n\}$ and

$$\Sigma(t) := \operatorname{diag}(\underbrace{e^{-2\operatorname{Re}(\lambda_1)t}, \dots, e^{-2\operatorname{Re}(\lambda_1)t}}_{l_1 \text{ times}}, \dots, \underbrace{e^{-2\operatorname{Re}(\lambda_N)t}, \dots, e^{-2\operatorname{Re}(\lambda_N)t}}_{l_N \text{ times}}) \in \mathbb{C}^{d\times d}.$$

This representation of $P(t)$ directly implies that $\det P(t) \equiv \det P(0)$ *(cf. Fig. 2.2.3).*

In the following remark and in Example 2.2.5, we investigate the geometry of the modified norms of Lemma 2.2.1.

Remark 2.2.4. *For an ODE* (2.2.1)*, we distinguish different eigenvalue settings of the matrix $C \in \mathbb{C}^{d\times d}$, in order to isolate the interesting phenomena.*

- Case 1: C is in Case 1 of Lemma 2.2.1: *Due to Lemma 2.2.1, there exists a time-independent P-norm such that solutions decay as* (2.2.15)*. The geometric reason for the strict decay is the following: This specific norm is modified such that the trajectories of solutions to the ODE are never tangential to the P-norm level curves* $\{x \in \mathbb{C}^d \mid |x|_P^2 = \text{const.}\}$ *(cf. Fig. 2.2.2).*

  To prove this, denote $f(x) := x^H P x$. Then, the normal vector of the P-norm level curve at point $x \in \mathbb{C}^d \setminus \{0\}$ is given as the P-norm gradient of $f(x)$, i.e. $\eta(x) := \nabla_P f(x) = P^{-1}\nabla f(x) = 2x$ (see, e.g. (2.1.13) in [18] for gradients of Riemannian manifolds). The (backwards-in-time facing) solution tangent vector at point x is given as $-\dot{x} = Cx$. Due to the matrix inequality (2.2.5)*, $\eta(x)$ and $-\dot{x}$ are never perpendicular, i.e. the P-norm angle between them is bounded from below:*

$$\begin{aligned}\frac{\langle x, Cx\rangle_P}{|x|_P|Cx|_P} &= \frac{x^H PCx}{|x|_P|Cx|_P} = \frac{1}{2}\frac{x^H(C^TP+PC)x}{|x|_P|Cx|_P}\\ &\geq \mu\frac{x^H Px}{|x|_P|Cx|_P} = \mu\frac{|x|_P}{|Cx|_P} \geq \frac{\mu}{|C|_P} > 0,\end{aligned}$$

  where $|C|_P$ denotes the matrix norm induced by the vector norm $|x|_P$.

- Case 2: C is in Case 2 of Lemma 2.2.1*: First, for the time-independent matrix P_ε, as defined in [5], Lemma 4.3, the analogous result as for the above case is true. The calculation is identical, up to replacing μ by $\mu-\varepsilon$ (cf. Example 2.2.5). Thus, also in the defective case, the* solutions are never tangential to the level curves of the P_ε-norm.

  The $P(t)$-norm of Case 2 in Lemma 2.2.1 has a different geometric effect on solutions due to its time-dependency. In fact, a solution $x(t)$ can even be tangential to the $P(t)$-norm level curves for all times, while still maintaining sharp exponential decay, as Example 2.2.5 below shows.

- Case 3: All eigenvalues of C have real part μ (defective or non-defective)*: For $t \geq 0$,* the $P(t)$-angle between $\eta(x(t))$ and $-\dot{x}(t)$ stays constant, *i.e.*

$$\frac{\langle x(t), Cx(t)\rangle_{P(t)}}{|x(t)|_{P(t)}|Cx(t)|_{P(t)}} = \frac{\langle x(0), Cx(0)\rangle_{P(0)}}{|x(0)|_{P(0)}|Cx(0)|_{P(0)}}. \qquad (2.2.17)$$

  Indeed, in this case, (2.2.14) *becomes an equality (with $P(t) \equiv P(0)$, if all eigenvalues are non-defective), and hence $|x(t)|_{P(t)}^2 = e^{-2\mu t}|x(0)|_{P(0)}^2$ for all $t \geq 0$. In the* *following computation we use the polarization identity in the first and last step. The second identity follows from the fact that, if $x(t)$ is a solution to* (2.2.1)*, so is $(I+C)x(t)$ and $(I+iC)x(t)$ (with initial conditions $(I+C)x(0)$ and $(I+iC)x(0)$, respectively):*

$$\begin{aligned}
\frac{\langle x(t), Cx(t)\rangle_{P(t)}}{|x(t)|_{P(t)}|Cx(t)|_{P(t)}} &= \frac{1}{4}\frac{|x(t)+Cx(t)|^2_{P(t)} - |x(t)-Cx(t)|^2_{P(t)}}{e^{-\mu t}|x(0)|_{P(0)}e^{-\mu t}|Cx(0)|_{P(0)}} \\
&\quad + \frac{1}{4}\frac{i|x(t)-iCx(t)|^2_{P(t)} - i|x(t)+iCx(t)|^2_{P(t)}}{e^{-\mu t}|x(0)|_{P(0)}e^{-\mu t}|Cx(0)|_{P(0)}} \\
&= \frac{1}{4}\frac{e^{-2\mu t}|x(0)+Cx(0)|^2_{P(0)} - e^{-2\mu t}|x(0)-Cx(0)|^2_{P(0)}}{e^{-\mu t}|x(0)|_{P(0)}e^{-\mu t}|Cx(0)|_{P(0)}} \\
&\quad + \frac{1}{4}\frac{e^{-2\mu t}i|x(0)-iCx(0)|^2_{P(0)} - e^{-2\mu t}i|x(0)+iCx(0)|^2_{P(0)}}{e^{-\mu t}|x(0)|_{P(0)}e^{-\mu t}|Cx(0)|_{P(0)}} \\
&= \frac{\langle x(0), Cx(0)\rangle_{P(0)}}{|x(0)|_{P(0)}|Cx(0)|_{P(0)}}
\end{aligned}$$

for all $t \geq 0$.

Example 2.2.5. Consider the IVP $\dot{x} = -Cx$, $x(0) = (6,6)^T$, with matrix

$$C = \begin{pmatrix} 1 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix},$$

which has the defective eigenvalue and spectral gap $\lambda = \mu = \frac{1}{2}$ and the (generalized) eigenvectors

$$w^{(0)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1, & -1 \end{pmatrix}^T, \qquad w^{(1)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1, & 1 \end{pmatrix}^T.$$

Our goal is to obtain a better geometric understanding of the necessity of a time-dependent norm for sharp decay estimates of solutions. To this end, we compare the here presented Lyapunov functional constructions with functionals considered in [5].

The naive approach of using the Euclidean norm of the solution $x(t)$ exhibits non strict decay (the dashed curve in Fig. 2.2.1 has a horizontal tangent at $t = 1$). The time-independent norm $|\cdot|_{P_\varepsilon}$, as defined in [5], Lemma 4.3, with the matrix

$$P_\varepsilon = \frac{1}{\sqrt{2}}\begin{pmatrix} \frac{1}{2\varepsilon^2}+1 & \frac{1}{2\varepsilon^2}-1 \\ \frac{1}{2\varepsilon^2}-1 & \frac{1}{2\varepsilon^2}+1 \end{pmatrix}, \quad \varepsilon > 0,$$

yields uniform exponential decay, but the rate $\mu - \varepsilon$ is not sharp. See Fig. 2.2.1 for the decay plot and Fig. 2.2.2 for a geometric reasoning why a modified norm can yield exponential decay.

The $P(t)$-norm, with matrix

$$P(t) = \frac{1}{2}\begin{pmatrix} t^2+2t+2 & t^2 \\ t^2 & t^2-2t+2 \end{pmatrix},$$

as defined in (2.2.13) (with weights $\beta^1 = \beta^2 = 1$), provides the sharp exponential decay

$$|x(t)|^2_{P(t)} = e^{-t}|x(0)|^2_{P(0)} = e^{-t}|x(0)|^2_2, \qquad t \geq 0.$$

See Figures 2.2.3–2.2.4 for the geometric evolution of the $P(t)$-norm.

Choosing the initial condition $\tilde{x}(0) = (0,7)^T$, yields the solution $\tilde{x}(t)$ that is tangential to the $P(t)$-norm level curve for each $t \geq 0$, while the exponential decay of the solution in $P(t)$-norm is still sharp, see Figure 2.2.5.

◊

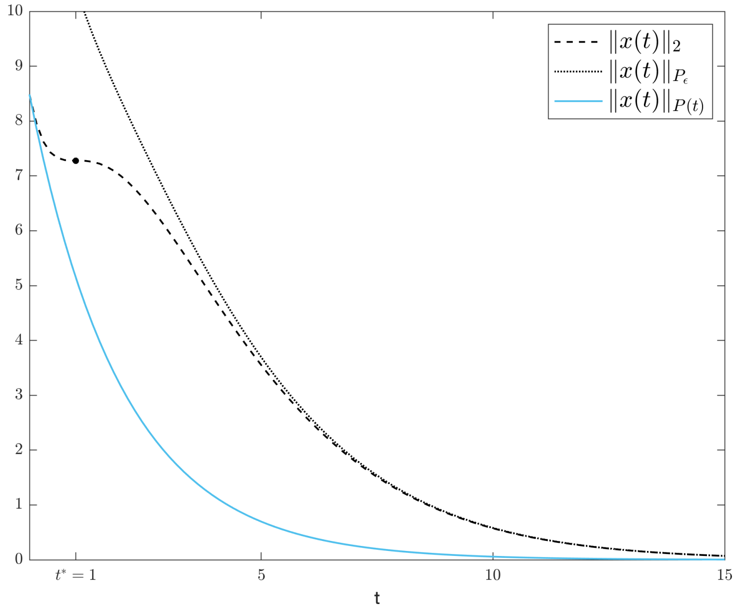


Figure 2.2.1: The dashed line shows the decay of the solution in the Euclidean norm. It initially exhibits a wavy behavior, where at time $t^* = 1$, there is no strict decay at all. The dotted line describes the solution in P_ε-norm with $\varepsilon = 0.4$. It yields uniform exponential decay, however, the decay rate is not sharp. The solid line shows the decay of $|x(t)|_{P(t)}$ with sharp exponential rate $e^{-\frac{1}{2}t}$.

Lemma 2.2.1 shows that the $P(t)$-norm of any solution to (2.2.1) decays exponentially. But due to the time dependence of the norm itself, it is not evident that this is an appropriate functional to capture the sharp decay rate of solutions. Hence, we shall next compare the $P(t)$-norm to the Euclidean norm.

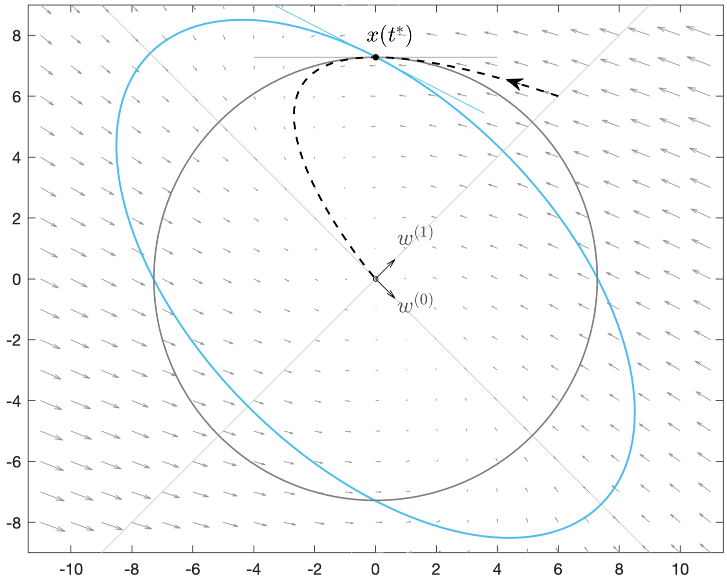


Figure 2.2.2: The dashed line shows the solution trajectory $x(t)$. At the marked point $x(t^*)$, the solution is tangential to the Euclidean level curve. This implies non-strict decay in the Euclidean norm (cf. Fig. 2.2.1) at $t^* = 1$. The ellipse represents a level curve of the P_ε-norm (with $\varepsilon = 0.4$). It modifies the geometry such that the solution is never tangential to the level curves of $|\cdot|_{P_\varepsilon}$. This assures strict exponential decay in the P_ε-norm, however the rate is not sharp.

An arbitrary Hermitian positive definite matrix $P \in \mathbb{C}^{n\times n}$ satisfies

$$\lambda^P_{\min} I \le P \le \lambda^P_{\max} I, \tag{2.2.18}$$

where $\lambda^P_{\min}$ is the smallest and $\lambda^P_{\max}$ is the largest eigenvalue of P. Using this inequality for $P(t)$, and the decay estimate (2.2.16), leads to the Euclidean decay estimate

$$\begin{aligned}|x(t)|_2^2 &\le (\lambda^{P(t)}_{\min})^{-1}|x(t)|^2_{P(t)} \le (\lambda^{P(t)}_{\min})^{-1} e^{-2\mu t}|x(0)|^2_{P(0)}\\ &\le (\lambda^{P(t)}_{\min})^{-1}\lambda^{P(0)}_{\max} e^{-2\mu t}|x(0)|_2^2.\end{aligned}$$

But here the decay behavior is “hidden” in the smallest eigenvalue of $P(t)$. The true qualitative behavior of $|x(t)|_2^2$ will be derived next.

The following technical Lemma 2.2.6 allows us, in the subsequent step, to estimate the time-dependency of the $P^m_n(t)$-semi-norm. The strategy of the proof has already been used for Corollary 12 of [22] in the setting of Fokker–Planck equations on $\mathbb{R}^d$. However, our refined estimate here, for one, provides an upper bound for the time-dependence

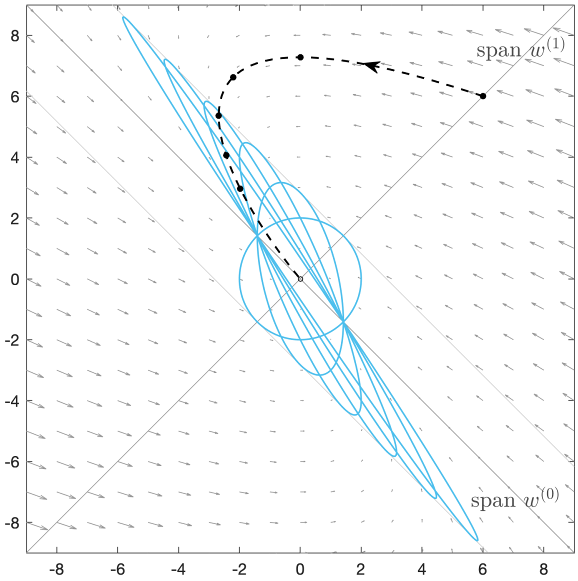


Figure 2.2.3: The dashed line in the plot describes the solution $x(t)$ with the marked points $x(0), x(1), \dots, x(5)$. Additionally, the level curves of $\{x \in \mathbb{R}^2 \mid |x|^2_{P(t)} = 4\}$ for $t = 0, 1, \dots, 5$ are plotted. In direction of the eigenvector of the matrix C, $w^{(0)}$, the distances stay constant in time. The area spanned by each ellipse-shaped level curve stays constant (cf. Remark 2.2.3), while the semi-major axis of the ellipse stretches out and tilts towards the eigenvector axis $w^{(0)}$ as time increases. The stretch is linear in t in the direction of $\pm w^{(0)}$: For arbitrary $t \geq 0$, the point $\pm\sqrt{2}(1-t, 1+t)^T$ is on the ellipse with the tangent $(\pm\sqrt{2}, \pm\sqrt{2})^T + \text{span}\{w^{(0)}\}$.

of a more general class of modified norms. And additionally, the constants appearing in the estimate are explicit, depending on m and a parameter θ which, later on, allows to optimize the constants in the decay estimate of solutions to the ODE (2.2.1).

Lemma 2.2.6. *For linearly independent vectors $v^1, \dots, v^m \in \mathbb{C}^d$ define*

$$\hat{w}^m(t) := \xi^m v^m + \sum_{j=1}^{m-1} \xi^j(t) v^j, \tag{2.2.19}$$

where $\xi^j(t)$ for $j \in \{1, \dots, m-1\}$ are (arbitrary) real-valued polynomials in $t \geq 0$, and $\xi^m > 0$. Furthermore let

$$\hat{P}^m(t) := \hat{w}^m(t) \otimes \hat{w}^m(t), \quad t \geq 0; \qquad Q^j := v^j \otimes v^j, \quad j \in \{1, \dots, m\}.$$

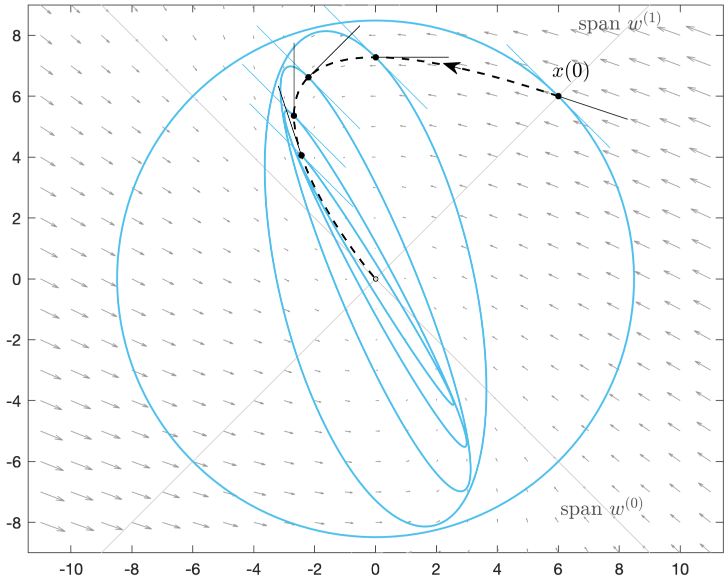


Figure 2.2.4: The level curves of $\{x \in \mathbb{R}^2 \mid |x|^2_{P(t)} = e^{-t}|x(0)|^2_2\}$ for $t = 0, 1, \ldots, 4$ are plotted (cf. Fig. 2.2.3). They intersect with the solution trajectories exactly at the marked points $x(0), x(1), \ldots, x(4)$, which corresponds to the statement of Lemma 2.2.1, Case 2. Notice that the tangents of the level curves of $|\cdot|_{P(t)}$ at $x(t)$ are all parallel to each other. The intersection angle in the $P(t)$-norm is time-independent, see Remark 2.2.4, Case 3.

Then, the following inequality holds for every $x \in \mathbb{C}^d$, $\theta \in (0,1)$ *and* $t \geq 0$ *:*

$$|x|^2_{\hat{P}^m(t)} \geq (1-\theta)(\xi^m)^2 |x|^2_{Q^m} - \left(\frac{(m-1)^2}{\theta} - 1\right) \sum_{k=1}^{m-1} (\xi^k(t))^2 |x|^2_{Q^k}.$$

Notice that the second (t-dependent) term of the r.h.s. involves only the semi-norms $|x|_{Q^1}, \ldots, |x|_{Q^{m-1}}$.

The technical proof is deferred to Appendix 2.A.

Remark 2.2.7. *Lemma 2.2.6 is formulated in a general form, to be applicable also to §2.4 below. Here, we use it to estimate the time-dependency of the* $P^m_n(t)$*-semi-norms (as defined in* (2.2.9)*) for arbitrary* $n \in I_\mu$ *and corresponding* $m \in \{2, \ldots, l_n\}$*. In notation of Lemma 2.2.6, choose* $v^1 = v_n^{(0)}, \ldots, v^m = v_n^{(m-1)}$ *and* $\xi^k(t) = \frac{t^{m-k}}{(m-k)!}$ *for* $k \in \{1, \ldots, m\}$*, which*

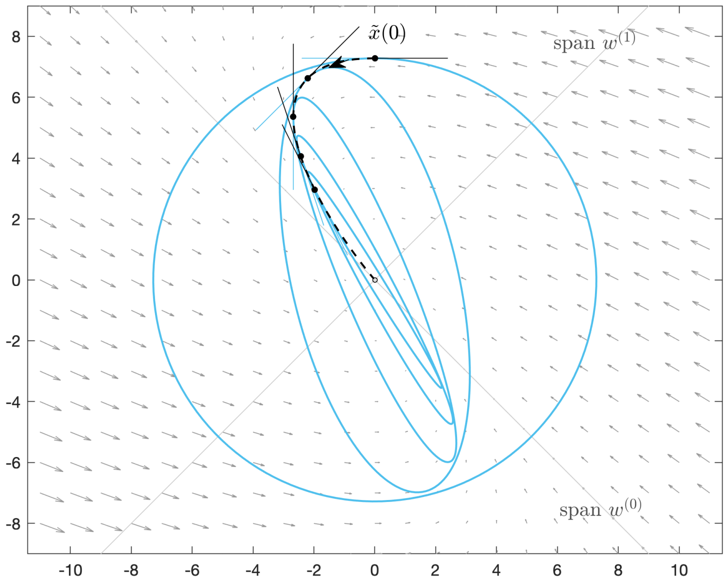


Figure 2.2.5: The level curves of $\{x \in \mathbb{R}^2 \mid |x|^2_{P(t)} = e^{-t}|\tilde{x}(0)|_2^2\}$ for $t = 0, 1, \dots, 4$ are plotted analogous to Fig. 2.2.4, but here for the initial value $\tilde{x}(0) = (0,7)^T$. The solution $\tilde{x}(t)$ is tangential to the $P(t)$-norm level curves for each $t \geq 0$.

leads to $\hat{w}_n^m(t) = w_n^m(t)$. *Then, Lemma 2.2.6 yields*

$$|x|^2_{P_n^m(t)} \geq \\ (1-\theta)\,|x|^2_{P_n^m(0)} - \left(\frac{(m-1)^2}{\theta} - 1\right) \sum_{k=1}^{m-1} \left(\frac{t^{m-k}}{(m-k)!}\right)^2 |x|^2_{P_n^k(0)}, \tag{2.2.20}$$

for every $n \in I_\mu$, *all corresponding* $m \in \{2, \dots, l_n\}$, $x \in \mathbb{C}^d$, $\theta \in (0,1)$ *and* $t \geq 0$.

In the next step, we combine the exponential decay in $P(t)$-norm of solutions, (2.2.16) with the lower bound (2.2.20). This allows to estimate the $P(0)$-norm decay of solutions and, consequently, the decay behavior in the Euclidean norm. In contrast to the result stated in [22] for the FP setting, we obtain an multiplicative constant $\mathscr{C}$ in the estimate depending explicitly on the maximal defect associated to μ and the choice of the weights β_n^m of $P(t)$.

The freedom of choice in the weights and their influence on the constant $\mathscr{C}$ is of great importance in §2.3–§2.5, as $\mathscr{C}$ will need to stay bounded in the *non-defective limit* (see Example 2.2.11 below).

Theorem 2.2.8. *Let* $C \in \mathbb{C}^{d\times d}$ *be positive stable and let* $\mu > 0$ *be the smallest real part of all eigenvalues. Let M be the maximal size of a Jordan block associated to* μ *(i.e. the* *maximal defect associated to* μ *is* $M-1$*). Then there exists a constant* $\mathscr{C} > 0$*, such that the solutions to* (2.2.1) *satisfy*

$$|x(t)|_2^2 \le \mathscr{C}(1+t^{2(M-1)})e^{-2\mu t}|x(0)|_2^2. \tag{2.2.21}$$

The constant $\mathscr{C}$ *can be chosen as*

$$\mathscr{C} := \begin{cases} (\lambda_{\min}^P)^{-1}\lambda_{\max}^P, & M=1, \\ 2(\lambda_{\min}^{P(0)})^{-1}\lambda_{\max}^{P(0)} c_M \max\limits_{n\in I_\mu}\Big[\sum\limits_{m=1}^{l_n} \dfrac{\beta_n^m}{\min\limits_{k\in\{1,\dots,m\}} \beta_n^k}\Big], & M\ge 2, \end{cases} \tag{2.2.22}$$

where $\lambda_{\min}^{P(0)}$ *is the smallest and* $\lambda_{\max}^{P(0)}$ *is the largest eigenvalue of the matrix* $P(0)$ *(with* $P(0)\equiv P$ *for* $M=1$*), which is defined in* (2.2.13)*. The constants* c_M *for* $M\ge 2$ *are given as:*

$$c_M = 2^{M-2}\left(\prod_{j=1}^{M-1} 4j^2-1\right)\left(\prod_{j=2}^{M}\sum_{k=1}^{j}\frac{1}{[(j-k)!]^2}\right).$$

The technical proof of this result is deferred to Appendix 2.A.

Remark 2.2.9. *If, for all* $n \in I_\mu$*, the weights* β_n^m *of the matrix* $P(0)$ *are monotonically decreasing in* m*, i.e.* $\beta_n^m \ge \beta_n^{m+1}$ *for* $m \in \{1,\dots,l_n-1\}$*, then*

$$\max_{n\in I_\mu}\Big[\sum_{m=1}^{l_n}\frac{\beta_n^m}{\min\limits_{k\in\{1,\dots,m\}}\beta_n^k}\Big] = M.$$

Remark 2.2.10. *Note that we could calculate the solution to the ODE system* (2.2.1) *directly, by means of its Jordan transformation, and get qualitatively the same decay behavior as in* (2.2.21)*, but possibly with a different multiplicative constant: Using* $x(t) = e^{-Ct}x(0)$*,* $C = (V^H)^{-1}J^HV^H$*, we obtain*

$$\begin{aligned} |x(t)|_2^2 &\le |V|_2^2|V^{-1}|_2^2|e^{-J^Ht}|_2^2|x(0)|_2^2 \\ &\le |V|_2^2|V^{-1}|_2^2\hat{c}_M(1+t^{2(M-1)})e^{-2\mu t}|x(0)|_2^2, \end{aligned} \tag{2.2.23}$$

where $|V|_2$ *denotes the matrix norm induced by the vector norm* $|\cdot|_2$*,* V *is the transformation matrix from* (2.2.2)*,* $J := \mathrm{diag}(J_1,\dots,J_N)$ *the corresponding Jordan matrix, and* $\hat{c}_M$ *depends only on the largest Jordan block. So what is the gain of the result in Theorem 2.2.8?*

Firstly, the construction of the matrix $P(t)$ *and the method of estimating the* $P(t)$*-norm decay of the solution can be translated almost directly to the infinite dimensional setting of the Fokker–Planck equation with linear drift, where a direct way of calculating the decay (as in the finite dimensional ODE case) is not possible (see [5] for the exponential decay in the non-defective case and [22] for an improved decay in the defective case). In the Fokker–Planck setting on* $\mathbb{R}^d$*, the place of the* $P(t)$*-norm is taken by the modified Fisher information involving* $P(t)$*.*

Secondly, the result also makes it possible to systematically calculate the multiplicative constant $\mathscr{C}$ *from* (2.2.21)*, which we will use in §2.3–§2.5, and which can be further exploited to get decay results for infinite dimensional ODE systems (see [1], §4.3).*

2.2.2 Uniform Decay Estimates in Non-Defective Limits

The advantage of the $P(t)$-norm estimation compared to (2.2.23) can be seen in the following example.

Example 2.2.11. Consider the matrix

$$C_\varepsilon := \begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix}$$

with arbitrary $\varepsilon \neq 0$. Its corresponding Jordan transformation matrix reads

$$V_\varepsilon := \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\varepsilon} \end{pmatrix},$$

and $M = 2$. For $\varepsilon \to 0$ the factor $|V_\varepsilon|_2 |V_\varepsilon^{-1}|_2$ in (2.2.23) becomes unbounded of order ε^{-1} (even though the true decay of the solution improves to $e^{-t}|x(0)|_2$ in the limit). This is due to the discontinuity of the Jordan transformation at the transition from defectiveness to non-defectiveness. We apply Theorem 2.2.8 to the ODE system $\dot{x} = -C_\varepsilon x$, with the following eigenvectors of C_ε^H:

$$v_1^{(0)} = \begin{pmatrix} 0, & 1 \end{pmatrix}^T \quad \text{and} \quad v_1^{(1)} = \begin{pmatrix} \frac{1}{\varepsilon}, & 0 \end{pmatrix}^T.$$

When using $\beta^1_{1,\varepsilon} = 1$ and $\beta^2_{1,\varepsilon} = \varepsilon^2$ in (2.2.13), we get $P_\varepsilon(0) = I$, and hence the constant

$$\mathscr{C}_\varepsilon = 12 \cdot \max\{2, 1 + \varepsilon^2\} \tag{2.2.24}$$

stays bounded in the *non-defective limit* $\varepsilon \to 0$. ◊

This example shows that, while the method presented here is still relying on the Jordan transformation (as $P(t)$ is constructed with generalized eigenvectors), the additional weights β_n^m in the matrix $P(t)$ allow for estimates more closely related to the actual behavior of the solutions. As sketched in Example 2.2.11 this can allow for (but does not guarantee) an estimate that is uniform in the *non-defective limit.* In §2.3–§2.5 we will see further examples of specific choices of the weights β_n^m that lead to uniform estimates in the non-defective limit.

Remark 2.2.12. *While the proof of Theorem 2.2.8 is formulated to work for arbitrarily large defects, more careful estimations can improve the decay estimate. We shall now show an improvement for defect one. For any $n \in I_\mu$ the inequality* (2.A.2) *with $m = 2$ and $\theta = \frac{1}{2}$ yields*

$$|x(t)|^2_{P_n^2(0)} \le 2e^{-2\mu t}\left(t^2 |x(0)|^2_{P_n^1(0)} + |x(0)|^2_{P_n^2(0)}\right),$$

where we used (2.2.10) *with $P_n^1(t) = P_n^1(0)$.*

For $n \in I_\mu$ with $l_n = 2$ the decay estimate for $P_n(0) = \beta_n^1 P_n^1(0) + \beta_n^2 P_n^2(0)$ follows as

$$|x(t)|^2_{P_n(0)} \le 2e^{-2\mu t}(1 + \frac{\beta_n^2}{\beta_n^1} t^2)|x(0)|^2_{P_n(0)}.$$

We can use this estimate to get an improved upper bound for solutions from Example 2.2.11 in the Euclidean norm (compared to (2.2.21) *with* (2.2.24)*): With the same matrix choice $P_\varepsilon(0) = P_1^1(0) + \varepsilon^2 P_1^2(0) = I$ as in Example 2.2.11, it follows that solutions to $\dot{x} = -C_\varepsilon x$ satisfy*

$$|x(t)|_2^2 = |x(t)|^2_{P_\varepsilon(0)} \le 2e^{-2t}(1 + \varepsilon^2 t^2)|x(0)|_2^2, \quad t \ge 0. \tag{2.2.25}$$

This estimate not only yields a bounded multiplicative constant for $\varepsilon \to 0$, but also yields a sharp decay rate — namely purely exponential — for the non-defective limit case $\varepsilon = 0$. In comparison, the solution propagator norm estimate for all $\varepsilon \in \mathbb{R}$ is given as

$$\begin{aligned} |e^{-C_\varepsilon t} x(0)|_2^2 &= e^{-2t} \left\| \begin{pmatrix} 1 & -\varepsilon t \\ 0 & 1 \end{pmatrix} \right\|_2^2 |x(0)|_2^2 \\ &= e^{-2t} \left(1 + \frac{\varepsilon^2 t^2}{2} + \sqrt{\varepsilon^2 t^2 + \frac{\varepsilon^4 t^4}{4}} \right) |x(0)|_2^2 \qquad (2.2.26) \\ &\overset{\varepsilon t \to \infty}{\approx} e^{-2t}(2 + \varepsilon^2 t^2)|x(0)|_2^2. \end{aligned}$$

This shows that (2.2.25) *is rather accurate.*

Remark 2.2.13. *Consider an ODE* (2.2.1) *with matrix C that has a Jordan block J_{n_2} in Case 2 from §2.2.1, i.e. $l_{n_2} > 1$ and $\mathrm{Re}(\lambda_{n_2}) > \mu$. Then, the construction of Case 2 can be replaced with the one of Case 3. This means, exchanging the time-constant matrix P_{n_2}, which has predetermined weights $b^m_{n_2}$, by a time-dependent matrix $\widetilde{P}_{n_2}(t)$, which allows for arbitrary weights $\beta^m_{n_2} > 0$. For simplicity let us assume $M = 1$. With the appropriate (straightforward) modifications of Lemma 2.2.1 and Theorem 2.2.8 (treating J_{n_2} as Case 3), this yields the decay estimate* (2.2.21) *with the modified constant*

$$\widetilde{\mathscr{C}} := 2(\lambda_{\min}^{\widetilde{P}(0)})^{-1} \lambda_{\max}^{\widetilde{P}(0)} c_{l_{n_2}} \Big[\sum_{m=1}^{l_{n_2}} \frac{\beta^m_{n_2}}{\min\limits_{k \in \{1,\dots,m\}} \beta^k_{n_2}} \Big], \tag{2.2.27}$$

where $\widetilde{P}(t)$ is the matrix defined by (2.2.13)*, but with $\widetilde{P}_{n_2}(t)$ instead of P_{n_2}.*

For ODE families $\dot{x}_\varepsilon = -C_\varepsilon x_\varepsilon$ and their non-defective limits $\varepsilon \to 0$, this modification can be beneficial: The additional weights of $\widetilde{P}_{n_2}(t)$ provide further possibilities to obtain a multiplicative constant $\widetilde{\mathscr{C}}_\varepsilon$ that is bounded for $\varepsilon \to 0$. In §2.5.2 (Case $k = 3$), we will see an example of an ODE family where Theorem 2.2.8 yields an unbounded constant $\mathscr{C}_\varepsilon$ (for all possible weights) but a bounded constant $\widetilde{\mathscr{C}}_\varepsilon$ for $\varepsilon \to 0$ (with the correct choice of weights).

2.2.3 Uniform Decay for a Family of ODEs

We shall consider now an extension of Example 2.2.11 which will be relevant for the PDEs discussed in §2.3–§2.5: We consider the matrix family with parameter $z \in \mathbb{R}$

$$C(z) := \begin{pmatrix} \mu(z) & \mu'(z) \\ 0 & \mu(z) \end{pmatrix} = \mu(z) \begin{pmatrix} 1 & \frac{\mu'(z)}{\mu(z)} \\ 0 & 1 \end{pmatrix} \tag{2.2.28}$$

with a given function $\mu \in C^1(\mathbb{R})$ and $\mu(z) \geq \mu_{\min} = \mu(z_0) > 0$. For simplicity let $z_0 \in \mathbb{R} \cup \{\infty, -\infty\}$ be the unique global point of minimum (infimum if $|z_0| = \infty$) of μ.

We are now interested in a uniform-in-z estimate on the matrix propagator $e^{-C(z)t}$ with $t \geq 0$, based on the estimate (2.2.25). To this end we have to consider the interplay of two effects: On the one hand the parameter value $z = z_0$ yields the smallest exponential decay rate $\mu_{\min}$ but it is without defect, since $\mu'(z_0) = 0$ makes $C(z_0)$ a diagonal matrix. On the other hand the parameters $z \neq z_0$ yield a larger decay, but with a defect (as long as $\mu'(z) \neq 0$). Hence we shall be interested in the question, whether or not the typical defective decay of the form $\mathcal{O}((1+t^2)e^{-2\mu_{\min}t})$ persists for the uniform estimate of $|e^{-C(z)t}|^2$ for $t \to +\infty$. In the subsequent examples we shall illustrate that both scenarios are in fact possible.

Example 2.2.14. Let $\mu(z) := \mu_{\min} + \alpha z^2$ with some $\alpha > 0$, and hence $z_0 = 0$. From Example 2.2.11 with (2.2.25) (using $\varepsilon = \frac{\mu'(z)}{\mu(z)}$, $t \mapsto t\mu(z)$) we obtain the uniform decay estimate

$$|e^{-C(z)t}|_2^2 \leq 2e^{-2\mu_{\min}t} \sup_{z\in\mathbb{R}} f_1(z,t), \quad z \in \mathbb{R}, t \geq 0, \tag{2.2.29}$$

with

$$f_1(z,t) := (1 + 4\alpha^2 z^2 t^2)e^{-2\alpha z^2 t}.$$

An elementary computation yields

$$\sup_{z\in\mathbb{R}} f_1(z,t) = \begin{cases} 1, & \alpha t \leq \frac{1}{2}, \\ 2\alpha t e^{-\frac{2\alpha t-1}{2\alpha t}}, & \alpha t > \frac{1}{2}, \end{cases}$$

with the asymptotic behavior $\sup_{z\in\mathbb{R}} f_1(z,t) = \mathcal{O}(\frac{2\alpha}{e} t)$ as $t \to +\infty$. Hence, estimate (2.2.29) exhibits the typical defective decay behavior, and the term $te^{-2\mu_{\min}t}$ cannot be dropped in the estimate. ◊

Example 2.2.15. Let $\mu(z) := \mu_0 + \alpha e^{\beta z}$ with some $\alpha > 0$ and $\beta \in \mathbb{R} \setminus \{0\}$ with $|\beta| < 2$. Here, $z_0 = -\infty \operatorname{sgn}(\beta)$. Then (2.2.25) yields the uniform decay estimate

$$|e^{-C(z)t}|_2^2 \leq 2e^{-2\mu_0 t} \sup_{z\in\mathbb{R}} f_2(z,t), \quad z \in \mathbb{R}, t \geq 0, \tag{2.2.30}$$

with

$$f_2(z,t) := (1 + \alpha^2\beta^2 e^{2\beta z} t^2)e^{-2\alpha e^{\beta z} t}.$$

Since $\partial_t f_2(z,0) = -2\alpha e^{\beta z} < 0$ and $\partial_t f_2(z,t) \neq 0$ for every $z \in \mathbb{R}$, $t \geq 0$, we conclude

$$|e^{-C(z)t}|_2^2 \leq 2e^{-2\mu_0 t} \sup_{z\in\mathbb{R}} f_2(z,0) = 2e^{-2\mu_0 t}, \quad z \in \mathbb{R}, t \geq 0.$$

Hence, this example shows a purely exponential decay behavior, which is rather typical for the non-defective case. ◊

We remark that we could not find an example of a parameter function $\mu \in C^1(\mathbb{R})$ with a minimum at $|z_0| < \infty$ for which the algebraic factor vanishes in the uniform estimate.

In the following sections §2.3–§2.5 we investigate parabolic and kinetic evolution equations in which equation coefficients depend on an uncertainty variable $z \in \mathbb{R}$. After a Fourier decomposition, the sensitivity analysis leads to families of defective ODE systems of type similar to (2.2.28), for which we are interested in uniform-in-z decay estimates of solutions with sharp rate. Sharpness is understood here in the sense that in the class of $C^1(\mathbb{R})$ parameter functions $\mu(z)$ satisfying $\mu_0 := \inf_{z\in\mathbb{R}} \mu(z) > 0$, the decay estimate is of type $\mathcal{O}(1 + t^m)e^{-2\mu_0 t}$ for large t, with minimal $m \in \mathbb{N}_0$. As Example 2.2.14 illustrates, $m > 0$ is necessary to cover arbitrary $C^1(\mathbb{R})$ parameter functions.

With the three examples of §3–§5 we shall illustrate the various challenges of this procedure to obtain estimates: *uniformity in the Fourier modes and in the non-defective limit(s).*

2.3 Linear Convection-Diffusion Equations with Uncertain Coefficients

First we consider the parabolic equation on the 1D torus

$$\partial_t u(x,z,t) = -a(z)\partial_x u(x,z,t) + b(z)\partial_x^2 u(x,z,t), \quad x \in \mathbb{T}^1, t \geq 0, \tag{2.3.1}$$

$$u(x,z,0) = u^0(x,z), \tag{2.3.2}$$

for $u(x,z,t) \in \mathbb{R}$ with the space variable x, convection coefficient $a(z) \in \mathbb{R}$ and diffusion coefficient $b(z)$ satisfying $b_0 := \inf_{z\in\mathbb{R}} b(z) > 0$. We are interested in the sensitivity of solutions with respect to the uncertainty parameter $z \in \mathbb{R}$ contained in the coefficients. We assume that the coefficients satisfy $a, b \in C^1(\mathbb{R})$. For each $z \in \mathbb{R}$, the equation is mass conserving (in time), i.e. $\frac{1}{2\pi}\int_0^{2\pi} u(x,z,t)dx = const.$, and the unique normalized steady state is given as $u^\infty(x,z) = 1$. Correspondingly we shall also assume that the initial condition is normalized as $\frac{1}{2\pi}\int_0^{2\pi} u^0(x,z)dx = 1$ for all $z \in \mathbb{R}$.

A Fourier expansion of u with respect to $x \in \mathbb{T}^1$, allows to rewrite the PDE (for each fixed z) as a family of ODEs. With the notation $u(x,z) = \sum_{k\in\mathbb{Z}} u_k(z)e^{ikx}$, the equation for each Fourier mode u_k, $k \in \mathbb{Z}$ reads

$$\partial_t u_k(z) = -ika(z)u_k(z) - k^2 b(z)u_k(z), \tag{2.3.3}$$

with the explicit solutions $u_k(z,t) = e^{-k^2 b(z)t - ika(z)t}u_k(z,0)$. Due to the above normalization we have for the $k = 0$ mode: $u_0(z,t) = u_0^\infty(z) = 1$.

2.3.1 First Order Parameter Sensitivity Analysis

Now we analyze the (linear order) sensitivity of the equation with respect to the uncertainty in the coefficients $a(z)$ and $b(z)$. Therefore we consider the evolution equation for $v(x,z,t) := \partial_z u(x,z,t)$, given as

$$\partial_t v(z) = -(\partial_z a(z))\partial_x u(z) + (\partial_z b(z))\partial_x^2 u(z) - a(z)\partial_x v(z) + b(z)\partial_x^2 v(z). \tag{2.3.4}$$

The Fourier modes $v_k(z,t) := \partial_z u_k(z,t)$ for $k \in \mathbb{Z}$ satisfy

$$\partial_t v_k(z) = -ik(\partial_z a(z))u_k(z) - k^2(\partial_z b(z))u_k(z) - ika(z)v_k(z) - k^2 b(z)v_k. \tag{2.3.5}$$

For $k \in \mathbb{Z} \setminus \{0\}$ the system of (2.3.3) and (2.3.5) reads

$$\partial_t \underbrace{\begin{pmatrix} u_k \\ v_k \end{pmatrix}}_{y_k(z,t):=} = -k^2 \underbrace{\begin{pmatrix} b(z) + \frac{ia(z)}{k} & 0 \\ \partial_z b(z) + \frac{i\partial_z a(z)}{k} & b(z) + \frac{ia(z)}{k} \end{pmatrix}}_{C_k(z):=} \begin{pmatrix} u_k \\ v_k \end{pmatrix}. \tag{2.3.6}$$

Our goal is to obtain a decay estimate with sharp decay rate for solutions to (2.3.6), uniform in $z \in \mathbb{R}$ by applying Theorem 2.2.8.

Due to the normalization of the initial conditions u^0, we have $v_0(z,t) = \frac{1}{2\pi}\int_0^{2\pi} v(x,z,t)dx \equiv 0$, and in particular for the initial condition $\int_0^{2\pi} v^0(x,z)dx = 0$. Hence, its steady state is $v_0^\infty(z) = 0$; the (expected) decay of all higher modes $v_k, k \in \mathbb{Z} \setminus \{0\}$ implies $v^\infty(x,z) \equiv 0$. With the notation $y := (u,v)^T$, the unique (normalized) steady state of the system (2.3.1), (2.3.4) is $y^\infty(x,z) \equiv (1,0)^T$.

For each Fourier mode $k \in \mathbb{Z} \setminus \{0\}$, the double eigenvalue of the matrix $C_k(z)$ is given as

$$\lambda_k(z) := b(z) + \frac{ia(z)}{k}.$$

Hence, the matrix is positive stable and the steady state is given as $y_k^\infty = 0 \in \mathbb{C}^2$. The spectral gap of the evolution operator in (2.3.6) is given by $\mu_k(z) := k^2 b(z) > 0$ and the eigenvalue $\lambda_k(z)$ of $C_k(z)$ is defective, if and only if $\partial_z \lambda_k(z) \neq 0$.

If we consider all Fourier modes, the spectral gap for the whole sequence $\{y_k(z)\}_{k\in\mathbb{Z}\setminus\{0\}}$ is given as

$$\mu(z) := \min_{k\in\mathbb{Z}\setminus\{0\}} \mu_k(z) = \mu_{\pm 1}(z) = b(z),$$

and it is realized by the modes $k = \pm 1$. The steady state is given by the sequence $\{y_k^\infty\}_{k\in\mathbb{Z}} = \{(\delta_{0k}, 0)^T\}_{k\in\mathbb{Z}}$, with δ_{0k} denoting the Kronecker delta.

In what follows, we investigate the decay rate of solutions to the system of equations (2.3.1) and (2.3.4) towards the steady state with a sharp rate, uniform in the uncertainty variable z.

The solution vector $y(\cdot, z, t) \in L^2(0, 2\pi)$ is equivalent to $\{y_k(z,t)\}_{k\in\mathbb{Z}}$ by Parseval's identity. As the ODE system for each Fourier mode $k \in \mathbb{Z}\setminus\{0\}$ can be defective for certain values in $z \in \mathbb{R}$, we expect a uniform-in-z decay rate that is not purely exponential. In fact, for non-constant a and b, defectiveness is the more typical behavior of the matrix $C_k(z)$.

For each fixed $k \in \mathbb{Z}\setminus\{0\}$, we proceed with a case distinction between the defective and non-defective case.

Case 1; $z \in \mathbb{R}$, such that $\partial_z \lambda_k(z) = 0$:

The matrix $C_k(z)$ is diagonal and the solutions are given as $y_k(z,t) = e^{-k^2\lambda_k(z)t} y_k(z,0)$. Hence, for each k and z, the decay of solutions to (2.3.6) is given as

$$|y_k(z,t)|_2^2 = e^{-2k^2 b(z)t} |y_k(z,0)|_2^2. \tag{2.3.7}$$

Case 2; $z \in \mathbb{R}$, such that $\partial_z \lambda_k(z) \neq 0$:

In this case the eigenvalue $\lambda_k(z)$ of $C_k(z)$ is defective of order 1, i.e. $M = 2$, for $k \in \mathbb{Z}\setminus\{0\}$. Hence we shall apply Theorem 2.2.8 to get a sharp decay estimate for solutions to (2.3.6).

To denote the dependence on the Fourier mode $k \in \mathbb{Z}$, we shall use the subscript k, e.g., we use $P_k(z,0)$ for the matrix $P(z,0)$, defined in (2.2.13), that corresponds to $C_k(z)$. Similarly, we denote $P_k(z,0)$'s weights β_n^m in (2.2.13) by $\beta_{n,k}^m$.

For $k \in \mathbb{Z}\setminus\{0\}$, the matrix $C_k(z)$ is positive stable and the eigenvector and generalized eigenvector of $C_k^H(z)$, corresponding to $\overline{\lambda}_k(z)$, are given as

$$v_{1,k}^{(0)} = \begin{pmatrix}1, & 0\end{pmatrix}^T, \qquad v_{1,k}^{(1)}(z) = \begin{pmatrix}0, & \frac{1}{\partial_z \overline{\lambda}_k(z)}\end{pmatrix}^T.$$

In analogy to Example 2.2.11 we choose $\beta_{1,k}^1 = 1$ and $\beta_{1,k}^2(z) = |\partial_z \lambda_k(z)|^2$, leading to

$$P_k(z,0) = v_{1,k}^{(0)} \otimes v_{1,k}^{(0)} + |\partial_z \lambda_k(z)|^2 v_{1,k}^{(1)} \otimes v_{1,k}^{(1)} = I.$$

An appropriate choice of $\beta_{1,k}^m$ is essential here, in order to make the matrix $P_k(z,0)$ (and hence the constant $\mathscr{C}_k(z)$ below) uniformly bounded for $\partial_z\lambda_k(z) \to 0$, i.e. for the *non-defective limit*.

Now we can apply Theorem 2.2.8 (for the rescaled time $\tau_k = k^2 t$) to get the following decay estimate for solutions to the system (2.3.6) for each $k \in \mathbb{Z}\setminus\{0\}$ and $z \in \mathbb{R}$:

$$|y_k(z,t)|_2^2 \le \mathscr{C}_k(z)(1 + k^4 t^2) e^{-2k^2 b(z)t} |y_k(z,0)|_2^2, \tag{2.3.8}$$

with the constant $\mathscr{C}_k(z) \ge 0$ defined in (2.2.22), given as

$$\mathscr{C}_k(z) = 12 \cdot \left(1 + \frac{|\partial_z \lambda_k(z)|^2}{\min\{1, |\partial_z \lambda_k(z)|^2\}}\right) = 12 \cdot \max\{2, 1 + |\partial_z \lambda_k(z)|^2\}. \tag{2.3.9}$$

Combining Case 1 and Case 2, we infer a decay estimate for the first order parameter sensitivity equations by applying Parseval's identity:

Theorem 2.3.1. *Let $a, b \in C^1(\mathbb{R})$ where $b_0 := \inf_{z\in\mathbb{R}} b(z) > 0$ and $\partial_z a, \partial_z b \in L^\infty(\mathbb{R})$. Then, there exists a constant $\mathscr{C} > 0$, such that normalized solutions $y(x,z,t) = (u(x,z,t), v(x,z,t))^T$ of the system* (2.3.1), (2.3.4) *with steady state $y^\infty := (1,0)^T$ satisfy*

$$\sup_{z\in\mathbb{R}} \|y(\cdot,z,t) - y^\infty\|^2_{L^2(0,2\pi;\mathbb{R}^2)} \le \mathscr{C}(1+t^2)e^{-2b_0 t} \sup_{z\in\mathbb{R}} \|y(\cdot,z,0) - y^\infty\|^2_{L^2(0,2\pi;\mathbb{R}^2)}$$

for $t \ge 0$.

Proof. Combining both estimates (2.3.7) and (2.3.8) leads to

$$|y_k(z,t)|_2^2 \le \widetilde{\mathscr{C}}(1+k^4t^2)e^{-2k^2b(z)t}|y_k(z,0)|_2^2, \quad k \in \mathbb{Z}\setminus\{0\}, \tag{2.3.10}$$

with the constant

$$\widetilde{\mathscr{C}} = \sup_{k\neq 0, z\in\mathbb{R}} \mathscr{C}_k(z) \le 12\max\{2, 1 + \|\partial_z a\|_\infty^2 + \|\partial_z b\|_\infty^2\}$$

independent of $z \in \mathbb{R}$ and $k \in \mathbb{Z}\setminus\{0\}$. With Parseval's identity we obtain

$$\begin{aligned}\|y(\cdot,z,t) - y^\infty\|^2_{L^2(0,2\pi;\mathbb{R}^2)} &= \frac{1}{2\pi}\sum_{k\in\mathbb{Z}} |y_k(z,t) - y_k^\infty|_2^2 \\ &\le \frac{1}{2\pi}\sum_{k\in\mathbb{Z}\setminus\{0\}} \widetilde{\mathscr{C}}(1+k^4t^2)e^{-2k^2b(z)t}|y_k(z,0)|_2^2 \\ &\le \mathscr{C}(1+t^2)e^{-2b_0t}\|y(\cdot,z,0) - y^\infty\|^2_{L^2(0,2\pi;\mathbb{R}^2)},\end{aligned}$$

where we used the estimate

$$(1+k^4t^2)e^{-2k^2b_0t} \le c(1+t^2)e^{-2b_0t}, \quad t \ge 0, k \neq 0,$$

with $c := \max_{t\ge0}(1+t^2)e^{-2b_0t}$. Taking the supremum over $z \in \mathbb{R}$ completes the proof. □

2.3.2 Second Order Parameter Sensitivity Analysis

Next we shall extend the above analysis to second order. This will also illustrate the challenges involved in obtaining uniform decay estimates in defective limits.

We assume $a, b \in C^2(\mathbb{R})$ and denote $w(x,z,t) := \partial_z^2 u(x,z,t)$. By differentiation of (2.3.4) with respect to z, the second order sensitivity equation is given as

$$\begin{aligned}\partial_t w(z) = &-(\partial_z^2 a(z))\partial_x u(z) + (\partial_z^2 b(z))\partial_x^2 u(z) \\ &- 2(\partial_z a(z))\partial_x v(z) + 2(\partial_z b(z))\partial_x^2 v(z) \\ &- a(z)\partial_x w(z) + b(z)\partial_x^2 w(z).\end{aligned} \tag{2.3.11}$$

The system for the Fourier mode $k \in \mathbb{Z} \setminus \{0\}$ of $(u,v,w)^T$, with $w_k(z,t) := \partial_z v_k(z,t)$, is given as

$$\partial_t \underbrace{\begin{pmatrix} u_k \\ v_k \\ w_k \end{pmatrix}}_{y_k(z,t):=} = -k^2 \underbrace{\begin{pmatrix} \lambda_k(z) & 0 & 0 \\ \partial_z \lambda_k(z) & \lambda_k(z) & 0 \\ \partial_z^2 \lambda_k(z) & 2\partial_z \lambda_k(z) & \lambda_k(z) \end{pmatrix}}_{D_k(z):=} \begin{pmatrix} u_k \\ v_k \\ w_k \end{pmatrix}. \tag{2.3.12}$$

As before, $w_0(z,t) \equiv 0$, and in particular for the initial condition $\int_0^{2\pi} w^0(x,v)dx = 0$. Hence the unique (normalized) steady state of the second order sensitivity system is $y^\infty(x,z) := (u^\infty(x,z), v^\infty(x,z), w^\infty(x,z))^T \equiv (1,0,0)^T$.

The triple eigenvalue is $\lambda_k(z)$, with $\mathrm{Re}(\lambda_k(z)) > 0$ and its defectiveness for $k \in \mathbb{Z} \setminus \{0\}$ depends on the values of $\partial_z \lambda(z)$ and $\partial_z^2 \lambda(z)$, i.e.

$$\operatorname{rank}(D_k(z) - \lambda_k(z)) = \begin{cases} 0, & \text{if } \partial_z \lambda_k(z) = \partial_z^2 \lambda_k(z) = 0, \\ 1, & \text{if } \partial_z \lambda_k(z) = 0 \,\&\, \partial_z^2 \lambda_k(z) \neq 0, \\ 2, & \text{if } \partial_z \lambda_k(z) \neq 0. \end{cases}$$

As in the first order analysis, we need to discuss the decay behavior of these three cases separately:

Case 1; $z \in \mathbb{R}$ such that $\partial_z \lambda(z) = \partial_z^2 \lambda(z) = 0$:

In this case, the eigenvalue $\lambda_k(z)$ is non-defective and the solutions are given as $y_k(z,t) = e^{-k^2 \lambda_k(z) t} y_k(z,0)$ from which we obtain the decay

$$|y_k(z,t)|_2^2 = e^{-2k^2 b(z) t} |y_k(z,0)|_2^2. \tag{2.3.13}$$

Case 2; $z \in \mathbb{R}$, such that $\partial_z \lambda(z) = 0$ and $\partial_z^2 \lambda(z) \neq 0$:

In this case the eigenvalue $\lambda_k(z)$ is defective of order one, i.e. $M = 2$. To obtain a sharp decay estimate of solutions, we construct $P_k(z,0)$ according to (2.2.13): $N = 2$, $l_1 = 2$, $l_2 = 1$, $M = 2$ and choosing $\beta_{1,k}^1 = 1$, $\beta_{1,k}^2 = |\partial_z^2 \lambda_k(z)|^2$ and $\beta_{2,k}^1 = 1$. The (generalized) eigenvectors of $D_k^H(z)$ are given as

$$v_{1,k}^{(0)} = \begin{pmatrix} 1, & 0, & 0 \end{pmatrix}^T, \qquad v_{1,k}^{(1)}(z) = \begin{pmatrix} 0, & 0, \frac{1}{\partial_z^2 \overline{\lambda}_k(z)} \end{pmatrix}^T, \qquad v_{2,k}^{(0)} = \begin{pmatrix} 0, & 1, & 0 \end{pmatrix}^T.$$

This leads to

$$P_k(z,0) = v_{1,k}^{(0)} \otimes v_{1,k}^{(0)} + |\partial_z^2 \lambda_k(z)|^2 v_{1,k}^{(1)}(z) \otimes v_{1,k}^{(1)}(z) + v_{2,k}^{(0)} \otimes v_{2,k}^{(0)} = I.$$

We can now apply Theorem 2.2.8 (for the rescaled time $\tau_k = k^2 t$) and get the decay estimate

$$|y_k(z,t)|_2^2 \leq \mathscr{C}_k(z)(1 + k^4 t^2) e^{-2k^2 b(z) t} |y_k(z,0)|_2^2 \tag{2.3.14}$$

with the constant $\mathscr{C}_k(z)$, defined in (2.2.22), given as

$$\mathscr{C}_k(z) = 12 \cdot \max\{2, 1 + |\partial_z^2 \lambda_k(z)|^2\}.$$

Note that this constant $\mathscr{C}_k(z)$ is uniformly bounded in the *non-defective limit* $\partial_z^2\lambda_k(z) \to 0$ (from defect 1 to non-defective), but it does not reduce to (2.3.13), hence it is not uniformly sharp.

Case 3; $z \in \mathbb{R}$, such that $\partial_z \lambda_k(z) \neq 0$:

The eigenvalue $\lambda_k(z)$ is defective of order two with $N = 1$, $l_1 = 3$ and $M = 3$. The (generalized) eigenvectors of $D_k^H(z)$ are given as

$$v_{1,k}^{(0)}(z) = \begin{pmatrix}1, & 0, & 0\end{pmatrix}^T, \quad v_{1,k}^{(1)}(z) = \begin{pmatrix}0, & \frac{1}{\partial_z \overline{\lambda}_k(z)}, & 0\end{pmatrix}^T,$$
$$v_{1,k}^{(2)}(z) = \begin{pmatrix}0, & \frac{-\partial_z^2 \overline{\lambda}_k(z)}{2(\partial_z \overline{\lambda}_k(z))^3}, & \frac{1}{2(\partial_z \overline{\lambda}_k(z))^2}\end{pmatrix}^T.$$

For this case the previous strategy of finding weights for $P(t)$ (as defined in (2.2.13)) that give a uniform in z decay estimate for solutions via Theorem 2.2.8 does not work directly. All choices of weights $\beta_{1,k}^j(z)$ for $j \in \{1,2,3\}$ lead to constants $\mathscr{C}_k(z)$ (defined in (2.2.22)) that are not bounded uniformly in z. The problem arises for the defective limit from defect 2 to defect 1, more precisely, for sequences $(z_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ such that $0 \neq \partial_z \lambda_k(z_n) \to 0$ in combination with $\frac{\partial_z^2 \lambda_k(z_n)}{\partial_z \lambda_k(z_n)} \not\to 0$ as $n \to \infty$. All weight choices of $\beta_{1,k}^j(z_n)$ lead to $(\lambda_{\min}^{P_k(z_n,0)})^{-1} \lambda_{\max}^{P_k(z_n,0)} \to \infty$ due to the three different powers of $\partial_z \overline{\lambda}_k(z_n)$ that appear in the (generalized) eigenvectors of $D_k^H(z_n)$.

However, this problem can be fixed with small adjustments to the proof of Theorem 2.2.8, which yield a uniform in z decay estimate.

Define

$$\widetilde{w}_k^3(z,t) := w_{1,k}^3(z,t) + \frac{\partial_z^2 \overline{\lambda}_k(z)}{2(\partial_z \overline{\lambda}_k(z))^2} w_{1,k}^2(z,t),$$

which is our replacement for $w_k^3(z,t)$, with $w_{1,k}^j(z,t)$ for $j \in \{2,3\}$ from (2.2.7). It satisfies

$$\widetilde{w}_k^3(z,0) = \begin{pmatrix}0, & 0, & \frac{1}{2(\partial_z \overline{\lambda}_k(z))^2}\end{pmatrix}^T,$$

which eliminates the problematic factor $2(\partial_z \overline{\lambda}_k(z))^3$ present in $w_k^3(z,0)$. The corresponding semi-norm matrix is

$$\widetilde{P}_k^3(z,t) := \widetilde{w}_k^3(z,t) \otimes \widetilde{w}_k^3(z,t).$$

Similar to definition (2.2.13), let

$$\widetilde{P}_k(z,t) := P^1_{1,k}(z,t) + |\partial_z\lambda_k(z)|^2 P^2_{1,k}(z,t) + 4|\partial_z\lambda_k(z)|^4 \widetilde{P}^3_k(z,t), \tag{2.3.15}$$

which is positive definite for all $k \neq 0$, $z \in \mathbb{R}$, $t \geq 0$ and satisfies $\widetilde{P}_k(z,0) = I$.

As definition (2.2.13) is modified, we cannot directly use Theorem 2.2.8 to get a decay estimate in the Euclidean norm. The idea of the proof of Theorem 2.2.8 is to estimate each P^m_n-semi-norm decay separately, see (2.A.3). Combining them yields a decay estimate in the P-norm. We will now follow the idea of the proof of Theorem 2.2.8 but have to carefully modify each step to work for $\widetilde{P}_k$.

First, one can easily verify that for $m = 1,2$ the estimate (2.A.3) remains true, if we replace $P_k(z,0)$ by $\widetilde{P}_k(z,0) = I$:

$$|y_k(z,t)|^2_{P^1_{1,k}(z,0)} \leq e^{-2k^2 b(z)t}|y_k(z,0)|^2_2, \quad t \geq 0, k \neq 0, \tag{2.3.16}$$

and (using $c_2 = 6$)

$$|y_k(z,t)|^2_{P^2_{1,k}(z,0)} \leq \frac{6}{\min\{1,|\partial_z\lambda_k(z)|^2\}}(1+k^4t^2)e^{-2k^2b(z)t}|y_k(z,0)|^2_2, \tag{2.3.17}$$

for $t \geq 0$, $k \neq 0$.

Next we shall derive a similar estimate for the semi-norm $|\cdot|_{\widetilde{P}^3_k(z,0)}$.

Lemma 2.3.2. *Let $y_k(z,t)$ be a solution of the ODE* (2.3.12) *and $z \in \mathbb{R}$, such that $\partial_z\lambda_k(z) \neq 0$. Then,*

$$|y_k(z,t)|^2_{\widetilde{P}^3_k(z,0)} \leq 146.25\frac{1+|\partial_z^2\lambda_k(z)|^2}{\min\{1,|\partial_z\lambda_k(z)|^4\}}(1+k^8t^4)e^{-2k^2b(z)t}|y(z,0)|^2_2 \tag{2.3.18}$$

for $t \geq 0$, $k \in \mathbb{Z}\setminus\{0\}$.

The technical proof is deferred to Appendix 2.A.

Finally, we can estimate solutions to (2.3.12) in Euclidean norm with the help of (2.3.15):

$$\begin{aligned}|y_k(z,t)|^2_2 &= |y_k(z,t)|^2_{\widetilde{P}_k(z,0)}\\ &= |y_k(z,t)|^2_{P^1_{1,k}(z,0)} + |\partial_z\lambda_k(z)|^2|y_k(z,t)|^2_{P^2_{1,k}(z,0)}\\ &\quad + 4|\partial_z\lambda_k(z)|^4|y_k(z,t)|^2_{\widetilde{P}^3_k(z,0)}.\end{aligned}$$

Using (2.3.16) for $P^1_{1,k}(z,0)$ and (2.3.17) for $P^2_{1,k}(z,0)$ allows to estimate the first two terms. For the third term including $\widetilde{P}^3_k(z,0)$, we use (2.3.18) to get

$$\begin{aligned}|y_k(z,t)|^2_2 \leq \Big[&1+6\max\{1,|\partial_z\lambda_k(z)|^2\}(1+k^4t^2)\\ &+4\max\{1,|\partial_z\lambda_k(z)|^4\}(1+|\partial_z^2\lambda_k(z)|^2)146.25(1+k^8t^4)\Big]e^{-2k^2b(z)t}|y_k(z,0)|^2_2\\ \leq \Big[&1+\big(12+585(1+|\partial_z^2\lambda_k(z)|^2)\big)\max\{1,|\partial_z\lambda_k(z)|^4\}\Big]\\ &\times(1+k^8t^4)e^{-2k^2b(z)t}|y_k(z,0)|^2_2\end{aligned} \tag{2.3.19}$$

for $t \geq 0$, $k \in \mathbb{Z} \setminus \{0\}$. Most notably, as $\partial_z a, \partial_z b, \partial_z^2 a, \partial_z^2 b \in L^\infty(\mathbb{R})$, the multiplicative constant

$$\widetilde{\mathscr{C}}_k(z) := 1 + \left(12 + 585(1 + |\partial_z^2 \lambda_k(z)|^2)\right) \max\{1, |\partial_z \lambda_k(z)|^4\}$$

is uniformly bounded in $z \in \mathbb{R}$. This includes the problematic limit $\partial_z \lambda_k(z) \to 0$ in combination with $\partial_z^2 \lambda_k(z) \neq 0$ (defect 2 to defect 1), which is our desired result for Case 3.

Combining Cases 1–3 for $z \in \mathbb{R}$ leads to:

Theorem 2.3.3. *Let $a, b \in C^2(\mathbb{R})$ where $b_0 := \inf_{z\in\mathbb{R}} b(z) > 0$ and $b, \partial_z a, \partial_z b, \partial_z^2 a, \partial_z^2 b \in L^\infty(\mathbb{R})$. Then, there exists a constant $\mathscr{C} > 0$, such that normalized solutions $y(x,z,t) = (u,v,w)^T$ to the system of equations* (2.3.1), (2.3.4) *and* (2.3.11) *with steady state $y^\infty := (1,0,0)^T$ satisfy*

$$\sup_{z\in\mathbb{R}} \| y(\cdot,z,t) - y^\infty \|^2_{L^2(0,2\pi;\mathbb{R}^3)} \leq \mathscr{C}(1+t^4) e^{-2b_0 t} \sup_{z\in\mathbb{R}} \| y(\cdot,z,0) - y^\infty \|^2_{L^2(0,2\pi;\mathbb{R}^3)}$$

for $t \geq 0$.

Proof. Analogous to the first order sensitivity equations, combining the three above cases of defects of $D_k(z)$, leads to a decay estimate uniform in $z \in \mathbb{R}$. Due to the estimates (2.3.13), (2.3.14) and (2.3.19), there exists an $\widetilde{\mathscr{C}} > 0$ independent of $z \in \mathbb{R}$ and $k \in \mathbb{Z} \setminus \{0\}$ such that

$$|y_k(z,t)|_2^2 \leq \widetilde{\mathscr{C}}(1 + k^8 t^4) e^{-k^2 b(z) t} |y_k(z,0)|_2^2, \quad t \geq 0.$$

With Parseval's identity (in analogy to the proof of Theorem 2.3.1) the desired result follows. □

2.3.3 Decay Estimates with Duhamel's Formula

Another method to get decay estimates with sharp rate for sensitivity equations is to use Duhamel's formula instead of the above presented Lyapunov functional method.

In the case of the Fourier transformed linear heat-convection equation (2.3.3), the solution $u_k(z,t)$ is given explicitly as $u_k(z,t) = e^{-k^2\lambda_k(z)t} u_k(z,0)$. We can interpret the Fourier transformed first order sensitivity equation (2.3.5) as an inhomogeneous equation of form

$$\partial_t v_k(z,t) + k^2 \lambda_k(z) v_k(z,t) = g_k(z,t)$$

with $g_k(z,t) := -k^2(\partial_z \lambda_k(z)) u_k(z,t)$. By Duhamel's formula we get

$$\begin{aligned} v_k(z,t) &= e^{-k^2\lambda_k(z)t} v_k(z,0) - k^2(\partial_z \lambda_k(z)) \int_0^t e^{-k^2\lambda_k(z)(t-s)} e^{-k^2\lambda_k(z)s} u_k(z,0)\, ds \\ &= e^{-k^2\lambda_k(z)t} v_k(z,0) - k^2(\partial_z \lambda_k(z)) t e^{-k^2\lambda_k(z)t} u_k(z,0). \end{aligned}$$

By using the solution propagator norm (2.2.26) with $\varepsilon t = k^2 t$, this yields the decay estimate for each Fourier-mode $y_k(z,t) = (u_k, v_k)^T$, $k \in \mathbb{Z} \setminus \{0\}$:

$$|y_k(z,t)|_2^2 \le \frac{4}{3}(1 + k^4 t^2) e^{-2k^2 b(z) t} |y_k(z,0)|_2^2, \quad t \ge 0.$$

By iteration, Duhamel's formula gives a decay estimate for sensitivity equations of *arbitrary* order. A similar method of iteratively deducing decay estimates was used e.g. in [15] (see Theorems 4.1 and 4.2) and [21] (see Theorems 2.1 and 4.4).

2.4 Goldstein–Taylor Model with Uncertain Coefficients

Our starting point is the linear one-dimensional BGK-model for the probability density $f(x,v,t) \ge 0$. This kinetic equation reads

$$\partial_t f + v \partial_x f = M_T(v) \int_{\mathbb{R}} f(x,v,t)\, dv - f(x,v,t), \tag{2.4.1}$$

for $x \in \mathbb{T}^1$, velocities $v \in \mathbb{R}$, and the Maxwell distribution $M_T(v) = (2\pi T)^{-\frac{1}{2}} e^{-\frac{|v|^2}{2T}}$, with given temperature T. Exponential decay towards the equilibrium for this v-continuous model was proved in §4.3 of [1]. We reduce the model drastically and allow only for two discrete velocities $v_\pm = \pm 1$, denoting $f_\pm(x,t) := f(x, \pm 1, t)$. This leads to the system of equations

$$\begin{aligned} \partial_t f_+(x,t) &= -\partial_x f_+(x,t) + \sigma(f_-(x,t) - f_+(x,t)), \\ \partial_t f_-(x,t) &= \partial_x f_-(x,t) - \sigma(f_-(x,t) - f_+(x,t)), \end{aligned}$$

called *Goldstein–Taylor model* with the relaxation coefficient $\sigma > 0$. These equations serve as a toy model that still exhibits many features of (2.4.1). For $\sigma = \frac{1}{2}$, an explicit exponential decay rate of the two velocity model by means of Lyapunov functionals was shown in §1.4 of [11]. The sharp decay estimate was found in [1], §4.1 with a refined functional. We are interested here in augmenting the large-time analysis with a sensitivity analysis.

2.4.1 First Order Parameter Sensitivity Analysis

Similarly to §2.3 we allow the relaxation coefficient to contain uncertainty and denote it by $\sigma(z)$. Throughout §2.4, assume $\sigma \in C^1(\mathbb{R})$, $\partial_z \sigma \in L^\infty(\mathbb{R})$, $\sigma_0 := \inf_{z\in\mathbb{R}} \sigma(z) > 0$, and $\sigma_1 := \sup_{z\in\mathbb{R}} \sigma(z) < 2$. This leads to the following equations for $x \in \mathbb{T}^1$, the parameter $z \in \mathbb{R}$ and $t \ge 0$:

$$\begin{aligned} \partial_t f_+(x,z,t) &= -\partial_x f_+(x,z,t) + \frac{\sigma(z)}{2}(f_-(x,z,t) - f_+(x,z,t)), \\ \partial_t f_-(x,z,t) &= \partial_x f_-(x,z,t) - \frac{\sigma(z)}{2}(f_-(x,z,t) - f_+(x,z,t)), \end{aligned} \tag{2.4.2}$$

with initial condition

$$f_\pm(x,z,0) = f_\pm^0(x,z).$$

For each $z \in \mathbb{R}$ assume $f_\pm^0(\cdot,z,t) \in L_+^1(\mathbb{T}^1)$.

The model is conserving total mass (in time), i.e. $\int_0^{2\pi}[f_+(x,z,t) + f_-(x,z,t)]\,dx = const.$ for all $z \in \mathbb{R}$. The unique normalized steady state for the system is given as $f_+^\infty(z) = f_-^\infty(z) = \frac{1}{2}$. Correspondingly, we shall also assume that the initial total mass is normalized, as $\frac{1}{2\pi}\int_0^{2\pi}[f_+^0(x,z) + f_-^0(x,z)]\,dx = 1$.

To analyze the (linear order) sensitivity of the equation with respect to the relaxation function $\sigma(z)$, we investigate the corresponding family of sensitivity equations for $g_\pm(x,z,t) := \partial_z f_\pm(x,z,t) \in \mathbb{R}$. For each $z \in \mathbb{R}$, they are given as

$$\begin{aligned}
\partial_t g_+(x,z,t) &= -\partial_x g_+(x,z,t) + \frac{\sigma(z)}{2}(g_-(x,z,t) - g_+(x,z,t)) \\
&\quad + \frac{\partial_z\sigma(z)}{2}(f_-(x,z,t) - f_+(x,z,t)), \\
\partial_t g_-(x,z,t) &= \partial_x g_-(x,z,t) - \frac{\sigma(z)}{2}(g_-(x,z,t) - g_+(x,z,t)) \\
&\quad - \frac{\partial_z\sigma(z)}{2}(f_-(x,z,t) - f_+(x,z,t)).
\end{aligned} \tag{2.4.3}$$

For each $z \in \mathbb{R}$, this system (2.4.3) is also conserving total mass (in time), i.e. $\int_0^{2\pi}[g_+(x,z,t) + g_-(x,z,t)]\,dx = const$. Due to the normalization of $f_\pm^0(x,z)$, we have $\frac{1}{2\pi}\int_0^{2\pi}[g_+^0(x,z) + g_-^0(x,z)]\,dx = 0$ with the corresponding steady state given as $g_+^\infty(x,z) = g_-^\infty(x,z) = \partial_z f_\pm^\infty(x,z) = 0$.

To analyze the decay behavior of solutions of the above system (2.4.2)–(2.4.3), we consider the Fourier series $f_\pm(x,z,t) = \sum_{k\in\mathbb{Z}} f_{\pm,k}(z,t)e^{ikx}$. It is convenient to introduce the following linear combinations of the Fourier modes $f_{\pm,k}$, $k \in \mathbb{Z}$:

$$u_k(z,t) := \begin{pmatrix} f_{+,k}(z,t) + f_{-,k}(z,t) \\ f_{+,k}(z,t) - f_{-,k}(z,t) \end{pmatrix}.$$

For each $k \in \mathbb{Z}$ they satisfy the decoupled ODE system

$$\partial_t u_k(z,t) = -\underbrace{\begin{pmatrix} 0 & ik \\ ik & \sigma(z) \end{pmatrix}}_{A_k(z):=} u_k(z,t). \tag{2.4.4}$$

For $k \in \mathbb{Z}$, the matrix $A_k(z)$ has the eigenvalues

$$\lambda_{\pm,k}(z) := \frac{\sigma(z)}{2} \pm i\sqrt{k^2 - \frac{\sigma^2(z)}{4}}.$$

Note that the discriminant is always positive for $k \neq 0$, due to our assumption $0 < \sigma(z) < 2$.

The eigenvectors are given by

$$\hat{v}_{\pm,k}(z) := \left(\tfrac{i\lambda_{\mp}(z)}{k}, \quad 1\right)^T, \, k \in \mathbb{Z} \setminus \{0\}, \qquad \text{and} \qquad \hat{v}_{+,0} := \begin{pmatrix}1\\0\end{pmatrix}, \, \hat{v}_{-,0} := \begin{pmatrix}0\\1\end{pmatrix}.$$

Similarly to (2.4.4), the Fourier modes

$$w_k(z,t) := \begin{pmatrix} g_{+,k}(z,t) + g_{-,k}(z,t) \\ g_{+,k}(z,t) - g_{-,k}(z,t) \end{pmatrix},$$

with $g_{\pm,k} = \partial_z f_{\pm,k}$, satisfy the ODE systems

$$\partial_t w_k(z,t) = -\begin{pmatrix} 0 & ik \\ ik & \sigma(z) \end{pmatrix} w_k(z,t) - \begin{pmatrix} 0 & 0 \\ 0 & \partial_z \sigma(z) \end{pmatrix} u_k(z,t), \quad k \in \mathbb{Z}.$$

Combining the two ODE systems for u_k and w_k leads to the following 4×4-systems describing the first order sensitivity equations for the model (2.4.2) in Fourier space:

$$\underbrace{\partial_t \begin{pmatrix} u_k(z,t) \\ w_k(z,t) \end{pmatrix}}_{y_k(z,t):=} = -\underbrace{\left(\begin{array}{cc|cc} 0 & ik & 0 & 0 \\ ik & \sigma(z) & 0 & 0 \\ \hline 0 & 0 & 0 & ik \\ 0 & \partial_z\sigma(z) & ik & \sigma(z) \end{array}\right)}_{D_k(z):=} \begin{pmatrix} u_k(z,t) \\ w_k(z,t) \end{pmatrix}, \quad k \in \mathbb{Z}. \tag{2.4.5}$$

Due to the block triangular form of the matrix $D_k(z)$, the eigenvalues of $D_k(z)$ are not affected by $\partial_z \sigma(z)$:

$$\lambda_{\pm,k}(z) := \frac{\sigma(z)}{2} \pm i\sqrt{k^2 - \frac{\sigma^2(z)}{4}}, \quad k \in \mathbb{Z},$$

where both eigenvalues have algebraic multiplicity two.

For $k \in \mathbb{Z}$ the matrix $D_k(z)$ is defective, if and only if $\partial_z \sigma(z) \neq 0$: For $k = 0$ only the eigenvalue $\lambda_{+,0}(z) = \sigma(z)$ is defective of order one, and for $k \neq 0$ both eigenvalues $\lambda_{\pm,k}(z)$ are defective of order one.

2.4.2 Sharp Decay Estimates for the Parameter Sensitivity Equations

The reasons for the assumptions from §2.4.1 imposed on $\sigma(z)$ will become evident in the following analysis: The lower bound $\sigma_0 > 0$ is needed in order to get a uniform in $z \in \mathbb{R}$ decay rate. The assumptions $\sigma_1 < 2$ and $\|\partial_z \sigma\|_\infty < \infty$ are necessary for the multiplicative constant in the decay estimate (obtained by Theorem 2.2.8) to be bounded for all $z \in \mathbb{R}$.

The decay rate of each mode $k \in \mathbb{Z}$ is determined by the size of the spectral gap of $D_k(z)$, which we denote by $\mu_k(z) > 0$, and its defectiveness. For $D_0(z)$ the eigenvalues are $\lambda_{+,0}(z) = \sigma(z)$ and $\lambda_{-,0}(z) = 0$. Hence, the spectral gap for the zeroth mode is $\mu_0(z) = \sigma(z)$. The zeroth mode of the steady state in our transformed setting is given as $y_0^\infty = (f_{+,0}^\infty + f_{-,0}^\infty, f_{+,0}^\infty - f_{-,0}^\infty, g_{+,0}^\infty + g_{-,0}^\infty, g_{+,0}^\infty - g_{-,0}^\infty)^T = (1,0,0,0)^T$. This implies that any solution to (2.4.5) for $k = 0$, $z \in \mathbb{R}$ fulfills the decay estimate

$$\left|y_0(z,t) - y_0^\infty\right|_2^2 \le \mathscr{C}_0(z)(1+t^2)e^{-2\sigma(z)t}\left|y_0(z,0) - y_0^\infty\right|_2^2, \quad t \ge 0, \tag{2.4.6}$$

with the constant

$$\mathscr{C}_0(z) = 12 \cdot \max\{2, 1 + |\partial_z \sigma(z)|^2\} \le 12 \cdot \max\{2, 1 + \|\partial_z \sigma\|_\infty^2\}$$

that can be computed in analogy to Case 2 in §2.3.1. Note that Theorem 2.2.8 can only be applied here to the two-dimensional subspace of $\mathbb{C}^4$ that pertains to $\lambda_{+,0}(z) = \sigma(z)$. In the orthogonal subspace corresponding to $\lambda_{-,0} = 0$, we have $y_{-,0}(z,t) = y_{-,0}(z,0) = y_{-,0}^\infty = (1,*,0,*)^T$, where '$*$' denotes the elements of $y_{+,0}$.

For the modes $k \in \mathbb{Z} \setminus \{0\}$ the matrix $D_k(z)$ is positive stable and the spectral gap is independent of k (in contrast to the examples in §2.3 and §2.5):

$$\mu_k(z) := \min\{\mathrm{Re}(\lambda_{+,k}(z)), \mathrm{Re}(\lambda_{-,k}(z))\} = \frac{\sigma(z)}{2}.$$

Moreover, the steady state is given as $y_k^\infty = 0 \in \mathbb{C}^4$.

In the next step, we apply Theorem 2.2.8 to the system (2.4.5) to get a sharp decay estimate for each Fourier mode $k \in \mathbb{Z} \setminus \{0\}$ of type

$$|y_k(z,t) - y_k^\infty|_2^2 \le \mathscr{C}_k(z)2(1+t^2)e^{-\sigma(z)t}|y_k(z,0) - y_k^\infty|_2^2.$$

A summation over the estimates for all Fourier modes will allow us to apply Parseval's identity on the left-hand side. In order to apply it also on the right-hand side one requires a uniform in k and z bound of the multiplicative decay constant $\mathscr{C}_k(z)$. We shall derive this bound now.

For each $k \in \mathbb{Z} \setminus \{0\}$, the matrix $P_k(z,0)$ of Theorem 2.2.8 has to be chosen depending on the defectiveness of the matrix $D_k(z)$, which is determined by the value of $\partial_z \sigma(z)$.

Case 1; $z \in \mathbb{R}$ such that $\partial_z \sigma(z) = 0$:

The matrix $D_k(z)$ is non-defective, and we construct the matrix $P_k(z,0)$ according to (2.2.13): $N = 4$, with $l_n = 1$ for $n \in \{1,\dots,4\}$, $M = 1$, i.e. each n is in Case 1 of §2.2. Choosing $\beta_{n,k} = 1$ leads to

$$P_k(z,0) := v_{1,+,k}^{(0)} \otimes v_{1,+,k}^{(0)} + v_{1,-,k}^{(0)} \otimes v_{1,-,k}^{(0)} + v_{2,+,k}^{(0)} \otimes v_{2,+,k}^{(0)} + v_{2,-,k}^{(0)} \otimes v_{2,-,k}^{(0)},$$

and the eigenvectors of $D_k^H(z)$ corresponding to $\overline{\lambda_{\mp,k}(z)} = \lambda_{\pm,k}(z)$ (i.e. satisfying the equation $D_k^H v_{i,\pm,k}^{(0)} = \lambda_{\pm,k} v_{i,\pm,k}^{(0)}$ for $i = 1,2$) are given as

$$v_{1,\pm,k}^{(0)}(z) = \begin{pmatrix} -\frac{i\lambda_{\mp,k}(z)}{k}, & 1, & 0, & 0 \end{pmatrix}^T, \qquad v_{2,\pm,k}^{(0)}(z) = \begin{pmatrix} 0, & 0, & -\frac{i\lambda_{\mp,k}(z)}{k}, & 1 \end{pmatrix}^T.$$

For each fixed value $\sigma \in [\sigma_0, \sigma_1]$, we have

$$\lim_{k\to+\infty} \tilde{v}_{1,\pm,k}^{(0)}(\sigma) = \begin{pmatrix} \mp 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \lim_{k\to+\infty} \tilde{v}_{2,\pm,k}^{(0)}(\sigma) = \begin{pmatrix} 0 \\ 0 \\ \mp 1 \\ 1 \end{pmatrix},$$

as well as

$$\lim_{k\to-\infty} \tilde{v}_{1,\pm,k}^{(0)}(\sigma) = \begin{pmatrix} \pm 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \lim_{k\to-\infty} \tilde{v}_{2,\pm,k}^{(0)}(\sigma) = \begin{pmatrix} 0 \\ 0 \\ \pm 1 \\ 1 \end{pmatrix},$$

where we used the notations $\tilde{v}_{i,\pm,k}^{(0)}(\sigma(z)) = v_{i,\pm,k}^{(0)}(z)$ for $i = 1,2$. Denoting $\tilde{P}_k(\sigma(z),0) = P_k(z,0)$, it follows that

$$\lim_{|k|\to\infty} \max_{\sigma\in[\sigma_0,\sigma_1]} |\tilde{P}_k(\sigma,0) - 2I|_2 = 0. \tag{2.4.7}$$

For each $\sigma \in [\sigma_0, \sigma_1]$ and $k \neq 0$, the four vectors $\tilde{v}_{1,\pm,k}(\sigma)$ and $\tilde{v}_{2,\pm,k}(\sigma)$ are linearly independent and hence, the matrix $\tilde{P}_k(\sigma,0)$ is positive definite. As all entries of $\tilde{P}_k(\sigma,0)$ are continuous in $\sigma \in [\sigma_0,\sigma_1]$ and the eigenvalues are continuous with respect to the matrix entries, we get

$$\inf_{z\in\mathbb{R}} \lambda_{\min}^{P_k(z,0)} \geq \min_{\sigma\in[\sigma_0,\sigma_1]} \lambda_{\min}^{\tilde{P}_k(\sigma,0)} =: \lambda_{k,\min} > 0.$$

Similarly, we get

$$\sup_{z\in\mathbb{R}} \lambda_{\max}^{P_k(z,0)} \leq \max_{\sigma\in[\sigma_0,\sigma_1]} \lambda_{\max}^{\tilde{P}_k(\sigma,0)} =: \lambda_{k,\max} < \infty.$$

Because of (2.4.7) we have $\lambda_{k,\max}, \lambda_{k,\min} \to 2$ for $|k| \to \infty$, and therefore

$$\lambda_{\min} := \min_{k\in\mathbb{Z}\setminus\{0\}} \lambda_{k,\min} > 0, \qquad \lambda_{\max} := \max_{k\in\mathbb{Z}\setminus\{0\}} \lambda_{k,\max} < \infty.$$

We summarize Case 1: For all $z \in \mathbb{R}$ such that $D_k(z)$ is non-defective, Theorem 2.2.8 yields the decay estimate

$$\left| y_k(z,t) - y_k^\infty \right|_2^2 \leq 2\mathscr{C}_k(z) e^{-\sigma(z)t} |y_k(z,0) - y_k^\infty|_2^2, \tag{2.4.8}$$

with a uniform bound for the constants $\mathscr{C}_k(z)$ (defined in (2.2.22)), i.e.

$$0 < \mathscr{C}_k(z) = (\lambda_{\min}^{P_k(z,0)})^{-1} \lambda_{\max}^{P_k(z,0)} \leq (\lambda_{\min})^{-1} \lambda_{\max} =: \mathscr{C} < \infty,$$

for $z \in \mathbb{R}$ and $k \neq 0$.

Case 2; $z \in \mathbb{R}$ such that $\partial_z \sigma(z) \neq 0$:

The two eigenvalues $\lambda_{\pm,k}(z)$ of $D_k(z)$ are both defective. The eigenvectors and generalized eigenvectors of $D_k^H(z)$ corresponding to $\overline{\lambda}_{\mp,k}(z) = \lambda_{\pm,k}(z)$ (i.e. the generalized eigenvectors satisfy $D_k^H v_{\pm,k}^{(1)} = \lambda_{\pm,k} v_{\pm,k}^{(1)} + v_{\pm,k}^{(0)}$) are given as

$$v_{\pm,k}^{(0)}(z) = \Big(-\tfrac{i\lambda_{\mp,k}(z)}{k}, \quad 1, \quad 0, \quad 0\Big)^T,$$
$$v_{\pm,k}^{(1)}(z) = \Big(\tfrac{i\lambda_{\mp,k}^2(z)}{2k^3}, \quad \tfrac{\lambda_{\mp,k}(z)}{2k^2}, \quad \tfrac{-i\lambda_{\mp,k}(z)}{\sigma_z(z)k}(1 - \tfrac{\lambda_{\mp,k}^2(z)}{k^2}), \quad \tfrac{1}{\sigma_z(z)}(1 - \tfrac{\lambda_{\mp,k}^2(z)}{k^2})\Big)^T.$$

Now we construct the matrix $P_k(z,0)$ according to (2.2.13): $N = 2$, with $l_+ = l_- = M = 2$, i.e. both $n \in \{+,-\}$ are in Case 3 of §2.2. Choosing $\beta_{\pm,k}^1 = 1$ and $\beta_{\pm,k}^2 = \frac{(\sigma_z(z))^2}{4}$, leads to

$$\begin{aligned}P_k(z,0) := v_{+,k}^{(0)} \otimes v_{+,k}^{(0)} + \frac{(\sigma_z(z))^2}{4} v_{+,k}^{(1)} \otimes v_{+,k}^{(1)} \\ + v_{-,k}^{(0)} \otimes v_{-,k}^{(0)} + \frac{(\sigma_z(z))^2}{4} v_{-,k}^{(1)} \otimes v_{-,k}^{(1)}.\end{aligned}$$

As mentioned in §2.3 the specific choice of weights $\beta_\pm^m$ is crucial in order to get a uniformly bounded constant $\mathscr{C}_k(z)$ in the *non-defective limit* $\partial_z\sigma(z) \to 0$.

Abbreviating $L := \|\partial_z\sigma\|_\infty$, for each value $(\sigma, \sigma_z) \in [\sigma_0, \sigma_1] \times [-L, L]$ (with notation in analogy to Case 1), we have

$$\lim_{k\to+\infty} \tilde{v}_{\pm,k}^{(0)}(\sigma,\sigma_z) = \begin{pmatrix} \mp 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \lim_{k\to+\infty} \frac{\sigma_z}{2} \tilde{v}_{\pm,k}^{(1)}(\sigma,\sigma_z) = \begin{pmatrix} 0 \\ 0 \\ \mp 1 \\ 1 \end{pmatrix},$$

and

$$\lim_{k\to-\infty} \tilde{v}_{\pm,k}^{(0)}(\sigma,\sigma_z) = \begin{pmatrix} \pm 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \lim_{k\to-\infty} \frac{\sigma_z}{2} \tilde{v}_{\pm,k}^{(1)}(\sigma,\sigma_z) = \begin{pmatrix} 0 \\ 0 \\ \pm 1 \\ 1 \end{pmatrix}.$$

It follows that

$$\lim_{|k|\to\infty} \max_{(\sigma,\sigma_z)\in[\sigma_0,\sigma_1]\times[-L,L]} |\tilde{P}_k(\sigma,\sigma_z,0) - 2I|_2 = 0. \tag{2.4.9}$$

For each $(\sigma, \sigma_z) \in [\sigma_0, \sigma_1] \times [-L, L]$, the four vectors $\tilde{v}_{\pm,k}^{(0)}(\sigma,\sigma_z)$ and $\sigma_z \tilde{v}_{\pm,k}^{(1)}(\sigma,\sigma_z)$ are linearly independent. This can be checked by considering the last two components of

$\sigma_z \tilde{v}^{(1)}_{\pm,k}(\sigma,\sigma_z)$: They have the same form as the last two components of $\tilde{v}^{(0)}_{2,\pm,k}(\sigma,\sigma_z)$ from Case 1 above, up to a multiplicative factor that is non-zero for all $k \neq 0$ due to $\sigma_1 < 2$.

Hence, the matrix $\tilde{P}_k(\sigma,\sigma_z,0)$ is positive definite on $[\sigma_0,\sigma_1] \times [-L,L]$. Due to the specific choice of $\beta^2_{\pm,k} = \frac{(\sigma_z(z))^2}{4}$ all entries of $\tilde{P}_k(\sigma,\sigma_z,0)$ are continuous with respect to (σ,σ_z). With the same argument as in Case 1, we get

$$\inf_{z\in\mathbb{R}} \lambda^{P_k(z,0)}_{\min} \geq \min_{(\sigma,\sigma_z)\in[\sigma_0,\sigma_1]\times[-L,L]} \lambda^{\tilde{P}_k(\sigma,\sigma_z,0)}_{\min} =: \lambda_{k,\min} > 0,$$

and

$$\sup_{z\in\mathbb{R}} \lambda^{P_k(z,0)}_{\max} \leq \max_{(\sigma,\sigma_z)\in[\sigma_0,\sigma_1]\times[-L,L]} \lambda^{\tilde{P}_k(\sigma,\sigma_z,0)}_{\max} := \lambda_{k,\max} < \infty.$$

The limit (2.4.9) implies

$$\lambda_{\min} := \min_{k\in\mathbb{Z}\setminus\{0\}} \lambda_{k,\min} > 0, \qquad \lambda_{\max} := \max_{k\in\mathbb{Z}\setminus\{0\}} \lambda_{k,\max} < \infty.$$

We summarize Case 2: For all $z \in \mathbb{R}$ such that $D_k(z)$ is defective, Theorem 2.2.8 yields the decay estimate

$$\left|y_k(z,t) - y^\infty_k\right|^2_2 \leq \mathscr{C}_k(z)(1+t^2)e^{-\sigma(z)t}|y_k(z,0) - y^\infty_k|^2_2. \tag{2.4.10}$$

Here, the constants $\mathscr{C}_k(z)$ from (2.2.22), using $c_2 = 6$, are uniformly bounded since $\partial_z\sigma \in L^\infty(\mathbb{R})$:

$$\begin{aligned} 0 < \mathscr{C}_k(z) &= 12\cdot(\lambda^{P_k(z,0)}_{\min})^{-1}\lambda^{P_k(z,0)}_{\max}\left(1 + \frac{\frac{(\sigma_z(z))^2}{4}}{\min\left\{1, \frac{(\sigma_z(z))^2}{4}\right\}}\right) \\ &\leq 12\cdot(\lambda_{\min})^{-1}\lambda_{\max}\max\{2, 1 + \frac{\|\partial_z\sigma\|^2_\infty}{4}\} =: \mathscr{C} < \infty \end{aligned}$$

for all $z \in \mathbb{R}$ and $k \neq 0$.

Now, we have all the necessary ingredients to estimate the decay of solutions to the system (2.4.2)–(2.4.3). Denote

$$\begin{aligned} \Phi &:= \begin{pmatrix} f_+, & f_-, & g_+, & g_- \end{pmatrix}^T, & \Phi^\infty &:= \begin{pmatrix} \frac{1}{2}, & \frac{1}{2}, & 0 & 0 \end{pmatrix}^T, \\ y &:= \begin{pmatrix} f_+ + f_-, & f_+ - f_-, & g_+ + g_-, & g_+ - g_- \end{pmatrix}^T, & y^\infty &:= \begin{pmatrix} 1, & 0, & 0, & 0 \end{pmatrix}^T. \end{aligned}$$

Theorem 2.4.1. *Let* $\sigma \in C^1(\mathbb{R})$ *where* $\sigma_0 := \inf_{z\in\mathbb{R}}\sigma(z) > 0$, $\sigma_1 := \sup_{z\in\mathbb{R}}\sigma(z) < 2$ *and* $\partial_z\sigma \in L^\infty(\mathbb{R})$. *Then, there exists a constant* $\mathscr{C} > 0$, *such that normalized solutions* $\Phi(x,z,t)$ *of the system* (2.4.2)–(2.4.3) *satisfy*

$$\sup_{z\in\mathbb{R}} \|\Phi(\cdot,z,t) - \Phi^\infty\|^2_{L^2(0,2\pi;\mathbb{R}^4)} \leq \mathscr{C}(1+t^2)e^{-\sigma_0 t} \sup_{z\in\mathbb{R}} \|\Phi(\cdot,z,0) - \Phi^\infty\|^2_{L^2(0,2\pi;\mathbb{R}^4)}$$

for $t \geq 0$.

Proof. Applying Parseval's identity and the decay estimates (2.4.6), (2.4.8) and (2.4.10), where the constants $\mathscr{C}_k(z)$ are uniformly bounded in $z \in \mathbb{R}$, $k \in \mathbb{Z}$, one obtains

$$\begin{aligned}\|\Phi(\cdot,z,t)-\Phi^\infty\|^2_{L^2(0,2\pi;\mathbb{R}^4)} &= \frac{1}{2}\|y(\cdot,z,t)-y^\infty\|^2_{L^2(0,2\pi;\mathbb{R}^4)}\\ &= \frac{1}{4\pi}\sum_{k\in\mathbb{Z}}|y_k(z,t)-y_k^\infty|_2^2\\ &\le \frac{1}{4\pi}\sum_{k\in\mathbb{Z}}2\mathscr{C}_k(z)(1+t^2)e^{-\sigma(z)t}|y_k(z,0)-y_k^\infty|_2^2\\ &\le \mathscr{C}(1+t^2)e^{-\sigma_0 t}\|\Phi(\cdot,z,0)-\Phi^\infty\|^2_{L^2(0,2\pi;\mathbb{R}^4)}\end{aligned}$$

for all $t \ge 0$. □

2.5 Fokker–Planck Equations with Uncertain Coefficients

In this section we consider the Fokker–Planck equation (FPE) with the spatial variable $x \in \mathbb{R}$,

$$\partial_t f(x,z,t) = \partial_x[\partial_x f(x,z,t) + a(z)xf(x,z,t)] =: L(z)f(x,z,t), \quad t \ge 0, \tag{2.5.1}$$

with initial condition

$$f(x,z,0) = f^0(x,z).$$

The drift parameter $a(z)$ depends only on the uncertainty parameter $z \in \mathbb{R}$ and we assume $a \in C^1(\mathbb{R})$ with $a_0 := \inf_{z\in\mathbb{R}} a(z) > 0$. As in §2.3–§2.4, we want to analyze the sensitivity of the decay for solutions to the steady state w.r.t. $z \in \mathbb{R}$. Contrary to the previous examples, *the steady state here also depends on z.*

For each $z \in \mathbb{R}$ the unique normalized steady state of the equation, i.e. $L(z)f^\infty(x,z) = 0$ with $\int_{\mathbb{R}} f^\infty(x,z)dx = 1$, is given as

$$f^\infty(x,z) = \sqrt{\frac{a(z)}{2\pi}}e^{-\frac{x^2}{2}a(z)}.$$

Denoting $g(x,z,t) := \partial_z f(x,z,t)$, the first order linear sensitivity equation is given as

$$\partial_t g(x,z,t) = L(z)g(x,z,t) + a_z(z)[x\partial_x f(x,z,t) + f(x,z,t)], \quad x \in \mathbb{R}, t \ge 0, \tag{2.5.2}$$

for each fixed $z \in \mathbb{R}$. The corresponding steady state is given as $g^\infty(x,z) := \partial_z f^\infty(x,z)$ which satisfies $\int_{\mathbb{R}} g^\infty(x,z)dx = 0$.

A direct approach to estimate the decay of g via Duhamel's formula (cf. §2.3.3) is not (easily) feasible: While a decay estimates with sharp rate for $f(x,z,t)$ is available, the decay behavior of the term $x\partial_x f(x,z,t)$ is not immediate. In [17], the Duhamel approach was taken to obtain decay estimates for nonlinear Vlasov–Fokker–Planck equations, but those estimates were not sharp.

We choose to expand $f(x,z,t)$ and $g(x,z,t)$ into eigenfunctions of $L(z)$. This allows us to use a recursive relation of the eigenfunctions $h_k(x,z)$ and $x\partial_x h_k(x,z)$ for $k \in \mathbb{N}_0$.

2.5.1 Eigenfunctions of the FP-Operator $L(z)$

The normalized eigenfunctions of $L(z)$ on the weighted space $L^2((f^\infty)^{-1})$ (with the inner product $\langle f,g\rangle_{L^2((f^\infty)^{-1})} = \int_{\mathbb{R}} fg(f^\infty)^{-1}dx$) are rescaled Hermite functions. The *probabilists' Hermite polynomials* are defined as

$$H_k(x) := (-1)^k e^{\frac{x^2}{2}} \frac{d^k}{dx^k} e^{-\frac{x^2}{2}}, \quad k \in \mathbb{N}_0, x \in \mathbb{R},$$

and satisfy the recursion

$$H_k'(x) = xH_k(x) - H_{k+1}(x) = kH_{k-1}(x), \quad k \in \mathbb{N}, x \in \mathbb{R}. \tag{2.5.3}$$

The *Hermite functions* are given as

$$\tilde{h}_k(x) := \frac{1}{\sqrt{2\pi k!}} H_k(x) e^{-\frac{x^2}{2}}, \quad k \in \mathbb{N}_0, x \in \mathbb{R}, \tag{2.5.4}$$

satisfying (due to (2.5.3)) $\partial_x \tilde{h}_k(x) = -\sqrt{k+1}\,\tilde{h}_{k+1}(x)$. We further denote the *Hermite functions rescaled by* $a(z)$ as

$$h_k(x,z) := \sqrt{a(z)}\,\tilde{h}_k(x\sqrt{a(z)}). \tag{2.5.5}$$

This rescaling is chosen such that the Hermite functions $h_k(z,\cdot)$ are normalized in $L^2((f^\infty)^{-1})$ for all $k \in \mathbb{N}_0$. Notice that $f^\infty(x,z) = h_0(x,z)$ and $g^\infty(x,z) = -\frac{\partial_z a(z)}{\sqrt{2}a(z)} h_2(x,z)$.

For later use we note that, due to (2.5.3), the rescaled Hermite functions satisfy

$$xh_k(x,z) = \frac{1}{\sqrt{a(z)}}[\sqrt{k+1}\,h_{k+1}(x,z) + \sqrt{k}\,h_{k-1}(x,z)], \tag{2.5.6}$$

for $k \in \mathbb{N}, x, z \in \mathbb{R}$. Using (2.5.3) again, this implies

$$x\partial_x h_k(x,z) = -\sqrt{k+1}[\sqrt{k+2}\,h_{k+2}(x,z) + \sqrt{k+1}\,h_k(x)], \tag{2.5.7}$$

for $k \in \mathbb{N}_0, x, z \in \mathbb{R}$.

The spectrum of $L(z)$, with $z \in \mathbb{R}$ fixed, is given as

$$\sigma(L(z)) = \{-a(z)k \mid k \in \mathbb{N}_0\}.$$

An orthonormal basis of eigenfunctions for the FP-operator $L(z)$ on $L^2((f^\infty)^{-1})$ is given by the rescaled Hermite functions defined in (2.5.5) (see e.g. [24], §5.5.1, §10.1.4), i.e. for $z \in \mathbb{R}$ fixed:

$$L^2((f^\infty)^{-1}) = \bigoplus_{k\in\mathbb{N}_0} \operatorname{span}\{h_k(\cdot,z)\}, \qquad L(z)h_k(\cdot,z) = -a(z)k\,h_k(\cdot,z).$$

2.5.2 Sharp Decay Estimate for the Parameter Sensitivity Equations

Let us assume $f^0(\cdot,z), g^0(\cdot,z) \in L^2((f^\infty)^{-1})$ where $f(x,z,t)$ is a probability density, $\int_{\mathbb{R}} f^0(x,z)dx = 1$, and $g^0(x,z)$ does not carry any mass, i.e. $0 = \int_{\mathbb{R}} g^0(x,z)dx = (g^0(\cdot,z), h_0(\cdot,z))_{L^2((f^\infty)^{-1})}$.
The eigenfunction expansions for the corresponding solutions $f(x,z,t)$ of (2.5.1) and $g(x,z,t)$ of (2.5.2) are given as

$$f(x,z,t) = \sum_{k=0}^{\infty} f_k(z,t)h_k(x,z), \qquad g(x,z,t) = \sum_{k=1}^{\infty} g_k(z,t)h_k(x,z),$$

for $x,z \in \mathbb{R}$ and $t \geq 0$. Due to (2.5.1), each eigenmode evolves as

$$\partial_t f_k(z,t) = -a(z)kf_k(z,t), \quad k \in \mathbb{N}_0, \tag{2.5.8}$$

and hence $f_0(z,t) = 1$. Plugging the eigenfunction expansion for g into (2.5.2) leads to

$$\begin{aligned}\sum_{k=1}^{\infty} \partial_t g_k(z,t)h_k(x) = -a(z)&\sum_{k=1}^{\infty} kg_k(z,t)h_k(x,z)\\ &+ a_z(z)\sum_{k=0}^{\infty} f_k(z,t)\left[h_k(x,z) + x\partial_x h_k(x,z)\right].\end{aligned}$$

Applying identity (2.5.7) gives

$$\begin{aligned}&\sum_{k=1}^{\infty} \partial_t g_k(z,t)h_k(x)\\ &\quad= -a(z)\sum_{k=1}^{\infty} kg_k(z,t)h_k(x,z) + a_z(z)\sum_{k=0}^{\infty} f_k(z,t)h_k(x,z)\\ &\qquad + a_z(z)\sum_{k=0}^{\infty} -(k+1)f_k(z,t)h_k(x,z) - \sqrt{(k+1)(k+2)}f_k(z,t)h_{k+2}(x,z)\\ &\quad= -a(z)\sum_{k=1}^{\infty} kg_k(z,t)h_k(x,z) - a_z(z)f_1(z,t)h_1(x,z)\\ &\qquad - a_z(z)\sum_{k=2}^{\infty}[kf_k(z,t) + \sqrt{k(k-1)}f_{k-2}(z,t)]h_k(x,z).\end{aligned}$$

Separating the eigenmodes then yields

$$\begin{aligned}
&\partial_t g_1(z,t) = -a(z) g_1(z,t) - a_z(z) f_1(z,t),\\
&\partial_t g_k(z,t) = -k a(z) g_k(z,t) - a_z(z)\big[k f_k(z,t) + \sqrt{k(k-1)} f_{k-2}(z,t)\big], \quad k \geq 2.
\end{aligned} \tag{2.5.9}$$

In contrast to f, the kth modes of g do not decouple for $k \geq 1$. They are rather coupled as the pair (f_1, g_1), respectively the triples (f_{k-2}, f_k, g_k) for $k \geq 2$. For $k = 1$ the evolution equation for $f_1(z,t)$, $g_1(z,t)$ can be written as the ODE system

$$\partial_t \underbrace{\begin{pmatrix} f_1 \\ g_1 \end{pmatrix}}_{y_1(z,t):=} = -a(z) \underbrace{\begin{pmatrix} 1 & 0 \\ \alpha(z) & 1 \end{pmatrix}}_{C_1(z):=} \begin{pmatrix} f_1 \\ g_1 \end{pmatrix}, \tag{2.5.10}$$

for $z \in \mathbb{R}$, $t \geq 0$, with the notation

$$\alpha(z) := \frac{a_z(z)}{a(z)}.$$

For $k = 2$, equation (2.5.9) can be written as

$$\partial_t \widetilde{g}_2(z,t) = -2a(z)[\widetilde{g}_2(z,t) + \alpha(z) f_2(z,t)],$$

with $\widetilde{g}_2(z,t) := g_2(z,t) + \frac{\alpha(z)}{\sqrt{2}}$, since $f_0(z,t) = \int_{\mathbb{R}} f(x,z,t)dx \equiv 1$. The corresponding system of equations is given as

$$\partial_t \underbrace{\begin{pmatrix} f_2 \\ \widetilde{g}_2 \end{pmatrix}}_{y_2(z,t):=} = -2a(z) \underbrace{\begin{pmatrix} 1 & 0 \\ \alpha(z) & 1 \end{pmatrix}}_{C_2(z):=} \begin{pmatrix} f_2 \\ \widetilde{g}_2 \end{pmatrix}, \tag{2.5.11}$$

for $z \in \mathbb{R}$, $t \geq 0$. Since the matrices $C_1(z) = C_2(z)$ are defective, if and only if $a_z(z) \neq 0$, we shall now distinguish these cases.

Case $k = 1,2$ and $z \in \mathbb{R}$ such that $a_z(z) = 0$:

The matrices $C_1(z) = C_2(z)$ are diagonal and the solutions of the eigenmodes $k = 1,2$ are given explicitly as

$$y_k(z,t) = e^{-ka(z)t} y_k(z,0), \quad t \geq 0, k = 1,2.$$

The decay estimate

$$|y_k(z,t)|_2^2 \leq e^{-2ka(z)t} |y_k(z,0)|_2^2, \quad t \geq 0, k = 1,2, \tag{2.5.12}$$

follows.

Case $k = 1,2$ and $z \in \mathbb{R}$ such that $a_z(z) \neq 0$:

The matrices $C_1(z) = C_2(z)$ are defective of order 1 and we apply Theorem 2.2.8 to get a uniform-in-z decay estimate. The construction of the matrices $P_1(z,t) = P_2(z,t)$ resembles Example 2.2.11 (with $\varepsilon = \alpha(z)$ and rescaling $t \mapsto ka(z)t$). It yields the decay estimate

$$|y_k(z,t)|_2^2 \leq \mathscr{C}_k(z)(1 + k^2 a(z)^2 t^2) e^{-2ka(z)t} |y_k(z,0)|_2^2, \quad k = 1,2, \tag{2.5.13}$$

with the uniform in $z \in \mathbb{R}$ bounded constant (for $k = 1,2$)

$$\mathscr{C}_k(z) = 12 \cdot \max\{2, 1 + \alpha(z)^2\} \leq 12 \max\left\{2, 1 + \frac{\|a_z\|_\infty^2}{a_0^2}\right\} =: \mathscr{C}_{1,2}. \tag{2.5.14}$$

Notice that by definition of $y_2(z,t)$ the decay $y_2(z,t) \xrightarrow{t\to\infty} 0$ implies $g_2(z,t) \xrightarrow{t\to\infty} -\frac{\alpha(z)}{\sqrt{2}} = (g^\infty(\cdot,z), h_2(\cdot,z))_{L^2((f^\infty)^{-1})}$. This concludes the analysis for the modes $k = 0,1,2$.

For our goal to get a decay estimate with sharp uniform-in-z decay rate of the system (2.5.1)–(2.5.2) as formulated in Theorem 2.5.2 below, it is important to get a "precise" decay estimate for the modes $k = 1,3$. Only these two modes have the spectral gap $a(z)$ of the system of equations (2.5.1)–(2.5.2). The other modes have larger spectral gaps and decay much faster. On the level of the modal equations for $k \geq 4$ all we need are "sufficient" decay estimates, namely rates at least as good as the ones of the modes $k = 1,3$. This is in contrast to §2.4 where every Fourier mode has the spectral gap $\frac{\sigma(z)}{2}$ of the sensitivity equations and needs "precise" treatment.

For the modes $k \geq 3$, the equation for $g_k(z,t)$ corresponds to

$$\underbrace{\partial_t \begin{pmatrix} f_{k-2} \\ f_k \\ g_k \end{pmatrix}}_{y_k(z,t):=} = -ka(z) \underbrace{\begin{pmatrix} \frac{k-2}{k} & 0 & 0 \\ 0 & 1 & 0 \\ \gamma(k)\alpha(z) & \alpha(z) & 1 \end{pmatrix}}_{C_k(z):=} \begin{pmatrix} f_{k-2} \\ f_k \\ g_k \end{pmatrix}, \tag{2.5.15}$$

for $z \in \mathbb{R}$, $t \geq 0$, denoting

$$\gamma(k) := \sqrt{\frac{k-1}{k}} \in [\sqrt{\tfrac{2}{3}}, 1).$$

For each $k \geq 3$, the eigenvalues of $C_k(z)$ are $\lambda_{1,k} = \frac{k-2}{k}$ and $\lambda_{2,k} = 1$, where $\lambda_{2,k}$ is defective of order 1, if and only if $a_z(z) \neq 0$. The (non-defective) spectral gap of $C_k(z)$ is given as

$$\mu_k = \frac{k-2}{k}, \quad k \geq 3.$$

Case $k=3$ and $z\in\mathbb{R}$ such that $a_z(z)=0$:

The matrix $C_3^H(z)$ is diagonal and the solutions for the eigenmodes are given explicitly as

$$\begin{aligned}&f_1(z,t)=e^{-a(z)t}f_1(z,0),\\ &f_3(z,t)=e^{-3a(z)t}f_3(z,0),\qquad g_3(z,t)=e^{-3a(z)t}g_3(z,0),\quad t\geq 0.\end{aligned}$$

The decay estimate

$$|y_3(z,t)|_2^2\leq e^{-2a(z)t}|y_3(z,0)|_2^2,\quad t\geq 0 \tag{2.5.16}$$

follows.

Case $k=3$ and $z\in\mathbb{R}$ such that $a_z(z)\neq 0$:

In this case the eigenvalue $\lambda_{2,3}(z)=1$ is defective, but this eigenvalue does not correspond to the spectral gap $\mu_3=\frac{1}{3}$.

The matrix $C_3(z)$ corresponds to two Jordan blocks. In notation from §2.2 this means: $N=2$, $l_1=1$, $l_2=2$ and $M=1$. We use Theorem 2.2.8 with the modification of Remark 2.2.13 for $n_2=2$. The (generalized) eigenvectors of $C_3^H(z)$ are given as

$$\begin{aligned}&v_{1,3}^{(0)}=(1,0,0)^T,\\ &v_{2,3}^{(0)}(z)=(0,\ \alpha(z),\ 0)^T,\qquad\qquad v_{2,3}^{(1)}(z)=(\sqrt{\tfrac{3}{2}}\alpha(z),\,0,\,1)^T.\end{aligned}$$

With the modifications described in Remark 2.2.13, the matrix $\widetilde{P}_3(z,t)$ is constructed with three arbitrary weights: We choose them as $\beta_{1,3}^1(z)=1$, $\beta_{2,3}^1(z)=\alpha(z)^{-2}$ and $\beta_{2,3}^2(z)=1$, which leads to

$$\widetilde{P}_3(z,0)=\begin{pmatrix}1+\frac{3}{2}\alpha(z)^2 & 0 & \sqrt{\frac{3}{2}}\alpha(z)\\ 0 & 1 & 0\\ \sqrt{\frac{3}{2}}\alpha(z) & 0 & 1\end{pmatrix}.$$

Then, Remark 2.2.13 (with the rescaling $t\mapsto 3a(z)t$) leads to the decay estimate

$$|y_3(z,t)|_2^2\leq\widetilde{\mathscr{C}}_3(z)e^{-2a(z)t}|y_3(z,0)|_2^2,\quad t\geq 0,$$

with the constant

$$\widetilde{\mathscr{C}}_3(z)=\frac{\lambda_{\max}^{\widetilde{P}_3(z,0)}}{\lambda_{\min}^{\widetilde{P}_3(z,0)}}12\cdot\max\{2,1+\alpha(z)^2\}. \tag{2.5.17}$$

Denoting $\delta(z) := 1 + \frac{3}{4}\alpha(z)$, the eigenvalues of $\widetilde{P}_3(z,0)$ are given as

$$\lambda_{1,2}^{\widetilde{P}_3(z,0)} = \delta(z) \pm \sqrt{\delta(z)^2 - 1}, \qquad \lambda_3^{\widetilde{P}_3(z,0)} = 1,$$

with $\lambda_{\max}^{\widetilde{P}_3(z,0)} = \lambda_1^{\widetilde{P}_3(z,0)}$ and $\lambda_{\min}^{\widetilde{P}_3(z,0)} = \lambda_2^{\widetilde{P}_3(z,0)}$. Since

$$\begin{aligned}\frac{\lambda_{\max}^{\widetilde{P}_3(z,0)}}{\lambda_{\min}^{\widetilde{P}_3(z,0)}} &= 2\delta(z)^2 - 1 + 2\delta(z)\sqrt{\delta(z)^2 - 1}\\ &\le 4\delta(z)^2 - 1 = 3 + \frac{9}{4}\alpha(z)^4 + 6\alpha(z)^2,\end{aligned}$$

the constant is uniformly bounded in z by

$$\widetilde{\mathscr{C}_3}(z) \le (6 + \frac{21}{4}\frac{\|a_z\|_\infty^4}{a_0^4})12 \cdot \max\{2, 1 + \frac{\|a_z\|_\infty^2}{a_0^2}\} =: \mathscr{C}_3.$$

We arrive at

$$|y_3(z,t)|_2^2 \le \mathscr{C}_3 e^{-2a(z)t} |y_3(z,0)|_2^2 \tag{2.5.18}$$

for $t \ge 0$.

Case $k \ge 4$ and $z \in \mathbb{R}$:

In this case the equations for the kth mode (2.5.15) does not correspond to the spectral gap $a(z)$ of the system (2.5.1)–(2.5.2). In fact, the exponential decay rate of $|y_k(z,t)|_2^2$, $k \ge 4$ is at least $4a(z)$, which is double the rate of the slowest modes $k = 1,3$. Thus, there is more freedom of choice for the matrix $P_k(z)$ for $k \ge 4$. This is important in order to get a uniform in z and k estimate for $k \ge 4$. The additional difficulty compared to $k = 3$ is the uniform bound for $k \to \infty$. Indeed, even Remark 2.2.13 would not give a matrix $\widetilde{P}_k(z,0)$ with uniform condition number for fixed z and $k \to \infty$, and therefore a decay estimate constant $\widetilde{\mathscr{C}_k}(z)$ that is unbounded in k.

The following lemma builds on the fact that the Euclidean norm, i.e. using $\widetilde{P} = I$, yields already a "sufficient" decay estimate as long as $|\alpha(z)|$ is small enough. An appropriate rescaling of the third coordinate via a modified norm does the trick for all $z \in \mathbb{R}$.

Lemma 2.5.1. *For $k \ge 4$ and $z \in \mathbb{R}$ solutions to* (2.5.15) *satisfy the decay estimate*

$$|y_k(z,t)|_2^2 \le \mathscr{C}_{\ge 4} e^{-2a(z)t} |y_k(z,0)|_2^2, \quad t \ge 0, \tag{2.5.19}$$

with the constant $\mathscr{C}_{\ge 4} := 2(1 + \frac{\|\partial_z a\|_\infty^4}{a_0^4})$.

The elementary but technical proof is deferred to Appendix 2.A.

Combining the above five cases for $k \in \mathbb{N}_0$ and $z \in \mathbb{R}$ leads to the desired uniform-in-z decay estimate for arbitrary initial conditions on $L^2((f^\infty)^{-1}) \times L^2((f^\infty)^{-1})$ with sharp rate:

Theorem 2.5.2. *Let $a \in C^1(\mathbb{R})$ where $a_0 := \inf_{z\in\mathbb{R}} a(z) > 0$ and $\partial_z a \in L^\infty(\mathbb{R})$. Then, there exists a constant $\mathscr{C} > 0$, such that normalized solutions $\Phi(x,z,t) = (f,g)^T$ of the system* (2.5.1)–(2.5.2) *with steady state $\Phi^\infty(x,z) := (f^\infty, g^\infty)^T$ satisfy*

$$\begin{aligned}\sup_{z\in\mathbb{R}} &\|\Phi(\cdot,z,t)-\Phi^\infty(\cdot,z)\|^2_{L^2((f^\infty)^{-1})}\\ &\le \mathscr{C}(1+t^2)e^{-2a_0 t}\sup_{z\in\mathbb{R}}\|\Phi(\cdot,z,0)-\Phi^\infty(\cdot,z)\|^2_{L^2((f^\infty)^{-1})}\end{aligned}$$

for $t \ge 0$ with an explicit constant $\mathscr{C} > 0$ only depending on a_0 and $\|\partial_z a\|_\infty$, as given in (2.5.20) *below.*

Proof. With Parseval's identity and the collected decay estimates (2.5.12), (2.5.13), (2.5.16), (2.5.18) and (2.5.19), we obtain

$$\begin{aligned}\|\Phi(\cdot,z,t)-&\Phi^\infty(\cdot,z)\|^2_{L^2((f^\infty)^{-1})}\\ &= \sum_{k=1}^\infty |f_k(z,t)|^2 + |g_1(z,t)|^2 + \left|g_2(z,t)+\frac{\alpha(z)}{\sqrt{2}}\right|^2 + \sum_{k=3}^\infty |g_k(z,t)|^2\\ &\le \sum_{k=1}^\infty |y_k(z,t)|_2^2\\ &\le \mathscr{C}_{1,2}\sum_{k=1}^2 (1+k^2a(z)^2t^2)e^{-2ka(z)t}|y_k(z,0)|_2^2 + \mathscr{C}_3 e^{-2a(z)t}|y_3(z,0)|_2^2\\ &\quad + \mathscr{C}_{\ge 4}\sum_{k=4}^\infty e^{-2a(z)t}|y_k(z,0)|_2^2\\ &\le (\mathscr{C}_{1,2}+\mathscr{C}_3+\mathscr{C}_{\ge4})(1+a(z)^2t^2)e^{-2a(z)t}\left(\sup_{z\in\mathbb{R}}\sum_{k=1}^\infty |y_k(z,0)|_2^2\right).\end{aligned}$$

Using the fact that $(1+a(z)^2t^2)e^{-2a(z)t} \le (1+a_0^2t^2)e^{-2a_0t}$ for $t \ge 0$, $z \in \mathbb{R}$ leads to

$$\begin{aligned}\|\Phi(\cdot,z,t)-&\Phi^\infty(\cdot,z)\|^2_{L^2((f^\infty)^{-1})}\\ &\le \mathscr{C}(1+t^2)e^{-2a_0t}\sup_{z\in\mathbb{R}}\|\Phi(\cdot,z,0)-\Phi^\infty(\cdot,z)\|^2_{L^2((f^\infty)^{-1})}\end{aligned}$$

for each fixed $z \in \mathbb{R}$, $t \ge 0$ with the constant

$$\begin{aligned}\mathscr{C} &:= 2\max\{1,a_0^2\}(\mathscr{C}_{1,2}+\mathscr{C}_3+\mathscr{C}_{\ge4})\\ &= 2\max\{1,a_0^2\}\left[12\cdot\max\{2,1+\frac{\|a_z\|^2_\infty}{a_0^2}\}(7+\frac{21}{4}\frac{\|a_z\|^4_\infty}{a_0^4}) + 2(1+\frac{\|a_z\|^4_\infty}{a_0^4})\right] < \infty. \qquad (2.5.20)\end{aligned}$$

□

Remark 2.5.3. *Similar to §2.3.3, Duhamel's formula would also yield a decay estimate here. The eigenfunction modes of $f(x,z,t)$ are given explicitly due to* (2.5.8). *This allows us to use Duhamel's formula for the evolution equation* (2.5.9), *providing the eigenfunction modes of $g_k(z,t)$ in explicit form.*

2.5.3 Uncertain Diffusion Coefficient

As an alternative to (2.5.1) we consider now the following Fokker–Planck equation on $\mathbb{R}$ with uncertainty in the diffusion term:

$$\partial_t u(x,z,t) = \partial_x(d(z)u_x(x,z,t) + xu(x,z,t)) =: L_1(z)u(x,z,t), \tag{2.5.21}$$

for $x, z \in \mathbb{R}$, $t \geq 0$ and a diffusion coefficient $d \in C^1(\mathbb{R})$ satisfying $d_0 := \inf_{z\in\mathbb{R}} d(z) > 0$. For $v(x,z,t) := \partial_z u(x,z,t)$ the first order linear sensitivity equation is given as

$$\partial_t v(x,z,t) = L_1(z)v(x,z,t) + d_z(z)u_{xx}(x,z,t), \tag{2.5.22}$$

for $x, z \in \mathbb{R}, t \geq 0$. A strategy similar to the one used in §2.5.2 can also be applied here: The rescaled Hermite functions

$$\hat{h}_k(x,z) := d(z)^{-\frac{1}{2}}\tilde{h}_k(xd(z)^{-\frac{1}{2}})$$

with $\tilde{h}_k(x,z)$ defined in (2.5.4) are an orthonormal basis of $L^2((\hat{h}_0)^{-1})$ of eigenfunctions of $L_1(z)$, i.e.

$$L_1(z)\hat{h}_k(\cdot,z) = -k\hat{h}_k(\cdot,z), \quad k \in \mathbb{N}_0,$$

which determines the whole (z-independent) spectrum

$$\sigma(L_1(z)) = -\mathbb{N}_0.$$

It follows that the unique normalized steady state of $L_1(z)$ and the corresponding steady state for (2.5.22) are given as

$$\begin{aligned} u^\infty(x,z) &= \hat{h}_0(x,z) = \frac{1}{\sqrt{2\pi d(z)}}e^{-\frac{x^2}{2d(z)}}, \\ v^\infty(x,z) &= \partial_z u^\infty(x,z) = \frac{d_z(z)}{\sqrt{2}d(z)}\hat{h}_2(x,z), \end{aligned}$$

respectively. An eigenfunction expansion leads to the non-defective ODE systems for the eigenfunction modes $k \geq 2$:

$$\partial_t \begin{pmatrix} u_{k-2} \\ v_k \end{pmatrix} = -\underbrace{\begin{pmatrix} k-2 & 0 \\ \frac{d_z(z)}{d(z)}\sqrt{(k-1)k} & k \end{pmatrix}}_{A_k(z):=} \begin{pmatrix} u_{k-2} \\ v_k \end{pmatrix}, \quad z \in \mathbb{R}, t \geq 0,$$

and $\partial_t v_0(z,t) = 0$, $\partial_t v_1(z,t) = -v_1(z,t)$. The matrix $A_k(z)$, $k \geq 2$, has the eigenvalues $\lambda_{1,k} = k-2$ and $\lambda_{2,k} = k$, and hence, is not defective. Contrary to the models previously investigated, the FPE with added uncertainty in the diffusion term *does not* result in the typical defective decay behavior for $v(x,z,t)$. The decay behavior of solutions $\phi(x,z,t) := (u,v)^T$ of the system (2.5.21)–(2.5.22) can hence be estimated easily by

$$\sup_{z\in\mathbb{R}} \|\phi(\cdot,z,t)-\phi^\infty(\cdot,z)\|^2_{L^2((u^\infty)^{-1})} \le \mathscr{C} e^{-t} \sup_{z\in\mathbb{R}} \|\phi(z,0)-\phi^\infty(\cdot,z)\|^2_{L^2((u^\infty)^{-1})},$$

with $\phi^\infty := (u^\infty, v^\infty)^T$, a constant $\mathscr{C} > 0$ and $t \ge 0$.

To sum up, we observe that the FPE (2.5.1) with uncertainty in the drift term gives rise to a more complicated and interesting decay behavior.

2.6 Conclusion

In this chapter we perform a sensitivity analysis for several linear PDEs with uncertainty by a Lyapunov functional method, obtaining sharp decay rates to the global equilibrium.

First, a systematic derivation of Lyapunov functionals — in the form of modified norms — for arbitrary linear ODE systems is given. The Lyapunov functional approach has a simple geometric interpretation: In the deformed metric, the angle between any trajectory and the level curves of the P-norm is uniformly bounded away from zero (for P constant in t). The novelty here is the inclusion of defective ODEs, which demand time-dependence in the norms $|\cdot|_{P(t)}$ in order to obtain sharp decay estimates of order $(1+t^M)e^{-\mu t}$. This approach is realized via a matrix $P(t)$, which is constructed from the explicit (generalized) eigenvectors of the system matrix accompanied by arbitrary weights. In the presence of an uncertainty parameter z, we obtain decay estimates that are uniform in z, which includes non-defective limits. In such cases, the matrix $P(z,t)$ has to be constructed more carefully, taking advantage of the non-uniqueness of $P(z,t)$ in the above method.

This method is applied to three PDEs, a convection-diffusion equation, a two-velocity BGK equation, and a Fokker–Planck equation, where each of these equations feature uncertainty in the equation parameters. A linear sensitivity analysis is performed, where for the convection-diffusion equation a second order sensitivity is also included. The analysis works well with PDEs that allow for a Fourier mode decomposition, since each mode evolves according to an ODE. Hence, the decay estimates have to be uniform in the eigenmodes k. In the presented examples (with the exception of §2.5.3) defects appear in the resulting ODEs.

Sharp decay estimates which are uniform in the uncertainty parameter z were obtained for these PDEs. The technical difficulty here is the possible appearance of non-defective limits due to the z-dependence of the ODEs, when one considers derivatives of solutions with respect to z. This problem is solved with a careful choice of the matrix $P(z,t)$, exploiting the fact that its construction (and in particular its weights) is not unique.

Let us point out an aspect of the sensitivity analysis of the Fokker–Planck equation in §2.5 distinct from the other investigated PDEs: The equilibrium that solutions converge to, depends itself on the uncertainty parameter.

The method could be helpful even for nonlinear PDEs with uncertainties. For those problems, usually perturbative solutions, namely solutions near global equilibria, are studied, in which the exponential decay due to the linear(-ized) hypocoercivity dominates the nonlinear growth. See e.g. [23, 28, 9, 13, 2, 1] for deterministic settings and [21, 17, 25] for inclusion of uncertainty quantification. One expects that our analysis can lead to sharper decay rates than previously used energy estimates in Sobolev spaces.

Appendix

2.A Proofs

Proof of Lemma 2.2.6

Proof. For arbitrary $i, j \in \{1, \dots, m\}$ and $\alpha > 0$

$$v^i \otimes v^j + v^j \otimes v^i \geq -\frac{1}{\alpha} Q^i - \alpha Q^j \tag{2.A.1}$$

holds true, as can be directly validated in matrix representation.

We estimate $\hat{w}_n^m(t) \otimes \hat{w}_n^m(t)$ from below by using inequality (2.A.1) for the double sum:

$$\begin{aligned}
\hat{w}_n^m(t) \otimes \hat{w}_n^m(t) &= \sum_{k=1}^{m} (\xi^k(t))^2 Q^k + \sum_{\substack{i,j \in \{1,\dots,m\} \\ i \neq j}} \xi^i(t) \xi^j(t) v^i \otimes v^j \\
&\geq \sum_{k=1}^{m} (\xi^k(t))^2 Q^k - \sum_{\substack{k,l \in \{1,\dots,m\} \\ l<k}} \left(\frac{1}{\alpha} (\xi^k(t))^2 Q^k + \alpha (\xi^l(t))^2 Q^l \right).
\end{aligned}$$

We reorder the double sum and notice that each term depends on only one of the two indices:

$$\begin{aligned}
&\hat{w}_n^m(t) \otimes \hat{w}_n^m(t) \\
&\quad \geq \sum_{k=1}^{m} (\xi^k(t))^2 Q^k - \sum_{k=2}^{m} \sum_{l=1}^{k-1} \left(\frac{1}{\alpha} (\xi^k(t))^2 Q^k + \alpha (\xi^l(t))^2 Q^l \right) \\
&\quad = \sum_{k=1}^{m} (\xi^k(t))^2 Q^k - \sum_{k=2}^{m} \frac{k-1}{\alpha} (\xi^k(t))^2 Q^k - \alpha \sum_{l=1}^{m-1} (m-l) (\xi^l(t))^2 Q^l \\
&\quad = (1 - \frac{m-1}{\alpha}) (\xi^m)^2 Q^m + \sum_{k=1}^{m-1} \left(1 - \frac{k-1}{\alpha} - \alpha(m-k) \right) (\xi^k(t))^2 Q^k.
\end{aligned}$$

For the first coefficient to be positive, we need $\alpha > m-1$. Moreover one has $\min_{k=1,\dots,m-1}\{1 - \frac{k-1}{\alpha} - \alpha(m-k)\} = 1 - \alpha(m-1)$ and therefore

$$\hat{w}_n^m(t) \otimes \hat{w}_n^m(t) \geq (1 - \frac{m-1}{\alpha}) (\xi^m)^2 Q^m - (\alpha(m-1) - 1) \sum_{k=1}^{m-1} (\xi^k(t))^2 Q^k,$$

which yields the desired result for $\theta := \frac{m-1}{\alpha} \in (0,1)$. □

Proof of Theorem 2.2.8

If $M = 1$ there are no defective eigenvalues with real part μ. In this case, the result follows from (2.2.15) and the estimates (2.2.18) with the corresponding matrix P.

For $M > 1$, we fix an arbitrary $n \in I_\mu$ and first estimate the $P_n^m(0)$-semi-norm decay for the corresponding $m \in \{2, \ldots, l_n\}$. To achieve this, we combine the decay estimate (2.2.10) and (2.2.20) that gives a lower bound on the $P_n^m(t)$-semi-norm with terms only depending on $P_n^k(0)$-semi-norms ($k \in \{1, \ldots, m\}$). This yields

$$\begin{aligned}&(1-\theta)|x(t)|^2_{P_n^m(0)} - \left(\frac{(m-1)^2}{\theta} - 1\right)\sum_{k=1}^{m-1}\left(\frac{t^{m-k}}{(m-k)!}\right)^2 |x(t)|^2_{P_n^k(0)}\\ &\le |x(t)|^2_{P_n^m(t)} = e^{-2\mu t}|x(0)|^2_{P_n^m(0)}.\end{aligned}$$

Rearranging and dividing by $(1-\theta)$ leads to

$$\begin{aligned}|x(t)|^2_{P_n^m(0)} \le &\underbrace{\left(\frac{(m-1)^2}{\theta} - 1\right)\frac{1}{1-\theta}}_{d_m(\theta):=}\sum_{k=1}^{m-1}\left(\frac{t^{m-k}}{(m-k)!}\right)^2 |x(t)|^2_{P_n^k(0)}\\ &+\frac{1}{1-\theta}e^{-2\mu t}|x(0)|^2_{P_n^m(0)}.\end{aligned} \tag{2.A.2}$$

By induction we shall show that, for arbitrary but fixed $n \in I_\mu$ and all corresponding $m \in \{1, \ldots, l_n\}$, there exists a constant $c_m > 0$ only depending on m, such that

$$|x(t)|^2_{P_n^m(0)} \le \frac{1}{\min\limits_{k=1,\ldots,m} \beta_n^k} c_m(1 + t^{2(m-1)})e^{-2\mu t}|x(0)|^2_{P_n(0)}, \quad t \ge 0, \tag{2.A.3}$$

where $P_n(0) = \sum_{m=1}^{l_n} \beta_n^m P_n^m(0)$ by definition (2.2.11).

For $m = 1$, the matrix P_n^m is not time-dependent and (2.2.10) immediately yields (2.A.3) with $c_1 = \frac{1}{2}$.

For the inductive step, we assume the claim is true for all $k \in \{1, \ldots, m\}$ with some $m \ge 1$ and constants $c_k > 0$ monotonically increasing in k and start from (2.A.2), written for $m+1$:

$$\begin{aligned}|x(t)|^2_{P_n^{m+1}(0)} &\le d_{m+1}(\theta)\sum_{k=1}^{m}\left(\frac{t^{m+1-k}}{(m+1-k)!}\right)^2 |x(t)|^2_{P_n^k(0)}\\ &\quad+\frac{1}{1-\theta}e^{-2\mu t}|x(0)|^2_{P_n^{m+1}(0)}\\ &\le d_{m+1}(\theta)\sum_{k=1}^{m}\frac{t^{2(m+1-k)}}{[(m+1-k)!]^2}\frac{1}{\min\limits_{j\in\{1,\ldots,k\}}\beta_n^j}c_k(1+t^{2(k-1)})e^{-2\mu t}|x(0)|^2_{P_n(0)}\\ &\quad+\frac{1}{\beta_n^{m+1}}\frac{1}{1-\theta}e^{-2\mu t}|x(0)|^2_{P_n(0)}\end{aligned} \tag{2.A.4}$$

where (2.A.3) with $k \in \{1, \dots, m\}$ was used in the second estimate.

In order to combine both of the terms in (2.A.4) into one summation, we compute $\inf_{\theta\in(0,1)} \max\{d_{m+1}(\theta), \frac{1}{1-\theta}\}$. For $m = 1$, one has $\inf_{\theta\in(0,1)} \max\{d_2(\theta), \frac{1}{1-\theta}\} = 2$. For $m > 1$, the coefficient $d_{m+1}(\theta)$ has its minimum at $\theta_{\min} = m^2 - m\sqrt{m^2-1}$ with value $d_{m+1}(\theta_{\min}) = 2m^2 + 2m\sqrt{m^2-1} - 1$. As $d_{m+1}(\theta_{\min}) \geq \frac{1}{1-\theta_{\min}}$, one gets $\inf_{\theta\in(0,1)} \max\{d_{m+1}(\theta), \frac{1}{1-\theta}\} = d_{m+1}(\theta_{\min})$.

In total

$$\inf_{\theta\in(0,1)} \max\{d_{m+1}(\theta), \frac{1}{1-\theta}\} \leq 4m^2 - 1, \qquad m \geq 1. \tag{2.A.5}$$

Applying this estimate to (2.A.4) and, additionally, using the upper bound $\max_{k=1,\dots,m} t^{2(m+1-k)} + t^{2m} \leq 2(1 + t^{2m})$ for all $t \geq 0$, one gets

$$\begin{aligned} |x(t)|^2_{P_n^{m+1}(0)} \leq{}& \frac{1}{\min\limits_{k\in\{1,\dots,m+1\}} \beta_n^k} \\ &\times \underbrace{2(4m^2-1)c_m \sum_{k=1}^{m+1} \frac{1}{[(m+1-k)!]^2}}_{c_{m+1}:=} (1+t^{2m})\, e^{-2\mu t} |x(0)|^2_{P_n(0)}, \end{aligned}$$

which concludes the induction and hence the proof of (2.A.3) for $m \in \{1, \dots, l_n\}$.

The constant c_{m+1} for $m \in \{1, \dots, l_n - 1\}$ is given as

$$c_{m+1} = 2^{m-1} \left(\prod_{j=1}^{m} 4j^2 - 1 \right) \left(\prod_{j=2}^{m+1} \sum_{k=1}^{j} \frac{1}{[(j-k)!]^2} \right).$$

By definition (2.2.13) the matrix $P(0)$ is given as

$$P(0) = \underbrace{\sum_{n\notin I_\mu} \beta_n P_n}_{P_{I_\mu^c}:=} + \underbrace{\sum_{n\in I_\mu} \sum_{m=1}^{l_n} \beta_n^m P_n^m(0)}_{P_{I_\mu}:=}.$$

The first term, $P_{I_\mu^c}$, covers the Cases 1–2. Applying Gronwall's lemma directly to the inequalities (2.2.4) and (2.2.6) yields

$$|x(t)|^2_{P_{I_\mu^c}} \leq e^{-2\mu t} |x(0)|^2_{P_{I_\mu^c}}.$$

Now, we take a closer look at the decay behavior of solutions with respect to the P_{I_μ}-semi-norm that corresponds to Case 3.

$$|x(t)|^2_{P_{I_\mu}} = \left(\sum_{n\in I_\mu} \beta_n^1 |x(t)|^2_{P_n^1(0)} + \sum_{n\in I_\mu} \sum_{m=2}^{l_n} \beta_n^m |x(t)|^2_{P_n^m(0)} \right). \tag{2.A.6}$$

The first term includes only semi-norms that are time-independent, i.e. $P_n^1(t) \equiv P_n^1(0)$, and using (2.2.10) directly gives the decay behavior. For the second term in (2.A.6), we apply (2.A.3) and get

$$\begin{aligned}
|x(t)|^2_{P_{I_\mu}} &\leq \sum_{n\in I_\mu} \beta_n^1 e^{-2\mu t} |x(0)|^2_{P_n^1(0)} \\
&\quad + \sum_{n\in I_\mu} \sum_{m=2}^{l_n} \frac{\beta_n^m}{\min\limits_{k\in\{1,\dots,m\}} \beta_n^k} c_m (1 + t^{2(m-1)}) e^{-2\mu t} |x(0)|^2_{P_n(0)} \\
&\leq e^{-2\mu t} \sum_{n\in I_\mu} |x(0)|^2_{P_n(0)} \\
&\quad + \max_{n\in I_\mu} \Bigg[\sum_{m=2}^{l_n} \frac{\beta_n^m}{\min\limits_{k\in\{1,\dots,m\}} \beta_n^k} c_m \Bigg] 2(1 + t^{2(M-1)}) e^{-2\mu t} \sum_{n\in I_\mu} |x(0)|^2_{P_n(0)} \\
&\leq 2 c_M \max_{n\in I_\mu} \Bigg[\sum_{m=1}^{l_n} \frac{\beta_n^m}{\min\limits_{k\in\{1,\dots,m\}} \beta_n^k} \Bigg] (1 + t^{2(M-1)}) e^{-2\mu t} |x(0)|^2_{P_{I_\mu}}.
\end{aligned}$$

Now, using (2.2.18) for $P(0)$, the decay behavior in the Euclidean norm follows as

$$\begin{aligned}
|x(t)|_2^2 &\leq (\lambda_{\min}^{P(0)})^{-1} |x(t)|^2_{P(0)} \\
&= (\lambda_{\min}^{P(0)})^{-1} \left(|x(t)|^2_{P_{I_\mu^c}} + |x(t)|^2_{P_{I_\mu}} \right) \\
&\leq (\lambda_{\min}^{P(0)})^{-1} \Bigg(e^{-2\mu t} |x(0)|^2_{P_{I_\mu^c}} \\
&\quad + 2 c_M \max_{n\in I_\mu} \Bigg[\sum_{m=1}^{l_n} \frac{\beta_n^m}{\min\limits_{k\in\{1,\dots,m\}} \beta_n^k} \Bigg] (1 + t^{2(M-1)}) e^{-2\mu t} |x(0)|^2_{P_{I_\mu}} \Bigg) \\
&\leq 2 (\lambda_{\min}^{P(0)})^{-1} \lambda_{\max}^{P(0)} c_M \max_{n\in I_\mu} \Bigg[\sum_{m=1}^{l_n} \frac{\beta_n^m}{\min\limits_{k\in\{1,\dots,m\}} \beta_n^k} \Bigg] (1 + t^{2(M-1)}) e^{-2\mu t} |x(0)|_2^2,
\end{aligned}$$

where the constant $\mathscr{C} := 2(\lambda_{\min}^{P(0)})^{-1} \lambda_{\max}^{P(0)} c_M \max_{n\in I_\mu} \Big[\sum_{m=1}^{l_n} \frac{\beta_n^m}{\min_{k\in\{1,\dots,m\}} \beta_n^k} \Big]$ depends only on the matrix $P(0)$. □

Proof of Lemma 2.3.2

Proof. Since $w_{1,k}^2, w_{1,k}^3$ satisfy (2.2.8) with $n = 1$ and $m = 2,3$, their linear combination $\widetilde{w}_k^3$ satisfies

$$\frac{d}{dt} \widetilde{w}_k^3(z,t) = (D_k^H(z) - \overline{\lambda}_k(z)) \widetilde{w}_k^3(z,t).$$

The computation leading to (2.2.10) also applies here (with rescaled time $\tau_k = k^2 t$) and results in

$$|y_k(z,t)|^2_{\widetilde{P}^3_k(z,k^2 t)} = e^{-2k^2 b(z)t} |y_k(z,0)|^2_{\widetilde{P}^3_k(z,0)}. \tag{2.A.7}$$

Now, our goal is an estimate in the $\widetilde{P}^3_k(z,0)$-semi-norm with help of Lemma 2.2.6. By definition

$$\widetilde{w}^3_k(z,t) = \xi^1_k(z,t)\, w^1_{1,k}(z,0) + \xi^2_k(z,t)\, w^2_{1,k}(z,0) + \xi^3_k \widetilde{w}^3_k(z,0),$$

with the polynomials

$$\xi^1_k(z,t) = \frac{t^2}{2} + \frac{\partial_z^2 \overline{\lambda}_k(z)}{2(\partial_z \overline{\lambda}_k(z))^2} t, \quad \xi^2_k(z,t) = t \quad \text{and} \quad \xi^3_k(z,t) = 1.$$

Lemma 2.2.6 yields

$$\begin{aligned} |x|^2_{\widetilde{P}^3_k(z,t)} \geq{}& (1-\theta)|x|^2_{\widetilde{P}^3_k(z,0)} \\ & - \left(\frac{4}{\theta} - 1\right) \left[|\xi^1_k(z,t)|^2 |x|^2_{P^1_{1,k}(z,0)} + |\xi^2_k(z,t)|^2 |x|^2_{P^2_{1,k}(z,0)} \right] \end{aligned} \tag{2.A.8}$$

for any (fixed) $x \in \mathbb{C}^3$, $\theta \in (0,1)$, $t \geq 0$. Replace x by a solution $y_k(z,t)$ to (2.3.12) (rescaling $\xi_k(z,t)$ to $\xi_k(z,k^2 t)$ to account for the prefactor k^2 in the ODE). Then (in analogy to the estimate (2.A.2)), using (2.A.7) leads to

$$\begin{aligned} |y_k(z,t)|^2_{\widetilde{P}^3_k(z,0)} \leq{}& d_3(\theta) \left[|\xi^1_k(z,k^2 t)|^2 |y_k(z,t)|^2_{P^1_{1,k}(z,0)} + |\xi^2_k(z,k^2 t)|^2 |y_k(z,t)|^2_{P^2_{1,k}(z,0)} \right] \\ & + \frac{1}{1-\theta} e^{-2k^2 b(z)t} |y_k(z,0)|^2_{\widetilde{P}^3_k(z,0)}, \end{aligned} \tag{2.A.9}$$

with $d_3(\theta) := \frac{4-\theta}{\theta(1-\theta)}$ as in (2.A.2). Minimizing in θ with estimate (2.A.5) yields

$$\inf_{\theta \in (0,1)} \max \left\{ d_3(\theta), \frac{1}{1-\theta} \right\} \leq 15.$$

Next we use the estimates (2.3.16), (2.3.17) and the fact that $\widetilde{P}^3_k(z,0) \leq \frac{1}{4|\partial_z \lambda_k(z)|^4} I$ to proceed with (2.A.9):

$$
\begin{aligned}
|y_k(z,t)|^2_{\widetilde{P}^3_k(z,0)} \leq 15 &\left[\left|\frac{k^4t^2}{2} + k^2 t\frac{\partial_z^2\overline{\lambda}_k(z)}{2(\partial_z\overline{\lambda}_k(z))^2}\right|^2 + \frac{6k^4t^2}{\min\{1,|\partial_z\lambda_k(z)|^2\}}(1+k^4t^2)\right.\\
&\left.\qquad + \frac{1}{4|\partial_z\lambda_k(z)|^4}\right] e^{-2k^2b(z)t}|y_k(z,0)|_2^2\\
\leq 15\frac{1+|\partial_z^2\lambda_k(z)|^2}{\min\{1,|\partial_z\lambda_k(z)|^4\}} &\left[\frac{k^8t^4}{4} + \frac{k^4t^2}{4} + \frac{k^6t^3}{2}\right.\\
&\left.\qquad + 6(k^4t^2 + k^8t^4) + \frac{1}{4}\right] e^{-2k^2b(z)t}|y_k(z,0)|_2^2\\
\leq 146.25\frac{1+|\partial_z^2\lambda_k(z)|^2}{\min\{1,|\partial_z\lambda_k(z)|^4\}}&(1+k^8t^4)e^{-2k^2b(z)t}|y(z,0)|_2^2.
\end{aligned}
$$

□

Proof of Lemma 2.5.1

Proof. To show that the matrix

$$
\widetilde{P}(z) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2}\min\{1,\frac{1}{\alpha(z)^4}\} \end{pmatrix} \tag{2.A.10}
$$

satisfies

$$
C_k^H(z)\widetilde{P}(z) + \widetilde{P}(z)C_k(z) \geq \frac{1}{2}\widetilde{P}(z), \quad z\in\mathbb{R}, k\geq 4, \tag{2.A.11}
$$

we show that

$$
\begin{aligned}
A_k(z) &:= C_k^H(z)\widetilde{P}(z) + \widetilde{P}(z)C_k(z) - \frac{1}{2}\widetilde{P}(z)\\
&= \begin{pmatrix} \frac{3}{2}-\frac{4}{k} & 0 & \frac{1}{2}\gamma(k)\min\{\alpha(z),\frac{1}{\alpha(z)^3}\} \\ 0 & \frac{3}{2} & \frac{1}{2}\min\{\alpha(z),\frac{1}{\alpha(z)^3}\} \\ \frac{1}{2}\gamma(k)\min\{\alpha(z),\frac{1}{\alpha(z)^3}\} & \frac{1}{2}\min\{\alpha(z),\frac{1}{\alpha(z)^3}\} & \frac{3}{4}\min\{1,\frac{1}{\alpha(z)^4}\} \end{pmatrix}
\end{aligned}
$$

is positive definite.

As $k\geq 4$, the first two leading minors are positive. The third minor is positive, if

$$
\begin{aligned}
\det A_k(z) = \frac{9}{8}\left(\frac{3}{2}-\frac{4}{k}\right)\min\{1,\frac{1}{\alpha(z)^4}\} - \frac{3}{8}\gamma(k)^2\min\{\alpha(z)^2,\frac{1}{\alpha(z)^6}\}\\
-\frac{1}{4}\left(\frac{3}{2}-\frac{4}{k}\right)\min\{\alpha(z)^2,\frac{1}{\alpha(z)^6}\} > 0
\end{aligned}
$$

for all $z \in \mathbb{R}$, $k \geq 4$. For all $k \geq 4$, we distinguish the following two cases:

For $z \in \mathbb{R}$ such that $|\alpha(z)| \geq 1$ the condition $\det A_k(z) > 0$ is equivalent to

$$f(k, \alpha(z)^2, \gamma(k)) := \left(\frac{3}{2} - \frac{4}{k}\right)\left(\frac{9}{4} - \frac{1}{2\alpha(z)^2}\right) - \frac{3}{4\alpha(z)^2}\gamma(k)^2 > 0.$$

The function $[4,\infty) \times [1,\infty) \times [\sqrt{\frac{2}{3}}, 1) \ni (k, \alpha^2, \gamma) \mapsto f(k, \alpha^2, \gamma)$, is monotonously increasing in k and α^2 but monotonously decreasing in γ, hence

$$f(k, \alpha(z)^2, \gamma(k)) \geq f(4,1,1) = \frac{1}{8} > 0.$$

For $z \in \mathbb{R}$ such that $|\alpha(z)| \leq 1$ the condition $\det A_k(z) > 0$ is equivalent to

$$g(k, \alpha(z)^2, \gamma(k)) := \left(\frac{3}{2} - \frac{4}{k}\right)\left(\frac{9}{4} - \frac{1}{2}\alpha(z)^2\right) - \frac{3}{4}\gamma(k)^2\alpha(z)^2 > 0.$$

The function $[4,\infty) \times [0,1] \times [\sqrt{\frac{2}{3}}, 1) \ni (k, \alpha^2, \gamma) \mapsto g(k, \alpha^2, \gamma)$ is monotonously increasing in k and monotonously decreasing in α^2 and γ, hence

$$g(k, \alpha(z)^2, \gamma(k)) \geq g(4,1,1) = \frac{1}{8} > 0.$$

This proves the matrix inequality (2.A.11). With a similar calculation as (2.2.4) (and the rescaling $t \mapsto k\alpha(z)t$), this implies

$$|y_k(z,t)|^2_{\widetilde{P}(z)} \leq e^{-\frac{1}{2}k\alpha(z)t}|y_k(z,0)|^2_{\widetilde{P}(z)}, \quad t \geq 0, k \geq 4.$$

With $\frac{1}{2}\min\{1, \frac{1}{\alpha(z)^4}\}I \leq \widetilde{P}(z) \leq I$, $z \in \mathbb{R}$, we obtain

$$|y_k(z,t)|_2^2 \leq 2\max\{1, \alpha(z)^4\}e^{-2\alpha(z)t}|y_k(z,0)|_2^2, \quad t \geq 0, k \geq 4,$$

from which the desired result follows. □

Bibliography

[1] Achleitner, F., Arnold, A. and Carlen, E.A.: On linear hypocoercive BGK models. In: *"From particle systems to partial differential equations", III.* Springer Proc. Math. Stat., vol. 162, 1–37. Springer (2016).

[2] Achleitner, F., Arnold, A. and Carlen, E.A.: *On multi-dimensional hypocoercive BGK models.* Kinetic & Related Models, vol. 11 (4) 953–1009 (2018).

[3] Achleitner, F., Arnold, A. and Stürzer, D.: *Large-Time Behavior in Non-Symmetric Fokker–Planck Equations.* Rivista di Matematica della Università di Parma, vol. 6, 1–68, (2015).

[4] Arnold, A., Einav, A. and Wöhrer, T.: *On the rates of decay to equilibrium in degenerate and defective Fokker–Planck equations.* J. Differential Equations, vol. 264 (11), 6843–6872, (2018).

[5] Arnold, A. and Erb, J.: *Sharp Entropy Decay for Hypocoercive and Non-Symmetric Fokker-Planck Equations With Linear Drift.* Preprint, arXiv:1409.5425 (2014).

[6] Arnold, A., Markowich, P., Toscani, G. and Unterreiter, A.: *On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker–Planck type equations.* Comm. Partial Differential Equations, vol. 26, 43–100, (2001).

[7] Arnold, V.I.: *Ordinary differential equations.* MIT Press, Cambridge, Mass.-London, Translated from the Russian and edited by Richard A. Silverman, (1978).

[8] Bhatnagar, P. L., Gross, E. P. and Krook, M.: *A Model for Collision Processes in Gases. I. Small Amplitude Processes in Charged and Neutral One-Component Systems.* Phys. Rev., vol. 94 (3), 511–525, (1954).

[9] Briant, M.: *From the Boltzmann equation to the incompressible Navier–Stokes equations on the torus: A quantitative error estimate.* J. Differential Equations, vol. 259 (11), 6072–6141, (2015).

[10] Desvillettes, L. and Villani, C.: *On the trend to global equilibrium for spatially inhomogeneous kinetic systems: the Boltzmann equation.* Invent. Math., vol. 159, 245–316, (2005).

[11] Dolbeault, J., Mouhot, C. and Schmeiser, C.: *Hypocoercivity for linear kinetic equations conserving mass.* Trans. Amer. Math. Soc., vol. 367 (6), 3807–3828 (2015).

[12] Gunzburger, M., Webster, C. and Zhang, G.: *Stochastic finite element methods for partial differential equations with random input data.* Acta Numerica, vol. 23, 521–650, (2014).

[13] Guo, Y.: *Boltzmann diffusive limit beyond the Navier-Stokes approximation.* Comm. Pure Appl. Math., vol. 59, 626–687, (2006).

[14] Hu, J. and Jin, S.: *Uncertainty quantification for kinetic equations.* In “Uncertianty Quantification for Kinetic and Hyperbolic Equations”, SEMA-SIMAI Springer Series (ed. S. Jin and L. Pareschi), 193–229, (2018).

[15] Jin, S., Liu, J.-G. and Ma, Z.: *Uniform spectral convergence of the stochastic Galerkin method for the linear transport equations with random inputs in diffusive regime and a micro-macro decomposition based asymptotic preserving method.* Res. Math. Sci., vol. 4 (15), (2017).

[16] Jin, S., Liu, L.: *An asymptotic-preserving stochastic Galerkin method for the semiconductor Boltzmann equation with random inputs and diffusive scalings.* SIAM Multiscale Model. Simul., vol. 15 (1), 157–183, (2017).

[17] Jin, S. and Zhu, Y.: *Hypocoercivity and uniform regularity for the Vlasov-Poisson-Fokker-Planck system with uncertainty and multiple Scales.* SIAM J. Math. Anal., vol. 50, 1790–1816, (2018).

[18] Jost, J.: *Riemannian geometry and geometric analysis.* Universitext, (2017).

[19] Jüngel, A.: *Entropy Methods for Diffusive Partial Differential Equations.* BCAM Springer Briefs, (2016).

[20] Li, Q. and Wang, L.: *Uniform regularity for linear kinetic equations with random input based on hypocoercivity.* SIAM/ASA J. Uncertainty Quantification, vol. 5 (1), 1193–1219, (2017).

[21] Liu, L. and Jin, S.: *Hypocoercivity based sensitivity analysis and spectral convergence of the stochastic Galerkin approximation to collisional kinetic equations with multiple scales and random inputs,* Multiscale Model. Simul., vol. 16 (3), 1085–1114, (2017).

[22] Monmarché, P.: *Generalized* Γ *calculus and application to interacting particles on a graph,* Potential Analysis, vol. 50, 439–466 (2018).

[23] Mouhot, C. and Neumann, L.,: *Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus.*, Nonlinearity, vol. 19 (4), 969–998, (2006).

[24] Risken, H.: *The Fokker-Planck equation. Methods of solution and applications.* Springer-Verlag, (1989).

[25] Shu, R.W. and Jin, S.: *Uniform regularity in the random space and spectral accuracy of the stochastic Galerkin method for a kinetic-fluid two-phase flow model with random initial inputs in the light particle regime.* vol. 52 (5), 1651–1678 (2017).

[26] Smith, R.: *Uncertainty quantification: theory, implementation, and applications.* SIAM, (2013).

[27] Toscani, G.: *Kinetic approach to the asymptotic behaviour of the solution to diffusion equations.* Rend. di Matematica 16, 329–346, (1996).

[28] Villani, C.: *Hypocoercivity,* American Mathematical Soc., (2009).

3 On the Goldstein–Taylor Equation with Space-Dependent Relaxation

3.1 Introduction

The object of this chapter is the large-time analysis of the *Goldstein–Taylor equation* on the one-dimensional torus $\mathbb{T}$, i.e. on $[0,2\pi]$ with periodic boundary conditions, and for $t \in (0,\infty)$:

$$\begin{aligned}\partial_t f_+(x,t) + \partial_x f_+(x,t) &= \frac{\sigma(x)}{2}(f_-(x,t) - f_+(x,t)),\\ \partial_t f_-(x,t) - \partial_x f_-(x,t) &= -\frac{\sigma(x)}{2}(f_-(x,t) - f_+(x,t)),\\ f_\pm(x,0) &= f_{\pm,0}(x),\end{aligned} \tag{3.1.1}$$

where $f_\pm(x,t)$ are the density functions of finding an element with a velocity ±1 in a position $x \in \mathbb{T}$ at time $t > 0$. The function

$$\sigma \in L^\infty_+(\mathbb{T}) := \left\{ f \in L^\infty(\mathbb{T}) \,\middle|\, \text{ess min } f > 0 \right\}$$

is the relaxation coefficient, and $f_{\pm,0}$ are the initial conditions. Since (3.1.1) is mass conserving, its steady state is of the form

$$f_{\pm,\infty}(x) = f_\infty\,, \quad x \in \mathbb{T}\,; \qquad f_\infty := \frac{1}{2}(f_{+,0} + f_{-,0})_{\text{avg}},$$

with the notation

$$h_{\text{avg}} := \frac{1}{2\pi}\int_0^{2\pi} h(x)\,dx. \tag{3.1.2}$$

The Goldstein–Taylor model was originally considered as a diffusion process, resulting as a limit of a discontinuous random migration in 1D, where particles may change direction with rate σ. It appeared in the context of turbulent fluid motion and the telegrapher's equation, see [21, 14], respectively. (3.1.1) can also be seen as a special, 1D case of a BGK-model (named after the three physicists Bhatnagar, Gross, and Krook [10]), a kinetic equation with discrete velocities. They appear in applications like gas and fluid dynamics as velocity discretisations of various kinetic models (e.g. the Boltzmann equation). The mathematical analysis of such discrete velocity models has a long standing tradition, see [11, 17] and references therein.

Although the Goldstein–Taylor equation is very simple, it still exhibits an interesting and mathematically rich structure. Hence, it has been attracting continuous interest over the last 20 years. Most of its mathematical analyses were devoted to the following three topics: scaling limits, asymptotic preserving (AP) numerical schemes, and its large-time behaviour. In a diffusive scaling, the Goldstein–Taylor model can be viewed as a hyperbolic approximation to the heat equation [20]. Various AP-schemes for this model in the stiff relaxation regime (i.e. for $\sigma \to \infty$) were constructed and analysed in [16, 15, 4]. Since the large-time convergence of solutions to (3.1.1) towards its unique steady state is also the topic of this chapter, we shall review the related literature in some more detail.

Analytically, the main difficulty of (3.1.1) is its hypocoercivity, as defined in [23]: The relaxation operator on the r.h.s. is not coercive on $\mathbb{T} \times \mathbb{R}^2$. Hence, for each fixed x, the r.h.s. by itself would drive the system to its local equilibrium, it is to the kernel of the relaxation operator, $\operatorname{span}\{\binom{1}{1}\}$, but the local mass (density) might be different at different positions. Convergence to the global equilibrium $(f_\infty, f_\infty)^T$ only arises due to the interplay between local relaxation and the transport operator on the l.h.s. of (3.1.1).

The Goldstein–Taylor model is included in the analysis of [5], when choosing the velocity matrix $V = \operatorname{diag}(1,-1)$ and the relaxation matrix $\boldsymbol{A}(x) = \frac{\sigma(x)}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \geq 0$. Exponential convergence to the steady state is proved there for the system (3.1.1) with inflow boundary conditions. Such boundary conditions make the problem significantly easier than in the periodic set-up envisioned here, in particular when allowing for $\sigma(x)$ to be zero on a subset of $\mathbb{T}$.

In [12] the authors proved polynomial decay towards the equilibrium, allowing $\sigma(x)$ to vanish at finitely many points.

In [22] the author proved exponential decay for solutions to (3.1.1) with more general $\sigma(x) \geq 0$. That work is based on a (non-local in time) *weak coercive estimate* on the damping.

While the papers mentioned so far did not focus on the optimality of the (exponential) decay rate, Bernard and Salvarani [9] were able to prove exponential decay for the case $\sigma(x) \geq 0$ with the optimal rate

$$\mu(\sigma) = \min\left\{\sigma_{\text{avg}}, \tilde{D}(0)\right\},$$

where $\tilde{D}(0)$ is the spectral gap of the telegrapher's equation[1], but excluding the case when some of those eigenvalues with real part equal to $\mu(\sigma)$ are defective. On the one

[1] More precisely, $\tilde{D}(0) = \inf\left\{\operatorname{Re}\lambda_j \,\middle|\, \lambda_j \in \text{ spectrum of } \begin{pmatrix} 0 & -1 \\ -\partial_{xx} & \sigma \end{pmatrix} \setminus \{0\}\right\}$, with this matrix being the generator of the telegrapher's equation, see [9, Proposition 3.5], [18, Theorem 2]. We note that, for σ

hand, this result closes the large-time analysis of the Goldstein–Taylor model, up to the restrictive requirement $f_{\pm,0} \in H^1(\mathbb{T})$. But on the other hand, even for simple non-constant relaxation functions $\sigma(x)$, the precise value of the spectral gap $\tilde{D}(0)$ is hardly accessible, see e.g. Appendix 3.A. Moreover, as [9] heavily relies on the equivalence of (3.1.1) to the telegrapher's equation, it cannot be extended to other discrete velocity models in 1D.

This motivates our subsequent analysis: We aim for a method that can be extended to other discrete velocity BGK-models (as illustrated below on a system with 3 velocities), that admits L^2-initial data, and that yields sharp rates for constant σ. Moreover, it should also apply to general non-homogeneous $\sigma \in L^\infty_+(\mathbb{T})$. In the non-homogeneous case, however, it will not achieve an optimal rate of convergence to the appropriate equilibrium of the system. The method to be derived here will use a Lyapunov function technique in the spirit of the earlier works [23, 13, 1, 2].

This chapter is structured as follows: In §3.2 we give the analytical setting of the problem and present the main convergence theorem. In §3.3 we will consider some known results for the torus, and explore some properties of the entropy functional and the anti-derivative of a function on $\mathbb{T}$, defined in (3.2.2) and (3.2.3). In §3.4 we will consider the case where $\sigma(x) = \sigma$ is constant, explore the spectral properties of the operator which governs (3.2.1), see how E_θ arises as one recast the Fourier information in spatial variables, and conclude by proving part (a) of our main theorem. Continuing to §3.5, we will prove, using our perturbative approach to the problem, part (b) of our main theorem. The robustness of our method will be shown in §3.6 where we use it to obtain an explicit rate of convergence for a three velocities Goldstein–Taylor model, and in our Appendix 3.A we will discuss a potential way one can improve the technique we would have presented in §3.5, and explicitly show the lack of optimality of it for a particular case.

3.2 The Setting of the Problem and Main Results

To better understand the Goldstein–Taylor system, (3.1.1), one starts by recasting it in the macroscopic variables

$$u := f_+ + f_- \geq 0, \qquad v := f_+ - f_-,$$

which yield the system on $\mathbb{T} \times (0,\infty)$:

$$\begin{aligned} \partial_t u(x,t) + \partial_x v(x,t) &= 0, \\ \partial_t v(x,t) + \partial_x u(x,t) &= -\sigma(x) v(x,t), \\ u(\cdot,0) = u_0 := f_{+,0} + f_{-,0}, \quad & v(\cdot,0) = v_0 := f_{+,0} - f_{-,0}\,, \end{aligned} \tag{3.2.1}$$

constant and in Fourier space, the matrix $\begin{pmatrix} 0 & -1 \\ k^2 & \sigma \end{pmatrix}$ is related to C_k from (3.4.2) by a simple similarity transformation.

whose theory of existence and uniqueness is straightforward (since the r.h.s. is a bounded perturbation of the transport operator; see §2 in [12] or, more generally, [19]). Moreover, when one tries to understand the qualitative behaviour of (3.2.1), one notices that the equation for u speaks of "total mass conservation" (upon integration over the spatial interval $(0,2\pi)$), while the equation for v predicts a strong decay to zero for the function. This means, at least intuitively, that the difference between f_+ and f_- should go to zero, and that their sum retains its mass. As the main driving force of the equation is a transport operation on the torus, we will not be surprised to learn that the long time behaviour of u (and since v should go to zero, of f_+ and f_- as well) is convergence to a constant. All of this has been verified in several cases, most generally in [9].

We now set the framework that will assist us in the investigation of the long time behaviour of (3.2.1), in a relatively general case. The natural Hilbert space to consider this problem is $L^2(\mathbb{T})^2$, with the standard inner product for each component:

$$\langle f_1, f_2\rangle := \frac{1}{2\pi}\int_0^{2\pi} f_1(x)\overline{f_2(x)}dx,$$

where the bar denotes complex conjugation. Since (3.1.1) and (3.2.1) are (only) hypocoercive, the symmetric part of their generators (i.e. the operators on their r.h.s.) are not coercive on $L^2(\mathbb{T})^2$. Hence, the standard L^2-norm cannot serve as a usable Lyapunov functional. As it is typical for hypocoercive equations (see [23, 13, 1]), a possible remedy is to rather consider a "twisted" norm (often also referred to as *entropy functional*), constructed such that this functional strictly decays along each trajectory $(u(t), v(t))$. The following functional is not an ansatz, but it will be derived in §3.4 as a Lyapunov functional to yield the sharp exponential decay for constant σ, when using an appropriate $\theta = \theta(\sigma)$.

Definition 3.2.1. Let $f, g \in L^2(\mathbb{T})$ and let $\theta > 0$ be given. Then we define the entropy $E_\theta(f,g)$ as

$$E_\theta(f,g) := \|f\|^2 + \|g\|^2 - \frac{\theta}{2\pi}\int_0^{2\pi} \mathrm{Re}\left(\partial_x^{-1} f(x)\overline{g(x)}\right) dx, \tag{3.2.2}$$

where the *anti-derivative* of f is defined as

$$\partial_x^{-1} f(x) := \int_0^x f(y)dy - \left(\int_0^x f(y)dy\right)_{\text{avg}} \tag{3.2.3}$$

with the average defined in (3.1.2).

Several recent studies (like [13, 1]) considered the Goldstein–Taylor system with constant σ. This case is fairly easy as it can be based on a Fourier analysis, constructing a Lyapunov functional as a sum of quadratic operators for each Fourier mode. But the moment we change $\sigma(x)$ to a non-constant function — even to one that is natural in the Fourier setting, such as sine or cosine — the Fourier analysis becomes almost impossible to solve.

The main idea that guided us in our approach was to re-examine the case where σ is constant and *to recast the modal Fourier norm by using a pseudo-differential operator*, without needing its modal decomposition. This functional, which is exactly E_θ for particular choices of $\theta = \theta(\sigma)$, can then be *extended* to the case where $\sigma(x)$ is not constant, and it yields quantitative results for the convergence. As the nature of this approach is perturbative, our decay rates are then not optimal. However, the methodology itself is fairly robust, and is viable in other cases, such as the multi-velocity Goldstein–Taylor model (as we shall see).

The main theorem we will show in this chapter is the following, where we shall use the vector notation

$$f(t) := \begin{pmatrix} f_+(t) \\ f_-(t) \end{pmatrix}, \quad f_0 := \begin{pmatrix} f_{+,0} \\ f_{-,0} \end{pmatrix}. \tag{3.2.4}$$

Theorem 3.2.2. *Let $u, v \in C([0,\infty); L^2(\mathbb{T}))$ be mild[2] real valued solutions to* (3.2.1) *with initial datum u_0, $v_0 \in L^2(\mathbb{T})$. Denoting by $u_{\text{avg}} := (u_0)_{\text{avg}}$ follows:*

a) If $\sigma(x) = \sigma$ is constant we have that:

If $\sigma \neq 2$ then

$$E_{\theta(\sigma)}\left(u(t) - u_{\text{avg}}, v(t)\right) \leq E_{\theta(\sigma)}\left(u_0 - u_{\text{avg}}, v_0\right) e^{-2\mu(\sigma)t}$$

for all $t \geq 0$, where

$$\theta(\sigma) := \begin{cases} \sigma, & 0 < \sigma < 2, \\ \frac{4}{\sigma}, & \sigma > 2, \end{cases} \qquad \mu(\sigma) := \begin{cases} \frac{\sigma}{2}, & 0 < \sigma < 2, \\ \frac{\sigma}{2} - \sqrt{\frac{\sigma^2}{4} - 1}, & \sigma > 2. \end{cases}$$

If $\sigma = 2$ then for any $0 < \varepsilon < 1$

$$E_{\frac{2(2-\varepsilon^2)}{2+\varepsilon^2}}\left(u(t) - u_{\text{avg}}, v(t)\right) \leq E_{\frac{2(2-\varepsilon^2)}{2+\varepsilon^2}}\left(u_0 - u_{\text{avg}}, v_0\right) e^{-2(1-\varepsilon)t},$$

for all $t \geq 0$. Consequently if $\sigma \neq 2$

$$\left\| f(t) - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| \leq C_\sigma \left\| f_0 - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| e^{-\mu(\sigma)t} \tag{3.2.5}$$

where

$$C_\sigma := \begin{cases} \sqrt{\frac{2+\sigma}{2-\sigma}}, & 0 < \sigma < 2, \\ \sqrt{\frac{\sigma+2}{\sigma-2}}, & \sigma > 2, \end{cases} \qquad f_\infty = \frac{u_{\text{avg}}}{2}, \tag{3.2.6}$$

and the decay rate $\mu(\sigma)$ is sharp.

For $\sigma = 2$ we have that for any $0 < \varepsilon < 1$

$$\left\| f(t) - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| \leq \frac{\sqrt{2}}{\varepsilon} \left\| f_0 - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| e^{-(1-\varepsilon)t}. \tag{3.2.7}$$

[2] We use *mild solution* in the termonology of semigroup theory [19].

b) If $\sigma(x)$ is non-constant such that

$$0 < \sigma_{\min} := \operatorname{ess\,inf}_{x\in\mathbb{T}} \sigma(x) < \operatorname{ess\,sup}_{x\in\mathbb{T}} \sigma(x) =: \sigma_{\max} < \infty,$$

then by defining

$$\theta^* := \min\left(\sigma_{\min}, \frac{4}{\sigma_{\max}}\right) \tag{3.2.8}$$

and

$$\alpha^*\left(\sigma_{\min},\sigma_{\max}\right) := \begin{cases} \frac{\sigma_{\min}\left(4+2\sqrt{4-\sigma_{\min}^2}-\sigma_{\min}\sigma_{\max}\right)}{4+2\sqrt{4-\sigma_{\min}^2}-\sigma_{\min}^2}, & \sigma_{\min} < \frac{4}{\sigma_{\max}}, \\ \sigma_{\max} - \sqrt{\sigma_{\max}^2 - 4}, & \sigma_{\min} \geq \frac{4}{\sigma_{\max}}, \end{cases} \tag{3.2.9}$$

we have that

$$E_{\theta^*}\left(u(t) - u_{\text{avg}}, v(t)\right) \leq E_{\theta^*}\left(u_0 - u_{\text{avg}}, v_0\right) e^{-\alpha^*(\sigma_{\min},\sigma_{\max})t}$$

for all $t \geq 0$ and as result

$$\left\| f(t) - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| \leq \sqrt{\frac{2+\theta^*}{2-\theta^*}} \left\| f_0 - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| e^{-\frac{\alpha^*(\sigma_{\min},\sigma_{\max})}{2}t}, \tag{3.2.10}$$

with f_∞ defined in (3.2.6).

Remark 3.2.3. *If we consider a sequence of relaxation functions $\sigma_n(x)$, $n \in \mathbb{N}$, satisfying the conditions of* (b)*, then for $\sigma_{\min,n} \to \sigma$, $\sigma_{\max,n} \to \sigma$ with $\sigma \neq 2$ follows*

$$\theta^* \to \min\left(\sigma, \frac{4}{\sigma}\right), \quad \text{and} \quad \alpha^* \to \begin{cases} \sigma - \sqrt{\sigma^2 - 4}, & \sigma > 2, \\ \sigma, & \sigma < 2. \end{cases}$$

Hence, we recovering the results of part (a) *of the above theorem.*

In addition, one should note that when $\sigma_{\min} > \frac{4}{\sigma_{\max}}$, we have that

$$\alpha^*\left(\sigma_{\min},\sigma_{\max}\right) = 2\mu\left(\sigma_{\max}\right),$$

where $\mu(\sigma)$ was defined in part (a) *of the Theorem. This validates the intuition that, if $\sigma_{\max}$ is "dominant", the convergence rate of the solution can be estimated using the "worst convergence rate", corresponding to $\mu(\sigma_{\max})$, of the $\sigma(x) = \sigma$ case.*

Lastly, one notices that, when $\sigma_{\min} = \frac{4}{\sigma_{\max}}$,

$$\frac{\sigma_{\min}\left(4+2\sqrt{4-\sigma_{\min}^2}-\sigma_{\min}\sigma_{\max}\right)}{4+2\sqrt{4-\sigma_{\min}^2}-\sigma_{\min}^2} = \sigma_{\max} - \sqrt{\sigma_{\max}^2 - 4},$$

which shows the continuity of α^* with respect to $\sigma_{\min}$ and $\sigma_{\max}$.

3.3 Preliminaries

In this short section we will remind the reader of a few simple properties of functions on the torus, as well as explore properties of the anti-derivative function, $\partial_x^{-1} f$, and our functional $E_\theta(f,g)$. Most of the simple proofs will be deferred to Appendix 3.B.

We begin with the well known Poincaré inequality:

Lemma 3.3.1 (Poincaré Inequality)**.** *Let $f \in H^1_{\text{per}}(\mathbb{T})$ with $f_{\text{avg}} = 0$. Then*

$$\|f\| \leq \|f'\|. \tag{3.3.1}$$

Next we focus our attention on some simple, yet crucial properties, of the anti-derivative function which was defined in (3.2.3).

Lemma 3.3.2. *Let $f \in L^1(\mathbb{T})$. Then:*

i) $\left(\partial_x^{-1} f\right)_{\text{avg}} = 0$.

ii) $\partial_x^{-1} f$ *is differentiable a.e. on* $[0,2\pi]$ *and* $\partial_x \left(\partial_x^{-1} f\right)(x) = f(x)$ *a.e.*

iii) If in addition f is differentiable we have that $\partial_x^{-1}\left(\partial_x f\right)(x) = f(x) - f_{\text{avg}}$.

iv) If, in addition, we have that $f_{\text{avg}} = 0$, then $\partial_x^{-1} f$ is a continuous function on the torus, and

$$\widehat{\partial_x^{-1} f}(k) = \begin{cases} \frac{\widehat{f}(k)}{ik}, & k \neq 0, \\ 0, & k = 0. \end{cases} \tag{3.3.2}$$

Remark 3.3.3. *An important corollary of* (ii), (iv) *and the fact that f is a function on the torus is the fact that, as long as $f_{\text{avg}} = 0$, we are allowed to use integration by parts with $\partial_x^{-1} f(x)$ on this boundaryless manifold without qualms.*

The last part of this section is dedicated to the investigation of our newly defined functional, E_θ.

Lemma 3.3.4. *Let $f, g \in L^2(\mathbb{T})$ be such that $f_{\text{avg}} = 0$ and let $\theta \in \mathbb{R}$ be given. Then the entropy $E_\theta(f,g)$, defined in* (3.2.2)*, satisfies*

$$E_\theta(f,g) \leq \left(1 + \frac{|\theta|}{2}\right)\left(\|f\|^2 + \|g\|^2\right). \tag{3.3.3}$$

If in addition $|\theta| < 2$ we have that

$$E_\theta(f,g) \geq \left(1 - \frac{|\theta|}{2}\right)\left(\|f\|^2 + \|g\|^2\right). \tag{3.3.4}$$

In particular, if $0 \leq \theta < 2$ we have that

$$\left(1 - \frac{\theta}{2}\right)\left(\|f\|^2 + \|g\|^2\right) \leq E_\theta(f,g) \leq \left(1 + \frac{\theta}{2}\right)\left(\|f\|^2 + \|g\|^2\right). \tag{3.3.5}$$

Lastly, we shall prove the following theorem, which finally brings the system (3.2.1) into play, and on which we will rely on frequently in our future estimation.

Proposition 3.3.5. *Let $u, v \in C([0,\infty); L^2(\mathbb{T}))$ be (real valued) mild solutions to* (3.2.1) *with initial datum u_0, $v_0 \in L^2(\mathbb{T})$. Then for any $\theta \in \mathbb{R}$*

$$\begin{aligned}\frac{d}{dt}E_\theta\left(u(t)-u_{\text{avg}}, v(t)\right) = -\theta\|u(t)-u_{\text{avg}}\|^2 + \frac{1}{2\pi}\int_0^{2\pi}(\theta-2\sigma(x))v(x,t)^2dx \\ +\frac{\theta}{2\pi}\int_0^{2\pi}\sigma(x)\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx - \theta\left(v(t)_{\text{avg}}\right)^2 ,\end{aligned} \tag{3.3.6}$$

where

$$u_{\text{avg}} = \frac{1}{2\pi}\int_0^{2\pi}u_0(x)dx = \frac{1}{2\pi}\int_0^{2\pi}u(x,t)dx, \quad \forall t>0. \tag{3.3.7}$$

Proof. We begin by noticing that the validity of (3.3.7) follows immediately from the fact that u is a mild solution and the conservation of mass property of the system (3.2.1). Moreover, one can see that replacing $(u(t), v(t))$ by $\left(u(t)-u_{\text{avg}}, v(t)\right)$ yields an equivalent solution (up to a constant shift in the initial data) to the system of equations, with the additional condition that the average of the first component is zero for all $t \geq 0$. With this observation in mind, we can assume without loss of generality that $u_{\text{avg}} = 0$ and continue. For the proof of (3.3.6) we first assume that (u, v) is a classical solution, pertaining to u_0, $v_0 \in H^1_{\text{per}}(\mathbb{T})$. The general result then follows by a simple density argument.

Using the Goldstein–Taylor equations we see that

$$\frac{d}{dt}\|u(t)\|^2 = 2\langle u, \partial_t u\rangle = -2\langle u, \partial_x v\rangle .$$

$$\frac{d}{dt}\|v(t)\|^2 = 2\langle v, \partial_t v\rangle = -2\langle v, \partial_x u + \sigma v\rangle .$$

Since

$$\langle u, \partial_x v\rangle + \langle v, \partial_x u\rangle = \frac{1}{2\pi}\int_0^{2\pi}\partial_x(uv)(x,t)dx = 0 ,$$

we see that

$$\frac{d}{dt}\left(\|u(t)\|^2 + \|v(t)\|^2\right) = -\frac{1}{\pi}\int_0^{2\pi}\sigma(x)v(x,t)^2dx. \tag{3.3.8}$$

We now turn our attention to the mixed term of $E_\theta(u, v)$:

$$\begin{aligned}&\frac{d}{dt}\frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1}u(x,t)v(x,t)dx \\ &= \frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1}(\partial_t u)(x,t)v(x,t)dx + \frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1}u(x,t)\partial_t v(x,t)dx \\ &= -\frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1}(\partial_x v)(x,t)v(x,t)dx - \frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1}u(x,t)[\partial_x u(x,t) + \sigma(x)v(x,t)]dx.\end{aligned}$$

Using points (ii) and (iii) of Lemma 3.3.2, together with Remark 3.3.3, we find that the above equals

$$-\frac{\theta}{2\pi}\int_0^{2\pi}\left(v(x,t)-v(t)_{\text{avg}}\right)v(x,t)dx+\frac{\theta}{2\pi}\int_0^{2\pi}u(x,t)^2dx$$
$$-\frac{\theta}{2\pi}\int_0^{2\pi}\sigma(x)\partial_x^{-1}u(x,t)v(x,t)dx.$$

Subtracting this from (3.3.8) (as there is a minus in definition (3.2.2)) yields (3.3.6). □

3.4 Constant Relaxation Function

In recent years, the Goldstein–Taylor model on $\mathbb{T}$ with constant σ was frequently tackled with a modal decomposition w.r.t. x. This approach allows for an extension to other discrete velocity models and even continuous velocities [1], but of course not to the non-homogeneous case. We briefly review some recent results: In [13, §1.4] exponential convergence was shown, but not with the sharp rate. In [1, §4.1] a hypocoercive decay estimate of the form

$$\left\|f(t)-\begin{pmatrix}f_\infty\\f_\infty\end{pmatrix}\right\|_{L^2}\le c\,e^{-\mu(\sigma)t}\left\|f_0-\begin{pmatrix}f_\infty\\f_\infty\end{pmatrix}\right\|_{L^2},$$

with the notation $f(t):=(f_+(t),f_-(t))^T$ and the sharp rate

$$\mu(\sigma):=\begin{cases}\sigma, & 0<\sigma<2,\\ \frac{\sigma}{2}-\sqrt{\frac{\sigma^2}{4}-1}, & \sigma>2,\end{cases}$$

was obtained (see also Fig. 3.4.2 below). And in [3, Th. 1.1] also the minimal constant c was provided.

In this section we will focus our attention on the (recast) Goldstein–Taylor equation with a constant relaxation rate, $\sigma(x)=\sigma$, i.e.

$$\begin{aligned}\partial_t u(x,t)&=-\partial_x v(x,t),\\ \partial_t v(x,t)&=-\partial_x u(x,t)-\sigma v(x,t)\,.\end{aligned}\tag{3.4.1}$$

While the modal analysis is straightforward, we have to show it in detail, to obtain the explicit decay rates for each Fourier mode, as well as an “optimal Lyapunov functional” for each mode — in the sense of providing the sharp decay rates. This will allow to derive our entropy functional, first on a modal level and then without modes, in terms of a pseudo-differential operator as defined in (3.2.2). As was mentioned in §3.2, this will give us intuition to the long time behaviour of the equation in several cases even when $\sigma(x)$ is not constant.

3.4.1 Fourier Analysis and the Spectral Gap

One natural way to understand the long-time behaviour of (3.4.1) relies on a simple Fourier Analysis and a hypocoercivity technique that was developed by Arnold and Erb in [6]. We begin with the former, and focus on the latter from the next subsection onwards.

Using the Fourier transform on the torus (i.e. in the spatial variables), we see that (3.4.1) is equivalent to the infinite dimensional ODE system:

$$\partial_t \begin{pmatrix} \widehat{u}(k) \\ \widehat{v}(k) \end{pmatrix} = -\underbrace{\begin{pmatrix} 0 & ik \\ ik & \sigma \end{pmatrix}}_{\boldsymbol{C}_k :=} \begin{pmatrix} \widehat{u}(k) \\ \widehat{v}(k) \end{pmatrix}, \qquad k \in \mathbb{Z}. \tag{3.4.2}$$

The eigenvalues of $\boldsymbol{C}_k$ are given by

$$\lambda_{\pm,k} := \frac{\sigma}{2} \pm \sqrt{\frac{\sigma^2}{4} - k^2}, \quad k \in \mathbb{Z},$$

and as such:

- *Invariant space:* For $k = 0$, we find that $\lambda_{-,0} = 0$ and $\lambda_{+,0} = \sigma$. In fact, as

  $$\boldsymbol{C}_0 = \begin{pmatrix} 0 & 0 \\ 0 & \sigma \end{pmatrix}, \tag{3.4.3}$$

  we can conclude immediately that $\widehat{u}(t,0) = \widehat{u_0}(0)$ and $\widehat{v}(t,0) = \widehat{v_0}(0)e^{-\sigma t}$, corresponding to the mass conservation of the original equation and the rapid decay of the difference between the masses of f_- and f_+.

- *Case I:* For $0 < |k| \leq \lfloor \frac{\sigma}{2} \rfloor$, with $\frac{\sigma}{2} \notin \mathbb{N}$, one finds two real valued eigenvectors, the minimum between is

  $$\lambda_{-,k} = \frac{\sigma}{2} - \sqrt{\frac{\sigma^2}{4} - k^2} = \frac{2k^2}{\sigma + \sqrt{\sigma^2 - 4k^2}},$$

  i.e. the long-time behaviour of $\widehat{u}(k)$ and $\widehat{v}(k)$ is controlled by $e^{-\left(\frac{\sigma}{2} - \sqrt{\frac{\sigma^2}{4} - k^2}\right)t}$.

- *Case II:* For $0 < k = \frac{\sigma}{2} \in \mathbb{N}$ the two eigenvalues coincide and are equal to $\frac{\sigma}{2}$. Hence, that eigenvalue is defective, i.e. corresponds to a Jordan block of size 2, and the long time behaviour of $\widehat{u}(k)$ and $\widehat{v}(k)$ is controlled by $(1+t)\, e^{-\frac{\sigma}{2}t}$.

- *Case III:* For $|k| > \lfloor \frac{\sigma}{2} \rfloor$, one finds two complex and conjugated eigenvalues, whose real part equals $\frac{\sigma}{2}$. We can conclude that the long-time behaviour of $\widehat{u}(k)$ and $\widehat{v}(k)$ is controlled by $e^{-\frac{\sigma}{2}t}$.

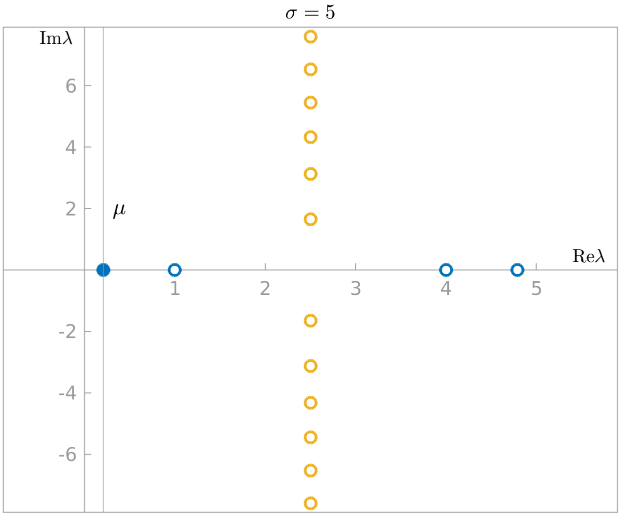


Figure 3.4.1: The eigenvalues $\lambda_{\pm,k}$ of $\boldsymbol{C}_k$, $|k| \in \mathbb{N}$ for $\sigma = 5$. The spectral gap is $\mu = (5 - \sqrt{21})/2$.

From the observations above, we notice that, as long as we subtract $\widehat{u}(0)$, i.e. as long as we remove the initial total mass from the original solution, all the modes converge *exponentially* to zero. Their rates have a sharp, and uniform-in-k lower bound that depends on σ. It will be denoted by $\mu(\sigma)$, the spectral gap of (3.4.1).

Case I, i.e. $0 < |k| < \lfloor \frac{\sigma}{2} \rfloor$, is the most "difficult case", as the real part of the eigenvalues depends on k. However, one notices that the lower eigenvalue, $\lambda_{-,k}$, increases with k, which implies that, if there are k-s such that $0 < |k| < \lfloor \frac{\sigma}{2} \rfloor$, the slowest possible convergence will be given by $\lambda_{-,\pm 1}$. As we need to compare the decay rates of *all* modes *simultaneously*, we find that it is enough to consider the following possibilities:

- $0 < \sigma < 2$: We only have Case III, i.e. all modes are controlled by $e^{-\frac{\sigma}{2}t}$.
- $\sigma = 2$: We have Case III, as well as defectiveness in $k = \pm 1$ (Case II). This means that the modes are controlled by $(1+t)\,e^{-t}$. If one searches for a pure exponential control, the best one can get is $e^{-(1-\varepsilon)t}$, for any given $\varepsilon > 0$.
- $\sigma > 2$: We have Cases I and III, and potentially Case II. All the modes that correspond to Case I are controlled by $e^{-\left(\frac{\sigma}{2}-\sqrt{\frac{\sigma 2}{4}-1}\right)t}$, while those that correspond to Case III are controlled by $e^{-\frac{\sigma}{2}t}$. If Case II is realised, i.e. $\frac{\sigma}{2} \in \mathbb{N} \setminus \{1\}$, we find that the modes $k = \pm\frac{\sigma}{2}$ are controlled by $(1+t)\,e^{-\frac{\sigma}{2}t}$. In total, thus, *all* the modes are controlled by $e^{-\left(\frac{\sigma}{2}-\sqrt{\frac{\sigma 2}{4}-1}\right)t}$, and the coefficient in the exponent is the spectral gap of the Goldstein–Taylor system (3.4.1).

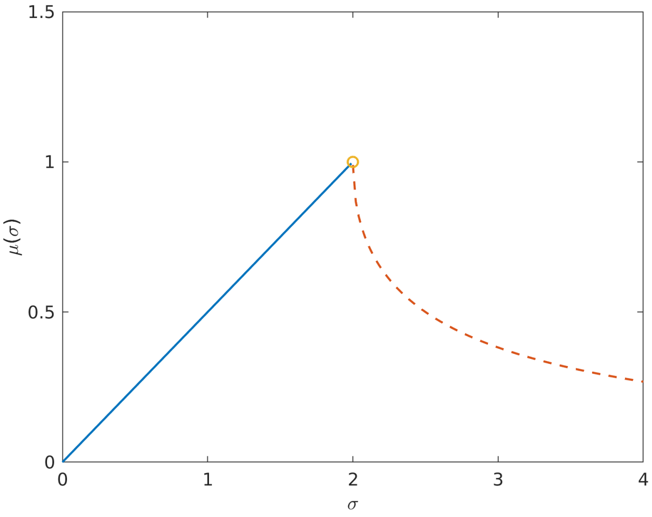


Figure 3.4.2: The exponential decay rate, $\mu(\sigma)$, of our solutions. One sees a linear growth until $\sigma = 2$ where the defectiveness appears (hence the circle). From that point onwards the decay rate decreases, and is of order $O\left(\frac{1}{\sigma}\right)$.

Before we turn our attention to properly consider these cases and "uncover" our spatial entropy, we remind the reader the hypocoercivity technique which will allow us to transform the spectral gap information of $\boldsymbol{C}_k$ into a an appropriate norm that will show the desired decay.

3.4.2 Hypocoercivity

In the previous subsection we have concluded that, barring the zero mode, all the Fourier modes of (3.4.2) decay exponentially (excluding potentially also those with $|k| = \frac{\sigma}{2}$ since the defective modes have a correction with a polynomial of degree 1). The lack of positive definiteness of the governing matrix, $\boldsymbol{C}_k$, impedes us in seeing this behaviour in the Euclidean norm on $\mathbb{C}^2$. However, an appropriately modified norm (determined by a positive definite matrix $\boldsymbol{P}_k$) can serve as a Lyaponov functional that decays with the expected rate (for $\boldsymbol{C}_k$ non-defective).

This is exactly the idea that motivated Arnold and Erb, and which is expressed in the following theorem (see [6], [1, Lemma 2]):

Theorem 3.4.1. *Let the matrix $\boldsymbol{C} \in \mathbb{C}^{n\times n}$ be positive stable (i.e. have only eigenvalues with positive real parts). Let*

$$\mu = \min\left\{\operatorname{Re}\lambda \mid \lambda \text{ is an eigenvalue of } \boldsymbol{C}\right\}.$$

Then:

i) If all eigenvalues with real part equal to μ are non-defective, there there exists a Hermitian, positive definite matrix $\boldsymbol{P}$ such that

$$\boldsymbol{C}^H\boldsymbol{P}+\boldsymbol{P}\boldsymbol{C}\geq 2\mu\boldsymbol{P}. \tag{3.4.4}$$

ii) If at least one eigenvalue with real part equal to μ is defective, then for any $\varepsilon>0$, one can find a Hermitian, positive definite matrix $\boldsymbol{P}_\varepsilon$ such that

$$\boldsymbol{C}^H\boldsymbol{P}_\varepsilon+\boldsymbol{P}_\varepsilon\boldsymbol{C}\geq 2\left(\mu-\varepsilon\right)\boldsymbol{P}_\varepsilon\,, \tag{3.4.5}$$

where $\boldsymbol{C}^H$ denotes the Hermitian transpose of $\boldsymbol{C}$.

We remark that the matrices $\boldsymbol{P}$ and $\boldsymbol{P}_\varepsilon$ are never unique. One can utilise the theorem in the following way: Assuming the eigenvalues associated to $\boldsymbol{C}$'s spectral gap, μ, are non-defective, then by defining the norm

$$\|y\|^2_{\boldsymbol{P}}:=\langle y,\boldsymbol{P}y\rangle=y^H\boldsymbol{P}y,$$

one sees that, if $y(t)$ solves the ODE $\dot{y}=-\boldsymbol{C}y$, then

$$\frac{d}{dt}\|y\|^2_{\boldsymbol{P}}=-\left\langle y,\left(\boldsymbol{C}^H\boldsymbol{P}+\boldsymbol{P}\boldsymbol{C}\right)y\right\rangle\leq-2\mu\|y\|^2_{\boldsymbol{P}}, \tag{3.4.6}$$

resulting in the correct decay rate. The same approach works in the second case of Theorem 3.4.1.

Besides the general idea of this methodology, Arnold and Erb have given a recipe (one that was later extended in [8] to defective cases, using a time dependent matrix $\boldsymbol{P}$) to finding the matrix $\boldsymbol{P},\boldsymbol{P}_\varepsilon$:

Assuming that $\boldsymbol{C}$ is diagonalisable, and letting $\{\omega_i\}_{i=1,\dots,n}$ be the eigenvectors of $\boldsymbol{C}^H$, the matrix $\boldsymbol{P}>0$ can be chosen to be[3]

$$\boldsymbol{P}=\sum_{i=1}^{n}b_i\omega_i\otimes\omega_i, \tag{3.4.7}$$

for any positive sequence $\{b_i\}_{i=1,\dots,n}$. The above formula remains true, for *a particular choice of* $\{b_i\}_{i=1,\dots,n}$, in the case where $\boldsymbol{C}$ is not diagonalisable. In that case we also need to augment the eigenvectors with the generalised eigenvectors. We refer the interested reader to Lemma 4.3 in [6]. Moreover, for $n=2$, the case we shall need below, and $\boldsymbol{C}$ non-defective, all matrices $\boldsymbol{P}$ satisfying (3.4.4) are indeed of the form (3.4.7), see [3, Lemma 3.1].

Now we return to the Fourier transformed Goldstein–Taylor system (3.4.2) to determine the modal Lyapunov functionals. A short computation, where the weights b_1, b_2 are chosen such that both diagonal elements of $\boldsymbol{P}$ are 1, finds the following matrices (For Case III we also require $b_1=b_2$, as this minimises the condition number of the resulting matrix $\boldsymbol{P}_k$ among the admissible ones.):

[3] For $v,w\in\mathbb{C}^d$ we denote $v\otimes w:=v\cdot w^H$ where $\cdot$ is the matrix-matrix multiplication.

- Case I: $0 < |k| < \lfloor \frac{\sigma}{2} \rfloor$, with $\frac{\sigma}{2} \not\in \mathbb{N}$. In this case we have:

$$\boldsymbol{P}_k^{(I)} := \begin{pmatrix} 1 & -\frac{i2k}{\sigma} \\ \frac{i2k}{\sigma} & 1 \end{pmatrix}. \tag{3.4.8}$$

- Case II: $|k| = \lfloor \frac{\sigma}{2} \rfloor \in \mathbb{N}$. As this case fosters defective eigenvalues, we will only consider the case $\sigma = 2$ (as was mentioned beforehand), and state the matrix corresponding to $k = \pm 1$:

$$\boldsymbol{P}_{\varepsilon,\pm 1}^{(II)} := \begin{pmatrix} 1 & \mp \frac{i(2-\varepsilon^2)}{2+\varepsilon^2} \\ \pm \frac{i(2-\varepsilon^2)}{2+\varepsilon^2} & 1 \end{pmatrix}. \tag{3.4.9}$$

- Case III: $|k| > \lfloor \frac{\sigma}{2} \rfloor$. In this case we have:

$$\boldsymbol{P}_k^{(III)} := \begin{pmatrix} 1 & -\frac{i\sigma}{2k} \\ \frac{i\sigma}{2k} & 1 \end{pmatrix}. \tag{3.4.10}$$

3.4.3 Derivation of the spatial entropy $E_\theta(u,v)$

The goal of this subsection is twofold: First we shall define a *modal entropy* to quantify the exponential decay of solutions to (3.4.2) towards its steady state:

$$\widehat{u_\infty}(k) = \begin{cases} \widehat{u_0}(k=0) = (u_0)_{\text{avg}}, & k = 0, \\ 0, & k \neq 0, \end{cases} \qquad \widehat{v_\infty}(k) = 0\ ,\ k \in \mathbb{Z}. \tag{3.4.11}$$

Since the matrix $\boldsymbol{C}_0$ from (3.4.3) has no spectral gap, the mode $k = 0$ plays a special role, and hence will be treated separately.

The second goal is to relate that modal-based entropy to the *spatial entropy* E_θ from Definition 3.2.1, which is not based on a modal decomposition. To this end we already remark that the off-diagonal factors ik in (3.4.8) and $\frac{1}{ik}$ in (3.4.10) correspond in physical space, roughly speaking, to a first derivative and an anti-derivative, respectively.

As in §3.4.1 we shall distinguish three cases of σ:

Case $0 < \sigma < 2$:

Then all modes $k \neq 0$ satisfy $|k| > \lfloor \frac{\sigma}{2} \rfloor$, and hence are in Case III. We recall from §3.4.1 that all modes decay here with the sharp rate $\frac{\sigma}{2}$. For a modal entropy to reflect this decay, we hence have to use for each mode a Lyapunov functional $\left\| \binom{\hat{u}(k,t)}{\hat{v}(k,t)} \right\|^2_{\boldsymbol{P}_k}$, where $\boldsymbol{P}_k$ satisfies the inequality (3.4.4) with $\mu = \frac{\sigma}{2}$. $\boldsymbol{P}_k = \boldsymbol{P}_k^{(III)}$ is the most convenient choice. Then we define the modal entropy for any $\{\widehat{u}(k), \widehat{v}(k)\}_{k \in \mathbb{Z}}$ such that $\widehat{u}(0) = 0$:

$$\begin{aligned} \mathcal{E}\,(\widehat{u}, \widehat{v}) \quad &:= \sum_{k \in \mathbb{Z} \setminus \{0\}} \left\| \begin{pmatrix} \widehat{u}(k) \\ \widehat{v}(k) \end{pmatrix} \right\|^2_{\boldsymbol{P}_k^{(III)}} + \left\| \begin{pmatrix} \widehat{u}(0) \\ \widehat{v}(0) \end{pmatrix} \right\|^2 && (3.4.12) \\ &= \sum_{k \in \mathbb{Z}} \left(|\widehat{u}(k)|^2 - \sigma \operatorname{Re}\left(\frac{\widehat{u}(k)}{ik} \overline{\widehat{v}(k)} \right) + |\widehat{v}(k)|^2 \right), && (3.4.13) \end{aligned}$$

where we used the convention $\frac{\hat{u}(0)}{0} = 0$. The mode $k = 0$ was included since $\hat{u}(0,t) = \hat{u}(0) = 0$ and $\hat{v}(0,t) = \hat{v}(0)e^{-\sigma t}$. Using Plancherel's equality, and (iv) from Lemma 3.3.2, we find that

$$\mathscr{E}\left(\hat{u},\hat{v}\right) = E_\sigma\left(u,v\right), \tag{3.4.14}$$

which shows why we consider the spatial entropy functional from Definition 3.2.1 in this case.

We note that, since $u_{\text{avg}}(t)$ is conserved, together with part (iv) of Lemma 3.3.2, explains why we have chosen to use the anti-derivative of u, and not of v.

Case $\sigma > 2$:

This situation is more complicated, as we have a mixture of the above three cases: finitely many k-s in $\mathbb{Z}$ for which $0 < |k| < \lfloor\frac{\sigma}{2}\rfloor$ (i.e. Case I), Case II for two k-s if $\frac{\sigma}{2} \in \mathbb{N}$, while the rest satisfy $|k| > \lfloor\frac{\sigma}{2}\rfloor$ (i.e. Case III). Following the above methodology to construct the modal entropy, we would need to use a combination of $\boldsymbol{P}_k^{(I)}$ and $\boldsymbol{P}_k^{(III)}$, given by (3.4.8) and (3.4.10), and potentially a matrix for the defective modes. This is feasible on the modal level, but does not easily translate back to the spatial variables. It would yield a complicated pseudo-differential operator "inside" the spatial entropy.

Recalling the discussion from §3.4.1 we see that the overall decay rate, $\mu = \frac{\sigma}{2} - \sqrt{\frac{\sigma^2}{4} - 1}$ is only determined by the modes $k = \pm 1$. Since all the other modes decay faster, we are not obliged to use "optimal" modal Lyapunov functionals for these higher modes. This gives some leeway for choosing the matrices $\boldsymbol{P}_k$, $|k| > 1$.

For $k \neq 0$ we shall use in fact the matrix

$$\boldsymbol{P}_k^{\text{suff}} := \boldsymbol{P}_k^{(III)}\left(\sigma \to \frac{4}{\sigma}\right) = \begin{pmatrix} 1 & -\frac{2i}{k\sigma} \\ \frac{2i}{k\sigma} & 1 \end{pmatrix} > 0\,, \tag{3.4.15}$$

which satisfies $\boldsymbol{P}_{\pm 1}^{\text{suff}} = \boldsymbol{P}_{\pm 1}^{(I)}$ for the crucial lowest modes. It also satisfies the following result, which implies exponential decay of all modal Lyapunov functionals $\left\|\binom{\hat{u}(k,t)}{\hat{v}(k,t)}\right\|^2_{\boldsymbol{P}_k^{\text{suff}}}$, $k \neq 0$ with rate $2\mu = \sigma - \sqrt{\sigma^2 - 4}$.

Lemma 3.4.2. *Let $\sigma > 2$. Then*

$$\boldsymbol{C}_k^H \boldsymbol{P}_k^{\text{suff}} + \boldsymbol{P}_k^{\text{suff}} \boldsymbol{C}_k - 2\mu \boldsymbol{P}_k^{\text{suff}} \geq 0, \qquad \forall\, k \neq 0\,.$$

The proof of this lemma is straightforward. Proceeding like in (3.4.12) we define the modal entropy for any $\{\hat{u}(k),\hat{v}(k)\}_{k\in\mathbb{Z}}$ such that $\hat{u}(0) = 0$:

$$\mathscr{E}\left(\hat{u},\hat{v}\right) := \sum_{k\in\mathbb{Z}\setminus\{0\}} \left\|\begin{pmatrix}\hat{u}(k)\\ \hat{v}(k)\end{pmatrix}\right\|^2_{\boldsymbol{P}_k^{\text{suff}}} + \left\|\begin{pmatrix}\hat{u}(0)\\ \hat{v}(0)\end{pmatrix}\right\|^2 .$$

Due to (3.4.14) and (3.4.15) it is related to the spatial entropy functional from Definition 3.2.1 as

$$\mathscr{E}\left(\hat{u},\hat{v}\right) = E_{\frac{4}{\sigma}}\left(u,v\right).$$

Case $\sigma = 2$:

Just like in the previous case, the lowest frequency modes, $k = \pm 1$, control the long time behaviour. However, the matrices $\boldsymbol{C}_{\pm 1}$ are now defective, which leads to a (purely) exponential decay rate reduced by ε.

We proceed similarly to the case $\sigma > 2$ and define for some $\varepsilon > 0$:

$$\boldsymbol{P}^{\text{suff}}_{\varepsilon,k} = \boldsymbol{P}^{(III)}_{k}\left(\sigma \to \frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}\right) = \begin{pmatrix} 1 & -\frac{i\left(2-\varepsilon^2\right)}{k\left(2+\varepsilon^2\right)} \\ \frac{i\left(2-\varepsilon^2\right)}{k\left(2+\varepsilon^2\right)} & 1 \end{pmatrix} > 0\,, \tag{3.4.16}$$

which satisfies $\boldsymbol{P}^{\text{suff}}_{\varepsilon,\pm 1} = \boldsymbol{P}^{(II)}_{\varepsilon,\pm 1}$ for the crucial lowest modes. It also satisfies the following result, which implies exponential decay of all modal Lyapunov functionals $\left\| \binom{\hat{u}(k,t)}{\hat{v}(k,t)} \right\|^2_{\boldsymbol{P}^{\text{suff}}_{\varepsilon,k}}$, $k \neq 0$ with rate of at least $2\mu = 2(1-\varepsilon)$.

Lemma 3.4.3. *Let $\sigma = 2$. Then*

$$\boldsymbol{C}_k^H \boldsymbol{P}^{\text{suff}}_{\varepsilon,k} + \boldsymbol{P}^{\text{suff}}_{\varepsilon,k} \boldsymbol{C}_k - 2\mu \boldsymbol{P}^{\text{suff}}_{\varepsilon,k} > 0 \qquad \forall\, k \neq 0\,.$$

Proceeding like in (3.4.12) we define the modal entropy for any $\{\widehat{u}(k), \widehat{v}(k)\}_{k\in\mathbb{Z}}$ such that $\widehat{u}(0) = 0$:

$$\mathscr{E}\left(\widehat{u}, \widehat{v}\right) := \sum_{k\in\mathbb{Z}\setminus\{0\}} \left\| \begin{pmatrix} \widehat{u}(k) \\ \widehat{v}(k) \end{pmatrix} \right\|^2_{\boldsymbol{P}^{\text{suff}}_{\varepsilon,k}} + \left\| \begin{pmatrix} \widehat{u}(0) \\ \widehat{v}(0) \end{pmatrix} \right\|^2 .$$

Due to (3.4.14) and (3.4.16) it is related to the spatial entropy functional from Definition 3.2.1 as

$$\mathscr{E}\left(\widehat{u}, \widehat{v}\right) = E_{\frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}}\left(u, v\right).$$

3.4.4 The Evolution of the Spatial Entropy

In the previous subsection we have shown how, depending on the value of σ, the entropies E_σ, $E_{\frac{4}{\sigma}}$ and $E_{\frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}}$ are the correct candidates to show the exponential convergence to equilibrium. A closer look at (3.4.6) shows that each modal Lyapunov functional $\left\| \binom{\hat{u}(k,t)}{\hat{v}(k,t)} \right\|^2_{\boldsymbol{P}_k}$ decays exponentially, and hence also the spatial entropy E_θ. Recalling the decay rates presented in §3.4.3 for the three regimes of σ, confirms that we have actually already proved most of part (a) of Theorem 3.2.2. However, as our main goal is to consider these functionals in the spatial variable alone (i.e. without a modal decomposition), we shall show how one achieves the correct convergence result following a direct calculation. This will also serve as a preparation for §3.5.

Theorem 3.4.4. *Under the same conditions of Theorem 3.2.2 with $\sigma(x) = \sigma$, one has that*

- *If $0 < \sigma < 2$ then*

$$E_\sigma\left(u(t) - u_{\text{avg}}, v(t)\right) \le E_\sigma\left(u_0 - u_{\text{avg}}, v_0\right) e^{-\sigma t}.$$

- *If* $\sigma > 2$ *then*

$$E_{\frac{4}{\sigma}}\left(u(t)-u_{\text{avg}},v(t)\right) \leq E_{\frac{4}{\sigma}}\left(u_0-u_{\text{avg}},v_0\right)e^{-\left(\sigma-\sqrt{\sigma^2-4}\right)t}.$$

- *If* $\sigma = 2$ *then for any* $0 < \varepsilon < 1$

$$E_{\frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}}\left(u(t)-u_{\text{avg}},v(t)\right) \leq E_{\frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}}\left(u_0-u_{\text{avg}},v_0\right)e^{-2(1-\varepsilon)t}.$$

Proof. Using Proposition 3.3.5, we find that:
If $0 < \sigma < 2$:

$$\begin{aligned}\frac{d}{dt}E_\sigma\left(u(t)-u_{\text{avg}},v(t)\right) &= -\sigma\|u(t)-u_{\text{avg}}\|^2-\sigma\|v(t)\|^2\\ &+\frac{\sigma^2}{2\pi}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx-\sigma\left(v(t)_{\text{avg}}\right)^2\end{aligned}$$

$$= -\sigma E_\sigma\left(u(t)-u_{\text{avg}},v(t)\right)-\sigma\left(v(t)_{\text{avg}}\right)^2 \leq -\sigma E_\sigma\left(u(t)-u_{\text{avg}},v(t)\right).$$

Note that since we know that $v_{\text{avg}(t)} = v_{0,\text{avg}}e^{-\sigma t}$ we can compute $E_\theta\left(u(t)-u_{\text{avg}},v(t)\right)$ explicitly.
If $\sigma > 2$:

$$\begin{aligned}\frac{d}{dt}E_{\frac{4}{\sigma}}\left(u(t)-u_{\text{avg}},v(t)\right) &= -\frac{4}{\sigma}\|u(t)-u_{\text{avg}}\|^2-\left(2\sigma-\frac{4}{\sigma}\right)\|v(t)\|^2\\ &+\frac{4}{2\pi}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx-\frac{4}{\sigma}\left(v(t)_{\text{avg}}\right)^2\end{aligned}$$

$$\leq -\left(\sigma-\sqrt{\sigma^2-4}\right)E_{\frac{4}{\sigma}}\left(u(t)-u_{\text{avg}},v(t)\right)+\left(\sigma-\sqrt{\sigma^2-4}-\frac{4}{\sigma}\right)\|u(t)-u_{\text{avg}}\|^2$$

$$+\left(\frac{4}{\sigma}-\sigma-\sqrt{\sigma^2-4}\right)\|v(t)\|^2+\frac{4}{2\pi}\left(1-\frac{\sigma-\sqrt{\sigma^2-4}}{\sigma}\right)\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx.$$

The desired inequality is valid if and only if

$$\begin{aligned}&\frac{4}{2\pi}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx\\ &\leq\left(\sigma-\sqrt{\sigma^2-4}\right)\|u(t)-u_{\text{avg}}\|^2+\left(\sigma+\sqrt{\sigma^2-4}\right)\|v(t)\|^2.\end{aligned} \tag{3.4.17}$$

Cauchy-Schwartz inequality, together with Poincaré inequality (Lemma 3.3.1) and Lemma 3.3.2, imply that

$$\frac{4}{2\pi}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx \leq 4\|u(t)-u_{\text{avg}}\|\|v(t)\|$$

$$= 2\left(\sqrt{\sigma - \sqrt{\sigma^2-4}}\|u(t)-u_{\text{avg}}\|\right)\left(\sqrt{\sigma + \sqrt{\sigma^2-4}}\|v(t)\|\right).$$

Together with the fact that $2|ab| \leq a^2+b^2$ shows (3.4.17) and concluding the proof in this case.

If $\sigma = 2$:

$$\begin{aligned}\frac{d}{dt}E_{\frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}}\left(u(t)-u_{\text{avg}}, v(t)\right) &= -\frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}\|u(t)-u_{\text{avg}}\|^2 - \frac{2\left(2+3\varepsilon^2\right)}{2+\varepsilon^2}\|v(t)\|^2\\ &+\frac{1}{2\pi}\cdot\frac{4\left(2-\varepsilon^2\right)}{2+\varepsilon^2}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx - \frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}\left(v(t)_{\text{avg}}\right)^2\\ &\leq -2(1-\varepsilon)E_{\frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}}\left(u(t)-u_{\text{avg}}, v(t)\right) - 2\varepsilon\left(1-\frac{2\varepsilon}{2+\varepsilon^2}\right)\|u(t)-u_{\text{avg}}\|^2\\ &-2\varepsilon\left(1+\frac{2\varepsilon}{2+\varepsilon^2}\right)\|v(t)\|^2 + \frac{1}{2\pi}\cdot\left(\frac{4\varepsilon\left(2-\varepsilon^2\right)}{2+\varepsilon^2}\right)\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx.\end{aligned}$$

Like before, the desired inequality will follow if

$$\begin{aligned}&\frac{1}{2\pi}\cdot\left(\frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}\right)\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx\\ &\leq\left(1-\frac{2\varepsilon}{2+\varepsilon^2}\right)\|u(t)-u_{\text{avg}}\|^2 + \left(1+\frac{2\varepsilon}{2+\varepsilon^2}\right)\|v(t)\|^2.\end{aligned}$$

This is valid since

$$\begin{aligned}&\frac{1}{2\pi}\cdot\left(\frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}\right)\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx\\ \leq\frac{2\sqrt{4+\varepsilon^4}}{2+\varepsilon^2}\|u(t)-u_{\text{avg}}\|\|v(t)\| &\leq 2\left(\sqrt{1-\frac{2\varepsilon}{2+\varepsilon^2}}\|u(t)-u_{\text{avg}}\|\right)\left(\sqrt{1+\frac{2\varepsilon}{2+\varepsilon^2}}\|v(t)\|\right)\\ &\leq\left(1-\frac{2\varepsilon}{2+\varepsilon^2}\right)\|u(t)-u_{\text{avg}}\|^2 + \left(1+\frac{2\varepsilon}{2+\varepsilon^2}\right)\|v(t)\|^2,\end{aligned}$$

where we used Cauchy-Schwartz inequality, Poincaré inequality, and Lemma 3.3.2 again.

The theorem is now complete. □

As the last part of this section, we finally prove part (a) of Theorem 3.2.2:

Proof of part (a) *of Theorem 3.2.2.* The decay estimates of E_θ, for the appropriate θ, follows immediately from Theorem 3.4.4. To show (3.2.5) and (3.2.7) we remind ourselves that

$$f_+ = \frac{u+v}{2}, \quad f_- = \frac{u-v}{2}$$

and

$$\|f\|^2 + \|g\|^2 \leq \frac{2}{2-\theta} E_\theta\left(g,f\right), \quad E_\theta\left(g,f\right) \leq \frac{2+\theta}{2}\left(\|f\|^2 + \|g\|^2\right)$$

for $0 < \theta < 2$ and $f_{\text{avg}} = 0$, according to Lemma 3.3.4. Thus, using the definition of f_∞ from (3.2.6) we see that

$$\begin{aligned}
&\|f_+(t) - f_\infty\|^2 + \|f_-(t) - f_\infty\|^2 \\
&= \frac{1}{2}\|u(t) - u_{\text{avg}}\|^2 + \frac{1}{2}\|v(t)\|^2 \leq \frac{1}{2-\theta} E_\theta\left(u(t) - u_{\text{avg}}, v(t)\right) \\
&\leq \frac{1}{2-\theta} E_\theta\left(u_0 - u_{\text{avg}}, v_0\right) e^{-2\mu(\sigma)t} \leq \frac{1}{2} \cdot \frac{2+\theta}{2-\theta}\left(\|u_0 - u_{\text{avg}}\|^2 + \|v_0\|^2\right) e^{-2\mu(\sigma)t} \\
&= \frac{2+\theta}{2-\theta}\left(\|f_{+,0} - f_\infty\|^2 + \|f_{-,0} - f_\infty\|^2\right) e^{-2\mu(\sigma)t},
\end{aligned}$$

which shows the result for the appropriate choices of $\theta(\sigma)$ and $\mu(\sigma)$. For $\sigma = 2$ we choose

$$\theta(2) = \frac{2\left(2-\varepsilon^2\right)}{2+\varepsilon^2}, \quad \mu(2) = 1 - \varepsilon\,.$$

The sharpness of the decay rate for $\sigma \neq 2$ can be verified easily on the first mode, e.g. for $u_0 = 0$, $v_0 = e^{ix}$. □

With the constant case fully behind us, we can now focus on the case where $\sigma(x)$ is a non-constant function.

3.5 Space-Dependent Relxation

The long-time behaviour of solutions to the Goldstein–Taylor equation (3.1.1), or equivalently its recast form (3.2.1), become increasingly harder to understand, if the relaxation function, $\sigma(x)$, is not a constant. However, as shown in §3.4, we have managed to find a potential spatial entropy, that captures the exact behaviour of the decay to equilibrium. The idea that we will employ in this section is to use the same type of entropy to try and estimate the convergence rate *even when $\sigma(x)$ is not constant.* This is, as mentioned in the introduction, a perturbative approach — yet the methodology, and ideas, are robust enough that the authors believe that it can be generalised to many other settings.

A central Lemma to establish our main result is the following:

Lemma 3.5.1. *Let $u, v \in L^2(\mathbb{T})$ be classical solutions to* (3.2.1) *with initial datum $u_0 \in L^1_+(\mathbb{T}), v_0 \in L^1(\mathbb{T})$. Denoting by $u_{\text{avg}} := (u_0)_{\text{avg}}$ we have that for any $0 < \alpha, \theta < 2$ the conditions*

$$\alpha < \theta, \quad 2\sigma_{\min} > \theta + \alpha, \tag{3.5.1}$$

$$\sup_{x\in\mathbb{T}}\left(\theta^2\left(\sigma(x) - \alpha\right)^2 - 4\left(\theta - \alpha\right)\left(2\sigma(x) - \theta - \alpha\right)\right) \leq 0, \tag{3.5.2}$$

imply that

$$E_\theta\left(u(t)-u_{\text{avg}},v(t)\right)\leq E_\theta\left(u_0-u_{\text{avg}},v_0\right)e^{-\alpha t}. \tag{3.5.3}$$

Proof. Using (3.3.6) from Proposition 3.3.5, and the fact that $\theta\left(v(t)_{\text{avg}}\right)^2\geq 0$, we find that

$$\frac{d}{dt}E_\theta\left(u(t)-u_{\text{avg}},v(t)\right)\leq -\alpha E_\theta\left(u(t)-u_{\text{avg}},v(t)\right)-(\theta-\alpha)\left\|u(t)-u_{\text{avg}}\right\|^2$$
$$-\frac{1}{2\pi}\int_0^{2\pi}(2\sigma(x)-\theta-\alpha)v(x,t)^2dx+\frac{\theta}{2\pi}\int_0^{2\pi}(\sigma(x)-\alpha)\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx.$$

The proof of the theorem will follow from the above inequality if we can show that

$$\begin{aligned}&\frac{\theta}{2\pi}\int_0^{2\pi}(\sigma(x)-\alpha)\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx\\&\leq(\theta-\alpha)\left\|u(t)-u_{\text{avg}}\right\|^2+\frac{1}{2\pi}\int_0^{2\pi}(2\sigma(x)-\theta-\alpha)v(x,t)^2dx.\end{aligned}\tag{3.5.4}$$

Due to condition (3.5.1) we have that

$$\inf_{x\in\mathbb{T}}(2\sigma(x)-\theta-\alpha)=2\sigma_{\min}-\theta-\alpha>0$$

and as such, together with Young inequality $|ab|\leq\frac{a^2}{\theta}+\frac{\theta b^2}{4}$ for any $\theta>0$, and Poincaré inequality, (3.3.1), we have that

$$\begin{aligned}&\left|\frac{\theta}{2\pi}\int_0^{2\pi}(\sigma(x)-\alpha)\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx\right|\\&\leq\frac{\theta}{2\pi}\int_0^{2\pi}\sqrt{2\sigma(x)-\theta-\alpha}\,|v(x,t)|\frac{|\sigma(x)-\alpha|}{\sqrt{2\sigma(x)-\theta-\alpha}}\left|\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)\right|dx\end{aligned}\tag{3.5.5}$$

$$\leq\frac{\theta}{2\pi}\left(\int_0^{2\pi}(2\sigma(x)-\theta-\alpha)\,v(x,t)^2dx\right)^{\frac{1}{2}}\left(\int_0^{2\pi}\frac{(\sigma(x)-\alpha)^2}{2\sigma(x)-\theta-\alpha}\left(\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)\right)^2dx\right)^{\frac{1}{2}}$$
$$\leq\frac{1}{2\pi}\int_0^{2\pi}\frac{\theta^2(\sigma(x)-\alpha)^2}{4(2\sigma(x)-\theta-\alpha)}\left(\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)\right)^2dx+\frac{1}{2\pi}\int_0^{2\pi}(2\sigma(x)-\theta-\alpha)v(x,t)^2dx$$
$$\leq\sup_{x\in\mathbb{T}}\left(\frac{\theta^2(\sigma(x)-\alpha)^2}{4(2\sigma(x)-\theta-\alpha)}\right)\left\|u(t)-u_{\text{avg}}\right\|^2+\frac{1}{2\pi}\int_0^{2\pi}(2\sigma(x)-\theta-\alpha)v(x,t)^2dx,$$

where we rewrote

$$\sigma(x)-\alpha=\sqrt{2\sigma(x)-\theta-\alpha}\cdot\frac{\sigma(x)-\alpha}{\sqrt{2\sigma(x)-\theta-\alpha}}$$

so that the term with $v(x,t)$ we obtain (with the help of Young and Poincaré inequalities) would be *exactly* the one that appears in the right hand side of (3.5.4).

The above implies that (3.5.4) will be valid when

$$\sup_{x\in\mathbb{T}} \frac{\theta^2\left(\sigma(x)-\alpha\right)^2}{4\left(2\sigma(x)-\theta-\alpha\right)} \leq \theta-\alpha,$$

which is equivalent, due to the positivity of the denominator, to (3.5.2). The proof is thus complete. □

Remark 3.5.2. *It is worth to note that the conditions expressed in* (3.5.1) *are crucial in our estimation. Indeed, they tell us that both*

$$(\theta-\alpha)\left\|u(t)-u_{\text{avg}}\right\|^2$$

and

$$\frac{1}{2\pi}\int_0^{2\pi}\left(2\sigma(x)-\theta-\alpha\right)v(x,t)^2\,dx$$

are non-negative. If one of the conditions would not be true, we would be able to cook initial data such that the mixed term is zero, and the above terms add up to something strictly negative - which in term might break the functional inequality we are aiming to attain.

The next step we consider, is to look for θ and α such that conditions (3.5.1) and (3.5.2) are satisfied and the decay rate in (3.5.3), α, is maximised.

We remind our readers the definition of θ^* from Theorem 3.2.2,

$$\theta^* = \min\left(\sigma_{\min}, \frac{4}{\sigma_{\max}}\right),$$

which in a sense captures the parameters behind the behaviour when $\sigma(x)$ is a constant which is not 2. With this at hand we have the following:

Lemma 3.5.3. *Assume that* $0<\sigma_{\min}<\sigma_{\max}<\infty$*, where* $\sigma_{\min}$ *and* $\sigma_{\max}$ *were defined in Theorem 3.2.2. Then*

$$\alpha^*\left(\sigma_{\min},\sigma_{\max}\right) := \begin{cases} \frac{\sigma_{\min}\left(4+2\sqrt{4-\sigma_{\min}^2}-\sigma_{\min}\sigma_{\max}\right)}{4+2\sqrt{4-\sigma_{\min}^2}-\sigma_{\min}^2}, & \sigma_{\min}<\frac{4}{\sigma_{\max}}, \\ \sigma_{\max}-\sqrt{\sigma_{\max}^2-4}, & \sigma_{\min}\geq\frac{4}{\sigma_{\max}}, \end{cases}$$

is such that θ^* *and* $\alpha^*\left(\sigma_{\min},\sigma_{\max}\right)$ *satisfy conditions* (3.5.1) *and* (3.5.2).

Proof. Clearly, since

$$\theta^* \leq \begin{cases} \sigma_{\min}, & \sigma_{\min}<\sigma_{\max}\leq 2, \\ \frac{4}{\sigma_{\max}}, & \sigma_{\max}>2, \end{cases}$$

we find that $0<\theta^*<2$.

We continue by considering condition (3.5.2), and constructing parameters which will give condition (3.5.1) automatically. Denoting by

$$f\left(\alpha,\theta,y\right) := \theta^2\left(y-\alpha\right)^2 - 4\left(\theta-\alpha\right)\left(2y-\theta-\alpha\right)$$

for (α,θ) that satisfy condition (3.5.1) and $y \in [\sigma_{\min},\sigma_{\max}]$, we find that for a fixed α and θ, f is an upward parabola in y whose non-positive part lies between its roots

$$y_{\pm}\left(\alpha,\theta\right) := \alpha + \frac{2\left(\theta-\alpha\right)}{\theta^2}\left(2\pm\sqrt{4-\theta^2}\right).$$

Thus, for condition (3.5.2) to be satisfied we need that

$$y_{-}\left(\alpha,\theta\right) \le \sigma_{\min}, \quad \text{and} \quad \sigma_{\max} \le y_{+}\left(\alpha,\theta\right).$$

A simple calculation, using the fact that for $0<\theta<2$

$$2\sqrt{4-\theta^2} > 4-\theta^2,$$

shows that for a fixed θ

$$y_{-}\left(\alpha,\theta\right) \le \sigma_{\min} \iff \alpha \le \frac{\theta\left(2\sqrt{4-\theta^2}-\left(4-\sigma_{\min}\theta\right)\right)}{2\sqrt{4-\theta^2}-\left(4-\theta^2\right)} = \gamma_{\min}\left(\theta\right),$$

$$\sigma_{\max} \le y_{+}\left(\alpha,\theta\right) \iff \alpha \le \frac{\theta\left(2\sqrt{4-\theta^2}+\left(4-\sigma_{\max}\theta\right)\right)}{2\sqrt{4-\theta^2}+\left(4-\theta^2\right)} = \gamma_{\max}\left(\theta\right).$$

Thus, if we choose $\alpha\left(\theta\right)$ for a fixed θ so that condition (3.5.2) is valid, we must have that

$$\alpha\left(\theta\right) \le \min\left(\gamma_{\min}\left(\theta\right),\gamma_{\max}\left(\theta\right)\right).$$

To continue and motivate our choice we show next that $\gamma_{\max}\left(\theta\right) \le \gamma_{\min}\left(\theta\right)$. Indeed, this is valid if and only if

$$\frac{2\sqrt{4-\theta^2}+\left(4-\sigma_{\max}\theta\right)}{2+\sqrt{4-\theta^2}} \le \frac{2\sqrt{4-\theta^2}-\left(4-\sigma_{\min}\theta\right)}{2-\sqrt{4-\theta^2}}$$

which is equivalent to

$$2\left(8-\theta\left(\sigma_{\min}+\sigma_{\max}\right)\right) + \sqrt{4-\theta^2}\,\theta\left(\sigma_{\max}-\sigma_{\min}\right) \le 4\left(4-\theta^2\right),$$

or

$$\frac{2\left(\sigma_{\max}+\sigma_{\min}-2\theta\right)}{\sigma_{\max}-\sigma_{\min}} \ge \sqrt{4-\theta^2},$$

an inequality that is satisfied when $\theta \leq \min\left(\sigma_{\min}, 2\right)$.[4]

Since, in addition, for any $\theta \leq \frac{4}{\sigma_{\max}}$ with $\theta < 2$ we have that

$$\gamma_{\max}\left(\theta\right) \leq \frac{2\theta\sqrt{4-\theta^2}}{2\sqrt{4-\theta^2}+\left(4-\theta^2\right)} < \theta,$$

we can deduce that for any $\theta \in \left(0,\theta^*\right] \subset \left(0, \min\left(\sigma_{\min}, \frac{4}{\sigma_{\max}}, 2\right)\right]$

$$\gamma_{\max}(\theta) = \min\left(\gamma_{\min}(\theta), \gamma_{\max}(\theta)\right), \qquad \gamma_{\max}\left(\theta\right) < \theta,$$

and as such the pair $\left(\theta, \gamma_{\max}(\theta)\right)$ satisfies condition (3.5.2) as well as

$$\gamma_{\max}(\theta) + \theta < 2\theta \leq 2\theta^* \leq 2\sigma_{\min}.$$

We conclude that θ and $\gamma_{\max}\left(\theta\right)$ satisfy both desired conditions, for any $\theta \in \left(0,\theta^*\right]$.

Aiming to maximise $\gamma_{\max}\left(\theta\right)$, which will correspond to our desired decay rate of Lemma 3.5.1, on $\left(0,\theta^*\right]$ we notice that

$$\frac{d}{d\theta}\left(\theta\left(2\sqrt{4-\theta^2}+4-\sigma_{\max}\theta\right)\right) = \frac{2}{\sqrt{4-\theta^2}}\left(4-2\theta^2+\left(2-\sigma_{\max}\theta\right)\sqrt{4-\theta^2}\right)$$

$$\frac{d}{d\theta}\left(2\sqrt{4-\theta^2}+4-\theta^2\right) = -\frac{2\theta}{\sqrt{4-\theta^2}}\left(1+\sqrt{4-\theta^2}\right)$$

and as such

$$\begin{aligned}\frac{d}{d\theta}\gamma_{\max}\left(\theta\right) &= \frac{2\left(4-2\theta^2+\left(2-\sigma_{\max}\theta\right)\sqrt{4-\theta^2}\right)\left(2\sqrt{4-\theta^2}+4-\theta^2\right)}{\left(4-\theta^2\right)^{\frac{3}{2}}\left(2+\sqrt{4-\theta^2}\right)^2}\\ &\quad+\frac{2\theta^2\left(2\sqrt{4-\theta^2}+4-\sigma_{\max}\theta\right)\left(1+\sqrt{4-\theta^2}\right)}{\left(4-\theta^2\right)^{\frac{3}{2}}\left(2+\sqrt{4-\theta^2}\right)^2}\\ &= \frac{\left(8-2\sigma_{\max}\theta\right)\left(8+4\sqrt{4-\theta^2}-\theta^2\right)}{\left(4-\theta^2\right)^{\frac{3}{2}}\left(2+\sqrt{4-\theta^2}\right)^2} = \frac{8-2\sigma_{\max}\theta}{\left(4-\theta^2\right)^{\frac{3}{2}}}.\end{aligned}$$

Thus, $\gamma_{\max}\left(\theta\right)$ increases in the domain $\theta \in \left(0,\theta^*\right] \subset \left(0, \frac{4}{\sigma_{\max}}\right]$.

Defining

$$\alpha^*\left(\sigma_{\min},\sigma_{\max}\right) := \max_{\theta\in\left(0,\theta^*\right]}\gamma_{\max}\left(\theta\right) = \gamma_{\max}\left(\theta^*\right)$$

[4]

$$\frac{2\left(\sigma_{\max}+\sigma_{\min}-2\theta\right)}{\sigma_{\max}-\sigma_{\min}} \geq \frac{2\left(\sigma_{\max}-\sigma_{\min}\right)}{\sigma_{\max}-\sigma_{\min}} = 2 \geq \sqrt{4-\theta^2}.$$

we find that the desired conditions are satisfied and

$$\alpha^*\left(\sigma_{\min},\sigma_{\max}\right)=\begin{cases}\frac{\sigma_{\min}\left(2\sqrt{4-\sigma_{\min}^2}+(4-\sigma_{\max}\sigma_{\min})\right)}{2\sqrt{4-\sigma_{\min}^2}+\left(4-\sigma_{\min}^2\right)}, & \sigma_{\min}<\frac{4}{\sigma_{\max}},\\ \frac{\frac{8}{\sigma_{\max}}\sqrt{4-\frac{16}{\sigma_{\max}^2}}}{2\sqrt{4-\frac{16}{\sigma_{\max}^2}}+4-\frac{16}{\sigma_{\max}^2}}, & \sigma_{\min}>\frac{4}{\sigma_{\max}},\end{cases}$$

$$=\begin{cases}\frac{\sigma_{\min}\left(2\sqrt{4-\sigma_{\min}^2}+(4-\sigma_{\max}\sigma_{\min})\right)}{2\sqrt{4-\sigma_{\min}^2}+\left(4-\sigma_{\min}^2\right)}, & \sigma_{\min}<\frac{4}{\sigma_{\max}},\\ \frac{4}{\sigma_{\max}+\sqrt{\sigma_{\max}^2-4}}, & \sigma_{\min}>\frac{4}{\sigma_{\max}},\end{cases}$$

which is exactly the formula given in the Lemma. The proof is thus complete. □

Remark 3.5.4. *It is worth to note that we could have chosen $\theta=\theta^*$ in the above proof, without considering the derivative of $\gamma_{max}(\theta)$. We have elected to consider it, though, to show that $\theta=\theta^*$ is the optimal choice (in the sense of getting the best α for Lemma 3.5.1), following this methodology.*

In addition, following the last statement of Remark 3.2.3, we see that at the boundary case of $\sigma_{\min}=\frac{4}{\sigma_{\max}}$ there is no ambiguity in the choice of α^* in our proof.

We now posses all the tools which are required to prove part (b) of Theorem 3.2.2.

Proof of part (b) *of Theorem 3.2.2.* The convergence estimation for $E_{\theta^*}\left(u(t)-u_{\text{avg}},v(t)\right)$ follows immediately from Lemma 3.5.1 and Lemma 3.5.3. To obtain (3.2.10) we use Lemma 3.3.4 in a similar fashion to the way we proved part (a). □

3.6 Convergence to Equilibrium in a Three Velocity Goldstein–Taylor Model

The Goldstein–Taylor model can be thought of as a simplification of a BGK system

$$\partial_t f(x,v,t)+v\cdot\nabla_x f(x,v,t)-\nabla_x V(x)\cdot\nabla_v f(x,v,t)=\boldsymbol{M}_{T(t)}\int f(x,v,t)dv-f(x,v,t)$$

in the discrete velocity space $v\in\{v_1,\dots,v_n\}$, with $x\in\mathbb{T}$, $V(x)=0$ and $\boldsymbol{M}_{T(t)}$ a constant matrix that speaks of the long-time behaviour. Under the natural physical assumption of the conservation of momentum, i.e. $\sum_{i=1}^n v_i=0$, and the expectation that the equilibrium state would be *equally distributed* and constant[5], we recover the general multi-

[5] The motivation for this is the two velocity Goldstein–Taylor model, where we expect the velocities distributions to behave the same. If one wants to approximate the BGK equation on $\mathbb{R}^d$ with a Maxwellian as a velocity distribution, M must be chosen as a discretisation of it, which yields unequal distribution of the appropriate equilibrium states.

velocity Goldstein–Taylor on $\mathbb{T} \times (0,\infty)$:

$$\partial_t f_i(x,t) + v_i f_i(x,t) = \sigma(x)\left(\begin{pmatrix}\frac{1}{n}\\ \vdots \\ \frac{1}{n}\end{pmatrix} \otimes (1,\dots,1) - \boldsymbol{I}\right)\begin{pmatrix}f_1(x,t)\\ \vdots \\ f_n(x,t)\end{pmatrix} \tag{3.6.1}$$

where we have added a relaxation rate, $\sigma(x)$, to the "collision side", and where

$$\{v_1,\dots,v_n\} = \begin{cases}\{-k,\dots,-1,1,\dots,k\}, & n=2k,\\ \{-k,\dots,-1,0,1,\dots,k\}, & n=2k-1,\end{cases} \qquad n\in\mathbb{N},\ n\geq 2.$$

A careful look shows that

$$\begin{pmatrix}\frac{1}{n}\\ \vdots \\ \frac{1}{n}\end{pmatrix} \otimes (1,\dots,1) - \boldsymbol{I} = \frac{1}{n}\begin{pmatrix}1-n & 1 & \dots & 1\\ 1 & 1-n & \dots & 1\\ \vdots & \vdots & \vdots & \vdots\\ 1 & 1 & \dots & 1-n\end{pmatrix}$$

which has $\boldsymbol{\xi_1} = (1,1,\dots,1)^T$ in its kernel, and $\mathscr{A} = \left\{(\xi_1,\dots,\xi_n)^T \in \mathbb{R}^n \mid \sum_{i=1}^n \xi_i = 0\right\}$ as its $n-1$ dimensional eigenspace corresponding to the eigenvalue $\lambda = -1$. This corresponds to the conservation of total mass, and the fact that the differences $\left\{f_i - f_j\right\}_{i,j=1,\dots,n}$ converge to zero. For more information we refer the interested reader to [1].

In our section we shall consider a simple three velocity Goldstein–Taylor model, which is governed by the following system of equations on $\mathbb{T} \times (0,\infty)$

$$\begin{aligned}\partial_t f_1(x,t) + \partial_x f_1(x,t) &= \frac{\sigma(x)}{3}\left(f_2(x,t) + f_3(x,t) - 2f_1(x,t)\right),\\ \partial_t f_2(x,t) &= \frac{\sigma(x)}{3}\left(f_1(x,t) + f_3(x,t) - 2f_2(x,t)\right),\\ \partial_t f_3(x,t) - \partial_x f_3(x,t) &= \frac{\sigma(x)}{3}\left(f_1(x,t) + f_2(x,t) - 2f_3(x,t)\right).\end{aligned} \tag{3.6.2}$$

Much like our Goldstein–Taylor equation, (3.1.1), we can recast the above with the variables

$$u_1 = f_1 + f_2 + f_3, \quad u_2 = f_1 - f_3, \quad u_3 = f_1 + f_3 - 2f_2,$$

and obtain the system

$$\begin{aligned}\partial_t u_1(x,t) + \partial_x u_2(x,t) &= 0,\\ \partial_t u_2(x,t) + \frac{1}{3}\partial_x\left(2u_1(x,t) + u_3(x,t)\right) &= -\sigma(x)u_2(x,t),\\ \partial_t u_3(x,t) + \partial_x u_2(x,t) &= -\sigma(x)u_3(x).\end{aligned} \tag{3.6.3}$$

Following our intuition we expect that by denoting

$$u_\infty := \frac{1}{2\pi}\int_{\mathbb{T}}\left(f_{1,0}(x) + f_{2,0}(x) + f_{3,0}(x)\right)dx,$$

we will find that

$$u_1(t,x) \xrightarrow[t\to\infty]{} u_\infty, \quad u_2(t,x) \xrightarrow[t\to\infty]{} 0, \quad u_3(t,x) \xrightarrow[t\to\infty]{} 0.$$

The case $\sigma(x) = \sigma > 0$ yields some fairly simple ODEs for $m_i(t) = \int_{\mathbb{T}} u_i(x,t)dx$, which confirm the above.

Looking at (3.6.3), we see that the first two equations are similar to (3.2.1), though with an additional "mixed term" and a different "weight" for $\partial_x u_1$. This is the intuition behind the following theorem:

Theorem 3.6.1. *Let u_1, u_2 and u_3 be classical solutions to* (3.6.3) *with initial datum $u_{1,0}, u_{2,0}, u_{3,0} \in L^1_+(\mathbb{T})$. Denoting by*

$$\mathfrak{E}_\theta\left(f,g,h\right) := \|f\|^2 + \frac{3}{2}\|g\|^2 + \frac{1}{2}\|h\|^2 - \frac{\theta}{2\pi}\int_0^{2\pi} \operatorname{Re}\left(\partial_x^{-1} f(x)\overline{g(x)}\right) dx,$$

we have that

$$\mathfrak{E}_\theta\left(u_1(t) - u_\infty, u_2(t), u_3(t)\right) \le \mathfrak{E}_\theta\left(u_{1,0} - u_\infty, u_{2,0}, u_{3,0}\right) e^{-\alpha t} \tag{3.6.4}$$

for any $\theta > 0$ and $\alpha > 0$ such that

$$\theta + \frac{3\alpha}{2} < 3\sigma_{\min}, \quad \alpha < 2\sigma_{\min}, \quad \alpha \le \frac{2\theta}{3}, \tag{3.6.5}$$

and

$$\left(\sup_{x\in\mathbb{T}} \frac{\theta^2\left(\sigma(x)-\alpha\right)^2}{12\sigma(x) - 4\theta - 6\alpha}\right) + \left(\sup_{x\in\mathbb{T}} \frac{\theta^2}{18\left(2\sigma(x)-\alpha\right)}\right) \le \left(\frac{2\theta}{3} - \alpha\right). \tag{3.6.6}$$

Remark 3.6.2. *It is important to note that the idea that guides us in defining $\mathfrak{E}_\theta$ is similar to that that helped us find E_θ from* (3.2.2)*. However, the norms of the functions are weighted differently, which corresponds, in part, to the different weighting of the transport parts of the system* (3.6.3).

Proof. We start by noticing that the transformation

$$u_1 \to u_1 - u_\infty, \quad u_2 \to u_2, \quad u_3 \to u_3$$

keeps (3.6.3) invariant, so we may assume, without loss of generality, that $u_\infty = 0$. This, together with the equation for $u_1(x,t)$ implies that

$$\left(u_1(t)\right)_{\text{avg}} = \left(u_{1,0}\right)_{\text{avg}} = u_\infty = 0.$$

Next, we compute the time derivatives of the L^2 norms:

$$\frac{d}{dt}\|u_1(t)\|^2 = 2\left\langle u_1, \partial_t u_1\right\rangle = -2\left\langle u_1, \partial_x u_2\right\rangle,$$

$$\frac{d}{dt}\|u_2(t)\|^2 = -\frac{4}{3}\langle u_2, \partial_x u_1\rangle - \frac{2}{3}\langle u_2, \partial_x u_3\rangle - 2\langle u_2, \sigma u_2\rangle,$$

and

$$\frac{d}{dt}\|u_3(t)\|^2 = -2\langle u_3, \partial_x u_2\rangle - 2\langle u_3, \sigma u_3\rangle.$$

Thus

$$\begin{aligned}\frac{d}{dt}\left(\|u_1(t)\|^2 + \frac{3}{2}\|u_2(t)\|^2 + \frac{1}{2}\|u_3(t)^2\|\right) = -\frac{3}{2\pi}\int_0^{2\pi}\sigma(x)u_2(x,t)^2 dx \\ -\frac{1}{2\pi}\int_0^{2\pi}\sigma(x)u_3(x,t)^2 dx\end{aligned} \tag{3.6.7}$$

Next, we see that

$$\begin{aligned}\frac{d}{dt}\int_0^{2\pi}\partial_x^{-1}u_1(x,t)u_2(x,t)dx = -\int_0^{2\pi}\partial_x^{-1}\left(\partial_x u_2\right)(x,t)u_2(x,t)dx \\ -\frac{1}{3}\int_0^{2\pi}\partial_x^{-1}u_1(x,t)\partial_x\left(2u_1(x,t)+u_3(x,t)\right)dx - \int_0^{2\pi}\sigma(x)\partial_x^{-1}u_1(x,t)u_2(x,t)dx \\ = 2\pi\left((u_2(t))^2_{\text{avg}} - \|u_2(t)\|^2\right) + \frac{4\pi}{3}\|u_1(t)\|^2 \\ +\frac{1}{3}\int_0^{2\pi}u_1(x,t)u_3(x,t)dx - \int_0^{2\pi}\sigma(x)\partial_x^{-1}u_1(x,t)u_2(x,t)dx\end{aligned}$$

where we used Lemma 3.3.2. As such, we find that

$$\begin{aligned}\frac{d}{dt}\left(-\frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1}u_1(x,t)u_2(x,t)\right)dx = \theta\|u_2(t)\|^2 - \theta\left(u_2(t)\right)^2_{\text{avg}} \\ -\frac{2\theta}{3}\|u_1(t)\|^2 - \frac{\theta}{6\pi}\int_0^{2\pi}u_1(x,t)u_3(x,t)dx + \frac{\theta}{2\pi}\int_0^{2\pi}\sigma(x)\partial_x^{-1}u_1(x,t)u_2(x,t)dx,\end{aligned}$$

from which, together with (3.6.7), we conclude that

$$\begin{aligned}\frac{d}{dt}\mathfrak{E}_\theta\left(u_1(t),u_2(t),u_3(t)\right) = -\frac{1}{2\pi}\int_0^{2\pi}\left(3\sigma(x)-\theta\right)u_2(x,t)^2dx \\ -\frac{1}{2\pi}\int_0^{2\pi}\sigma(x)u_3(x,t)^2dx - \frac{2\theta}{3}\|u_1(t)\|^2 - \theta\left(u_2(t)\right)^2_{\text{avg}} \\ -\frac{\theta}{6\pi}\int_0^{2\pi}u_1(x,t)u_3(x,t)dx + \frac{\theta}{2\pi}\int_0^{2\pi}\sigma(x)\partial_x^{-1}u_1(x,t)u_2(x,t)dx,\end{aligned} \tag{3.6.8}$$

and as such

$$\frac{d}{dt}\mathfrak{E}_\theta\left(u_1(t),u_2(t),u_3(t)\right) = -\alpha\mathfrak{E}_\theta\left(u_1(t),u_2(t),u_3(t)\right) + R_{\theta,\alpha,\sigma}(t)$$

$$
\begin{aligned}
R_{\theta,\alpha,\sigma}(t) = &-\frac{1}{2\pi}\int_0^{2\pi}\left(3\sigma(x)-\theta-\frac{3\alpha}{2}\right)u_2(x,t)^2dx \\
-\frac{1}{2\pi}\int_0^{2\pi}&\left(\sigma(x)-\frac{\alpha}{2}\right)u_3(x,t)^2dx-\left(\frac{2\theta}{3}-\alpha\right)\|u_1(t)\|^2-\theta\,(u_2(t))^2_{\text{avg}} \\
-\frac{\theta}{6\pi}\int_0^{2\pi}&u_1(x,t)u_3(x,t)dx+\frac{\theta}{2\pi}\int_0^{2\pi}(\sigma(x)-\alpha)\,\partial_x^{-1}u_1(x,t)u_2(x,t)dx,
\end{aligned}
\tag{3.6.9}
$$

To conclude the proof it is enough to show that under conditions (3.6.5) and (3.6.6) we have that $R_{\theta,\alpha,\sigma}(t)\leq 0$. A stronger statement, which we will prove, is that

$$
\begin{aligned}
&\left|-\frac{\theta}{6\pi}\int_0^{2\pi}u_1(x,t)u_3(x,t)dx+\frac{\theta}{2\pi}\int_0^{2\pi}(\sigma(x)-\alpha)\,\partial_x^{-1}u_1(x,t)u_2(x,t)dx\right| \\
\leq\ &\frac{1}{2\pi}\int_0^{2\pi}\left(3\sigma(x)-\theta-\frac{3\alpha}{2}\right)u_2(x,t)^2dx+\frac{1}{2\pi}\int_0^{2\pi}\left(\sigma(x)-\frac{\alpha}{2}\right)u_3(x,t)^2dx+\left(\frac{2\theta}{3}-\alpha\right)\|u_1(t)\|^2.
\end{aligned}
\tag{3.6.10}
$$

Similar to the techniques we've used in the proof of part (b) of Theorem 3.2.2, and using the positivity of appropriate functions that follows from (3.6.5), we see that

$$
\begin{aligned}
&\left|\frac{\theta}{2\pi}\int_0^{2\pi}(\sigma(x)-\alpha)\,\partial_x^{-1}u_1(x,t)u_2(x,t)dx\right| \\
\leq\ &\frac{\theta}{2\pi}\int_0^{2\pi}\frac{|\sigma(x)-\alpha|}{\sqrt{3\sigma(x)-\theta-\frac{3\alpha}{2}}}\left|\partial_x^{-1}u_1(x,t)\right|\cdot\sqrt{3\sigma(x)-\theta-\frac{3\alpha}{2}}\,|u_2(x,t)|\,dx \\
\leq\ &\left(\sup_{x\in\mathbb{T}}\frac{\theta^2\,(\sigma(x)-\alpha)^2}{12\sigma(x)-4\theta-6\alpha}\right)\|u_1(t)\|^2+\frac{1}{2\pi}\int_0^{2\pi}\left(3\sigma(x)-\theta-\frac{3\alpha}{2}\right)u_2(x,t)^2dx,
\end{aligned}
$$

and that

$$
\begin{aligned}
\left|\frac{\theta}{6\pi}\int_0^{2\pi}u_1(x,t)u_3(x,t)dx\right|&\leq\frac{\theta}{2\pi}\int_0^{2\pi}\frac{|u_1(x,t)|}{3\sqrt{\sigma(x)-\frac{\alpha}{2}}}\sqrt{\sigma(x)-\frac{\alpha}{2}}\,|u_3(x,t)|\,dx \\
&\leq\left(\sup_{x\in\mathbb{T}}\frac{\theta^2}{18\,(2\sigma(x)-\alpha)}\right)\|u_1(t)\|^2+\frac{1}{2\pi}\int_0^{2\pi}\left(\sigma(x)-\frac{\alpha}{2}\right)u_3(x,t)^2dx.
\end{aligned}
$$

Thus, one sees that (3.6.10) holds when

$$
\left(\sup_{x\in\mathbb{T}}\frac{\theta^2\,(\sigma(x)-\alpha)^2}{12\sigma(x)-4\theta-6\alpha}\right)+\left(\sup_{x\in\mathbb{T}}\frac{\theta^2}{18\,(2\sigma(x)-\alpha)}\right)\leq\left(\frac{2\theta}{3}-\alpha\right),
$$

which is (3.6.6). The proof is complete. □

While we have elected to not optimise the choice of α (like in §3.5), we can still infer the following, simpler yet far from optimal, corollary:

Corollary 3.6.3. Let $\theta > 0$ and $\alpha > 0$ be such that

$$\theta + \frac{3\alpha}{2} < 3\sigma_{\min}, \quad \alpha < 2\sigma_{\min}, \quad \alpha \leq \frac{2\theta}{3},$$

and

$$\frac{\theta^2\sigma_{\max}^2}{12\sigma_{\min} - 4\theta - 6\alpha} + \frac{\theta^2}{18\left(2\sigma_{\min} - \alpha\right)} \leq \left(\frac{2\theta}{3} - \alpha\right)$$

then

$$\mathfrak{E}_\theta\left(u_1(t) - u_\infty, u_2(t), u_3(t)\right) \leq \mathfrak{E}_\theta\left(u_{1,0} - u_\infty, u_{2,0}, u_{3,0}\right)e^{-\alpha t}.$$

In particular, for

$$\alpha^* := \min\left(\frac{\sigma_{\min}}{2}, \frac{3\sigma_{\min}}{9\sigma_{\max}^2 + 1}\right)$$

we have that $\theta_{\alpha^*} := 3\alpha^*$ and α^* satisfy the above requirements, and as such $\mathfrak{E}_{3\alpha^*}$ decays exponentially to zero with rate α^*.

Proof. Since $\alpha < 2\sigma_{\min} \leq \sigma_{\max} + \sigma_{\min}$ we see that

$$\alpha - \sigma_{\max} < \sigma_{\min} \leq \sigma(x) < \sigma_{\max} + \alpha.$$

Thus

$$\left(\sigma(x) - \alpha\right)^2 \leq \sigma_{\max}^2.$$

Using the above, with additional elementary estimation on the denominator of the expressions that appear in (3.6.6), we conclude that if

$$\frac{\theta^2\sigma_{\max}^2}{12\sigma_{\min} - 4\theta - 6\alpha} + \frac{\theta^2}{18\left(2\sigma_{\min} - \alpha\right)} \leq \left(\frac{2\theta}{3} - \alpha\right)$$

then (3.6.6) is valid. This, together with Theorem 3.6.1, shows the first statement of the corollary.

To show the second part we notice that with the choice $\theta_{\alpha^*} = 3\alpha^*$ and $\alpha^* \leq \frac{\sigma_{\min}}{2}$

$$\theta_{\alpha^*} + \frac{3\alpha^*}{2} = \frac{9\alpha^*}{2} \leq \frac{9\sigma_{\min}}{4} < 3\sigma_{\min}, \quad \alpha^* < \frac{\sigma_{\min}}{2} < 2\sigma_{\min}, \quad \alpha^* \leq 2\alpha^* = \frac{2\theta_{\alpha^*}}{3}.$$

Thus, in this case, the first condition of the corollary hold. Plugging θ_α^* in the second condition, and using the fact that

$$12\sigma_{\min} - 4\theta_{\alpha^*} - 6\alpha^* = 12\sigma_{\min} - 18\alpha^* \geq 3\sigma_{\min}$$

and

$$2\sigma_{\min} - \alpha^* \geq \frac{3}{2}\sigma_{\min}$$

when $\alpha^* \le \frac{\sigma_{\mathrm{min}}}{2}$, we see that in that case

$$\frac{\theta_{\alpha^*}^2 \sigma_{\mathrm{max}}^2}{12\sigma_{\mathrm{min}} - 4\theta_{\alpha^*} - 6\alpha^*} + \frac{\theta_{\alpha^*}^2}{18\,(2\sigma_{\mathrm{min}} - \alpha^*)} \le \left(9\sigma_{\mathrm{max}}^2 + 1\right) \frac{\alpha^{*2}}{3\sigma_{\mathrm{min}}}.$$

Thus, since $\frac{2\theta_\alpha^*}{3} - \alpha^* = \alpha^*$, our desired condition is valid when

$$\alpha^* \le \frac{3\sigma_{\mathrm{min}}}{9\sigma_{\mathrm{max}}^2 + 1},$$

which concludes the proof. □

Appendix

3.A Lack of Optimality

In this appendix we will briefly, and more formally than not, discuss the optimality of our main theorem by comparing the result we obtained to the optimal exponential rate of convergence to the Goldstein–Taylor equation, found in [9]. In fact, we aim to do a bit more. We will show how in some simple cases, our general methodology can be improved, and continue to show that even this improved bound is less than that given in [9].

To set the scene, we mention that the relaxation function we will eventually explore will be:

$$\sigma(x) = \begin{cases} 1, & 0 \leq x \leq \pi, \\ 4, & \pi < x \leq 2\pi. \end{cases} \tag{3.A.1}$$

The choice of 1 and 4 as the particular constants for $\sigma(x)$ is motivated by the fact that in this case $\sigma_{\min} = \frac{\sigma_{\max}}{4}$, and so our choice of $\theta^* = 1$ in our main theorem comes "from both directions".

Before we start with a more structured discussion, we would like to motivate our idea of how to improve the technique we developed in §3.5. A crucial point in the proof of the differential equation that governs the behaviour of E_{θ^*}, expressed in part (b) of Theorem 3.2.2, was the estimation:

$$\begin{aligned} &\left| \frac{\theta}{2\pi} \int_0^{2\pi} (\sigma(x) - \alpha) \partial_x^{-1} \left(u(x,t) - u_{\text{avg}} \right) v(x,t) dx \right| \\ &\leq \frac{1}{2\pi} \int_0^{2\pi} \frac{\theta^2 (\sigma(x) - \alpha)^2}{4(2\sigma(x) - \theta - \alpha)} \left(\partial_x^{-1} \left(u(x,t) - u_{\text{avg}} \right) \right)^2 dx + \frac{1}{2\pi} \int_0^{2\pi} (2\sigma(x) - \theta - \alpha) v(x,t)^2 dx \\ &\leq \sup_{x \in \mathbb{T}} \left(\frac{\theta^2 (\sigma(x) - \alpha)^2}{4(2\sigma(x) - \theta - \alpha)} \right) \| u(t) - u_{\text{avg}} \|^2 + \frac{1}{2\pi} \int_0^{2\pi} (2\sigma(x) - \theta - \alpha) v(x,t)^2 dx, \end{aligned}$$

which can be found in (3.5.5) in the proof of Lemma 3.5.1.

Passing from the second to the third line in the above is a result of an L^∞ estimation, plus the "normal" Poincaré inequality, (3.3.1). One idea that comes to mind on how one can improve this is to try and replace these two inequality with a *weighted Poincaré inequality*, i.e. to try and find a minimal constant C_ω, for a given weight $\omega(x)$, such that

$$\int_0^{2\pi} \left(f(x) - f_{\text{avg}} \right)^2 \omega(x) dx \leq C_\omega^2 \int_0^{2\pi} \left(f'(x) \right)^2 dx. \tag{3.A.2}$$

Indeed, denoting by

$$C_\omega := \inf\left\{C>0 \;\middle|\; \int_0^{2\pi}\left(f(x)-f_{\text{avg}}\right)^2\omega(x)dx \le C^2\int_0^{2\pi}\left(f'(x)\right)^2dx\right\},$$

we find that C_ω is in fact a minimum, which in turn satisfies (3.A.2). One can immediately see that if ω is bounded then

$$C_\omega \le \sqrt{\|\omega\|_\infty},$$

which shows how replacing the L^∞ approach with the weighted Poincaré constant in the estimations of (3.5.5) will gives a smaller upper bound, which in turn will translate to a larger range of choices for θ and α. Indeed, one find that

$$\begin{aligned}&\left|\frac{\theta}{2\pi}\int_0^{2\pi}(\sigma(x)-\alpha)\,\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx\right|\\ \le\ &\frac{\theta^2}{4}C^2_{\frac{(\sigma(x)-\alpha)^2}{2\sigma(x)-\theta-\alpha}}\|u(t)-u_{\text{avg}}\|^2+\frac{1}{2\pi}\int_0^{2\pi}(2\sigma(x)-\theta-\alpha)v(x,t)^2dx,\end{aligned}$$

yielding, according to the proof of Lemma 3.5.1, the following condition for the exponential convergence of E_θ with a rate α:

$$\frac{\theta^2}{4}C^2_{\frac{(\sigma(x)-\alpha)^2}{2\sigma(x)-\theta-\alpha}} \le \theta-\alpha. \tag{3.A.3}$$

Since

$$\sup_{x\in\mathbb{T}}\frac{\theta^2(\sigma(x)-\alpha)^2}{4(2\sigma(x)-\theta-\alpha)} \le \theta-\alpha$$

implies

$$\frac{\theta^2}{4}C^2_{\frac{(\sigma(x)-\alpha)^2}{2\sigma(x)-\theta-\alpha}} \le \sup_{x\in\mathbb{T}}\frac{\theta^2(\sigma(x)-\alpha)^2}{4(2\sigma(x)-\theta-\alpha)} \le \theta-\alpha,$$

we see that (3.A.3) gives us, as suggested, an improved decay rate.

Naturally, the study of weighted Poincaré inequalities is far from easy, which is the reason why we elected to present the L^∞ variant of the proof. However in the simple case we consider, and slightly more general cases, one can find the Poincaré constant.

The structure of the appendix is as follows: In §3.A.1 we will discuss the weighted Poincaré constant, and show some cases where one can explicitly compute it, which we will then use in our simple case in §3.A.2. We will then use [9] to evaluate the optimal convergence rate to (3.1.1) under the assumption that $\sigma(x)$ is given by (3.A.1) in §3.A.3, and conclude in §3.A.4 where we will compare the three available rates of convergence: The one from Theorem 3.2.2, the one we'll obtain in §3.A.2, and the optimal one which will be computed in §3.A.3.

3.A.1 Weighted Poincaré Inequality

Since our goal in this appendix is to show how one can utilise the weighted Poincaré inequality to improve our methodology, we will deal the finding of the associated constant more formally than rigorously. We start by recasting the problem of finding C_ω as a minimisation problem, and assume formally that it poses a solution, and that the differential equation we will find to classify the extremal points of our functional will provide this global minimum.

Consider the functional

$$\mathscr{F} : \boldsymbol{D} = H_0^1(\mathbb{T}) \cap H^2(\mathbb{T}) \to \mathbb{R}$$

given by

$$\mathscr{F}(u) := \int_0^{2\pi} \left(u'(x)\right)^2 dx,$$

and denote by

$$c_{\min} := \min\left\{\mathscr{F}(u) \,\middle|\, u \in \boldsymbol{D},\ \int_0^{2\pi} u(x)^2 \omega(x) dx = 1,\ \int_0^{2\pi} u(x) dx = 0\right\}. \tag{3.A.4}$$

Since for any $u \in \boldsymbol{D}$ we have that

$$\mathscr{F}(u) \geq \frac{1}{C_\omega^2} \int_0^{2\pi} \left(u(x) - u_{\text{avg}}\right)^2 \omega(x) dx$$

we see that $c_{\min} \geq \frac{1}{C_\omega^2}$. The converse is also true as for any $u \in \boldsymbol{D}$ we have that

$$v = \frac{u - u_{\text{avg}}}{\sqrt{\int_0^{2\pi} \left(u(x) - u_{\text{avg}}\right)^2 \omega(x) dx}}$$

satisfy the conditions in the definition of $c_{\min}$ and as such

$$\frac{\int_0^{2\pi} \left(u'(x)\right)^2 dx}{\int_0^{2\pi} \left(u(x) - u_{\text{avg}}\right)^2 \omega(x) dx} = \mathscr{F}(v) \geq c_{\min}.$$

Due to the sharpness of C_ω, the above inequality implies that $\frac{1}{C_\omega^2} \geq c_{\min}$, and we ca conclude that

$$C_\omega^2 = \frac{1}{c_{\min}}.$$

Thus, to find C_ω we focus our attention on finding $c_{\min}$. Assuming that u_* is a minimiser, we see that

$$\frac{d}{d\varepsilon}\mathscr{F}\left(\frac{u_* + \varepsilon\left(h - h_{\text{avg}}\right)}{\sqrt{\int_0^{2\pi} \left(u_*(x) + \varepsilon\left(h(x) - h_{\text{avg}}\right)\right)^2 \omega(x) dx}}\right)\Bigg|_{\varepsilon=0} \tag{3.A.5}$$

for any $h \in H_0^1(\mathbb{T}) \cap H^2(\mathbb{T})$. Since

$$\frac{d}{d\varepsilon}\int_0^{2\pi} \left(u_*'(x) + \varepsilon h'(x)\right)^2 dx\Big|_{\varepsilon=0} = 2\int_0^{2\pi} u_*'(x)h'(x)dx$$

$$= -2\int_0^{2\pi} u_*''(x)h(x)dx,$$

and

$$\frac{d}{d\varepsilon}\int_0^{2\pi} \left(u_*(x) + \varepsilon\left(h(x) - h_{\text{avg}}\right)\right)^2 \omega(x)dx\,\Big|_{\varepsilon=0} = 2\int_0^{2\pi} u_*(x)\left(h(x) - h_{\text{avg}}\right)\omega(x)dx$$

$$= 2\int_0^{2\pi} u_*(x)h(x)\omega(x)dx - 2h_{\text{avg}}\int_0^{2\pi} u_*(x)\omega(x)dx,$$

equation (3.A.5) together with the fact that u_* is in the minimisation set of (3.A.4) implies that

$$-2\int_0^{2\pi} u_*''(x)h(x)dx - \left(\int_0^{2\pi} \left(u'(x)\right)^2 dx\right)\left(2\int_0^{2\pi} u_*(x)h(x)\omega(x)dx - 2h_{\text{avg}}\int_0^{2\pi} u_*(x)\omega(x)dx\right) = 0,$$

which can be rewritten as

$$\int_0^{2\pi} \left(u_*''(x) + \lambda u_*(x)\omega(x) - \tau\right) h(x)dx = 0$$

with

$$\lambda = \int_0^{2\pi} \left(u_*'(x)\right)^2 dx = \mathscr{F}(u_*) > 0 \; {}^{6} \tag{3.A.6}$$

and

$$\tau = \frac{1}{2\pi}\left(\int_0^{2\pi} \left(u_*'(x)\right)^2 dx\right)\left(\int_0^{2\pi} u_*(x)\omega(x)dx\right). \tag{3.A.7}$$

From the above we can conclude two things:

- ◦ u_* solves the differential equation

  $$u'' + \lambda u(x)\omega(x) - \tau = 0 \tag{3.A.8}$$

  for $\lambda > 0$ and $\tau \in \mathbb{R}$.

- ◦ $\lambda = \mathscr{F}(u_*)$.

[6] $\mathscr{F}(u_*) = 0$ is and only if u_* is constant, and since $u_* \in H_0^1(\mathbb{T})$, this constant must be zero, which contradicts the fact that $\int_0^{2\pi} u_*(x)^2\omega(x)dx = 1$.

Thus, solving (3.A.8) and finding the minimal positive λ for which it is valid gives us the inverse of the desired weighted Poincaré constant.[7]

We would also like to mention that (3.A.8) can be obtained by means of constrained Euler-Lagrange equations.

The differential equation (3.A.8) is, in general, not simple to solve but in certain cases — which include our own, one can obtain some results. We will consider weights of the form:

$$\omega_{\sigma_1,\sigma_2}(x) = \begin{cases} \sigma_1, & 0 \le x \le \pi, \\ \sigma_2, & \pi < x \le 2\pi. \end{cases}$$

In that case, we find that the solution to our equation is given by

$$\begin{aligned} u(x) &= \begin{cases} c_1 \sin\left(\sqrt{\lambda\sigma_1}x\right) + c_2\cos\left(\sqrt{\lambda\sigma_1}x\right) + \frac{\tau}{\lambda\sigma_1}, & 0 < x < \pi, \\ c_3 \sin\left(\sqrt{\lambda\sigma_2}x\right) + c_4\cos\left(\sqrt{\lambda\sigma_2}x\right) + \frac{\tau}{\lambda\sigma_2}, & \pi < x < 2\pi, \end{cases} \\ &= \begin{cases} u_1(x), & 0 < x < \pi, \\ u_2(x) & \pi < x < 2\pi. \end{cases} \end{aligned}$$

Since $u \in H_0^1(\mathbb{T}) \cap H^2(\mathbb{T})$ and satisfies the minimisation conditions, we also have that:

$$\begin{aligned} u_1(0) &= u_2(2\pi), \\ u_1(\pi) &= u_2(\pi), \\ u_1'(0) &= u_2'(2\pi), \\ u_1'(\pi) &= u_2'(\pi), \\ \int_0^\pi u_1(x)dx + \int_0^\pi u_2(x)dx &= 0, \\ \int_0^\pi \sigma_1 u_1(x)^2 dx + \int_0^\pi \sigma_2 u_2^2(x)dx &= 1. \end{aligned}$$

As the existence of an extremal point implies that there is a non-zero vector $(c_1, c_2, c_3, c_4, c_5, \tau)^T$ and $\lambda > 0$ such all the above is satisfied, we conclude, just by considering the *first five* constraints, that there is a 5 × 5 matrix, $M(\lambda)$, such that

$$M(\lambda)\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ \tau \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

As the above system has a non-trivial solution, $\det(M(\lambda)) = 0$, which is how we search for options of λ. Adding the last constraint, one can use numerical methods to estimate

$$c_{\min}(\sigma_1, \sigma_2) = \lambda_{\min}(\sigma_1, \sigma_2).$$

[7] Here lies the assumption that a global minimum exists, and that it is an extremal point.

3.A.2 Improved Methodology

Returning to the proof of the differential inequality that governs the evolution of E_θ, expressed in Lemma 3.5.1, and in particular to (3.5.5), we see that with the choice of $\theta^* = 1$ and $\sigma(x)$ as in (3.A.1), we have that

$$\frac{d}{dt}E_1\left(u(t)-u_{\text{avg}}, v(t)\right) \le -\alpha E_1\left(u(t)-u_{\text{avg}}, v(t)\right) - \left(1-\alpha-\frac{C^2_{\omega_\alpha}}{4}\right)\|u(t)-u_{\text{avg}}\|^2$$

where

$$\omega_\alpha(x) = \begin{cases} 1-\alpha, & 0 \le x \le \pi, \\ \frac{(4-\alpha)^2}{7-\alpha}, & \pi < x \le 2\pi, \end{cases} = \frac{(\sigma(x)-\alpha)^2}{2\sigma(x)-1-\alpha}.$$

which implies that we search for α for which

$$\alpha \le 1 - \frac{C^2_{\omega_\alpha}}{4}.$$

Choosing

$$\alpha_0 = \alpha^*(1,4) = 4-\sqrt{12} = 2\left(2-\sqrt{3}\right),$$

which is rate one gets from our main theorem, Theorem 3.2.2, one can follow the process described in the previous subsection to find that

$$C^2_{\omega_{\alpha_0}} = 1.12013,$$

which satisfies the above condition. One can further imagine a way to improve this process: we can try to create a sequence $\{\alpha_n\}_{n\in\mathbb{N}}$ such that each α_n "maximises" the previous step, i.e.

$$\alpha_n = 1 - \frac{C^2_{\omega_{\alpha_{n-1}}}}{4}, \qquad n \in \mathbb{N},$$

while still satisfies $\frac{C^2_{\omega_{\alpha n}}}{4} \le 1-\alpha_n$. This means that for a given α_n, we compute $C_{\omega_{\alpha_n}}$ in a manner that was described in the previous subsection (note that ω_α is always of the form we've explored), and proceed to define α_{n+1}. Taking the limit in this process, if it exists, gives a viable candidate which in addition could satisfy

$$\alpha_{\max,\text{P}} = 1 - \frac{C^2_{\omega_{\alpha_{\max,\text{P}}}}}{4}.$$

Doing so in our case, and using numerical methods, one finds

$$\alpha_{\max,\text{P}} \approx 0.7234.$$

3.A.3 Optimal Rate of Convergence

The optimal rate of exponential convergence to the Goldstein–Taylor equation, (3.1.1), was found by Bernard and Salvarani in [9], and is expressed in the following theorem:

Theorem 3.A.1. *Consider the Goldstein–Taylor equations on $\frac{\mathbb{T}}{2\pi} \times (0,\infty)$:*

$$\begin{aligned} \partial_t f_+(x,t) &= -\partial_x f_+(x,t) + \widetilde{\sigma}(x)(f_-(x,t) - f_+(x,t)), \\ \partial_t f_-(x,t) &= \partial_x f_-(x,t) - \widetilde{\sigma}(x)(f_-(x,t) - f_+(x,t)), \\ f_\pm(x,0) &= f_{\pm,0}(x), \end{aligned} \tag{3.A.9}$$

where $f_{\pm,0} \in H^1\left(\frac{\mathbb{T}}{2\pi}\right)$ are non-negative functions, and $\widetilde{\sigma} \in L^\infty\left(\frac{\mathbb{T}}{2\pi}\right)$. Then, denoting by

$$f_\infty := \frac{1}{2}\int_0^1 \left(f_{+,0}(x) + f_{-,0}(x)\right) dx,$$

we find that there exists a constant A_* that depends only on $\|f_{\pm,0}\|_{H^1\left(\frac{\mathbb{T}}{2\pi}\right)}$ and $\|\widetilde{\sigma}\|_\infty$, such that

$$\|f_+(t) - f_\infty\|_{L^2\left(\frac{\mathbb{T}}{2\pi}\right)} + \|f_-(t) - f_\infty\|_{L^2\left(\frac{\mathbb{T}}{2\pi}\right)} \leq A_* e^{-\alpha_{BS} t}$$

with

$$\alpha_{BS} = 2\min\left(\|\widetilde{\sigma}\|_{L^1\left(\frac{\mathbb{T}}{2\pi}\right)}, -D(0)\right),$$

where

$$D(0) = \lim_{R\to 0^+} \sup\left\{\mathrm{Re}(\gamma) \mid \gamma \in sp(A_{\widetilde{\sigma}}),\, |\gamma| \geq R\right\}$$

and A_a is the operator whose domain is $D\left(A_{\widetilde{\sigma}}\right) = \left(H_0^1\left(\frac{\mathbb{T}}{2\pi}\right) \cap H^2\left(\frac{\mathbb{T}}{2\pi}\right)\right) \oplus H_0^1\left(\frac{\mathbb{T}}{2\pi}\right)$, and whose matrix representation is

$$A_{\widetilde{\sigma}} = \begin{pmatrix} 0 & 1 \\ \frac{d^2}{dx^2} & -2\widetilde{\sigma} \end{pmatrix}.$$

α_{BS} is the optimal exponential decay rate associated in these settings.

To compare the above result with our estimation we notice that a rescaling of both variables by a factor of 2π is required, and that the relaxation function in (3.A.9) lack the factor of a half which the relaxation function in our equations, (3.1.1), has. Taking all of this into account, we see that the connection between the relaxation functions is given by

$$\widetilde{\sigma}(\xi) = \pi\sigma(2\pi\xi), \qquad \xi \in \frac{\mathbb{T}}{2\pi},$$

and the appropriate decay rate one finds in our setting will be $\overline{\alpha} = \frac{\alpha_{\mathrm{BS}}}{2\pi}$.

For simplicity we will sometimes identify $\frac{\mathbb{T}}{2\pi}$ as $[0,1]$ with periodic conditions. From the definition of $\sigma(x)$ in our special case, (3.A.1), we find that

$$\widetilde{\sigma}(\xi) = \begin{cases} \pi, & 0 \leq \xi \leq \frac{1}{2}, \\ 4\pi, & \frac{1}{2} < \xi \leq 1. \end{cases} \tag{3.A.10}$$

As can be seen in [18], the spectrum of $A_{\widetilde{\sigma}}$, besides potentially $\{0\}$, is discrete and its eigenvalues, γ, satisfy

$$\operatorname{Re}\left(\gamma\right) \in [-2\|\widetilde{\sigma}\|_{\infty}, 0]\,. \tag{3.A.11}$$

The eigenvalue problem

$$A_{\widetilde{\sigma}}\begin{pmatrix} u \\ v \end{pmatrix} = \gamma\begin{pmatrix} u \\ v \end{pmatrix},$$

with $\gamma \neq 0$, is equivalent to the set of equations

$$u = \gamma v,$$

$$u'' - 2\widetilde{\sigma} v = \gamma v,$$

which translates to

$$v''(\xi) = \gamma\left(2\widetilde{\sigma}(\xi) + \gamma\right) v(\xi),$$

and $u = \gamma v$. As u and v are both real valued functions, $\gamma \in \mathbb{R}$, and due to (3.A.11) we find that

$$\gamma\left(2\widetilde{\sigma}(\xi) + \gamma\right) \leq 0.$$

Using (3.A.10), and the above consideration, we find that the solution to the our differential equation is given by

$$\begin{aligned} v(\xi) &= \begin{cases} c_1 \sin\left(\sqrt{-\gamma\left(2\pi + \gamma\right)}x\right) + c_2 \cos\left(\sqrt{-\gamma\left(2\pi + \gamma\right)}x\right), & 0 \leq \xi \leq \frac{1}{2} \\ c_3 \sin\left(\sqrt{-\gamma\left(8\pi + \gamma\right)}x\right) + c_4 \cos\left(\sqrt{-\gamma\left(8\pi + \gamma\right)}x\right), & \frac{1}{2} < \xi \leq 1, \end{cases} \\ &= \begin{cases} v_1(\xi) & 0 \leq \xi \leq \frac{1}{2}, \\ v_2(\xi) & \frac{1}{2} < \xi \leq 1, \end{cases} \end{aligned}$$

together with the “boundary” conditions

$$\begin{aligned} v_1(0) &= v_2\left(1\right), \\ v_1\left(\frac{1}{2}\right) &= v_2\left(\frac{1}{2}\right), \\ v_1'(0) &= v_2'(1), \\ v_1'\left(\frac{1}{2}\right) &= v_2'\left(\frac{1}{2}\right), \end{aligned}$$

which follow from the fact that $u, v \in H_0^1\left(\frac{\mathbb{T}}{2\pi}\right) \cap H^2\left(\frac{\mathbb{T}}{2\pi}\right)$. Much like the previous subsection, §3.A.2, we can find a (singular) 4×4 matrix, $M\left(\gamma\right)$, such that

$$M\left(\gamma\right)\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

and

$$\left\{\mathrm{Re}(\gamma) \mid \gamma \in \mathrm{sp}\left(A_{\widetilde{\sigma}}\right),\, |\gamma| \geq R\right\} = \left\{\gamma \in [-8\pi, 0) \cap (-R,R)^c \mid \det\left(M\left(\gamma\right)\right) = 0\right\}.$$

With the help of numerical methods one finds that

$$D(0) = \sup\left\{\gamma \in [-8\pi, 0) \mid \det\left(M\left(\gamma\right)\right) = 0\right\} \approx -2.7283.$$

With $D(0)$ computed, we find that

$$\alpha_{\mathrm{BS}} \approx 2\min\left(\|\widetilde{\sigma}\|_{L^1([0,1])}, 2.7283\right) = 2\min\left(\frac{5\pi}{2}, 2.7283\right) = 5.4566$$

and as such $\overline{\alpha} \approx 0.86844$.

3.A.4 Comparison of Convergence Rates

We now have three convergence rates for the case

$$\sigma(x) = \begin{cases} 1, & 0, \leq x \leq \pi, \\ 4, & \pi < x \leq 2\pi. \end{cases}$$

- The rate from our main theorem is $\alpha^* = 4 - \sqrt{12} \approx 0.5359$.
- The rate from our improved technique in §3.A.2 is $\alpha_{\mathrm{max,P}} \approx 0.7234$.
- The rate from the work of Bertrand and Salvarani is $\overline{\alpha} \approx 0.86844$.

This shows, as expected, the lack of optimality in our technique.

3.B Deferred proofs

Proof of Lemma 3.3.1. While the proof is standard, we show it here for completion and to fix the sharp constant. Denoting the k-th Fourier coefficient of f by

$$\widehat{f}(k) := \frac{1}{2\pi}\int_0^{2\pi} f(x) e^{-ikx} dx$$

we find that

$$\widehat{f'}(k) = ik\widehat{f}(k) \tag{3.B.1}$$

for all $k \in \mathbb{Z}$ (including $k = 0$). The condition $f_{\mathrm{avg}} = 0$ is equivalent to $\widehat{f}(0) = 0$ and as such, using Plancherel's equality, we find that

$$\|f\|^2 = \sum_{k\in\mathbb{Z}} \left|\widehat{f}(k)\right|^2 = \sum_{k\in\mathbb{Z}\setminus\{0\}} \left|\widehat{f}(k)\right|^2$$

$$= \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{\left|\widehat{f'}(k)\right|^2}{k^2} \leq \sum_{k\in\mathbb{Z}\setminus\{0\}} \left|\widehat{f'}(k)\right|^2 = \sum_{k\in\mathbb{Z}} \left|\widehat{f'}(k)\right|^2 = \|f'\|^2,$$

completing the proof. □

Proof of Lemma 3.3.2. Since for any function $h \in L^1(\mathbb{T})$

$$\left(h - h_{\text{avg}}\right)_{\text{avg}} = 0\,,$$

we conclude (i) from the definition of $\partial_x^{-1} f(x)$. To show (ii) we invoke the fundamental theorem of calculus (the version from Lebesgue theory), and to show (iii) we notice that if f is differentiable

$$\begin{aligned}\partial_x^{-1}\left(\partial_x f\right)(x) &= \int_0^x \partial_y f(y)dy - \frac{1}{2\pi}\int_0^{2\pi}\left(\int_0^x \partial_y f(y)dy\right)dx\\ &= f(x) - f(0) - \frac{1}{2\pi}\int_0^{2\pi}\left(f(x) - f(0)\right)dx = f(x) - f_{\text{avg}}.\end{aligned}$$

Lastly, we notice that the continuity of $\partial_x^{-1} f(x)$ as a function on the interval $[0,2\pi]$ is a standard result from Analysis. To conclude the continuity on the torus, though, we must also show that $\partial_x^{-1} f(0) = \partial_x^{-1} f(2\pi)$. This is equivalent to

$$0 = \int_0^0 f(x)dx = \int_0^{2\pi} f(x)dx = 2\pi f_{\text{avg}},$$

which is exactly the additional assumption. In addition, (3.3.2) for $k \neq 0$ follows immediately from (3.B.1) and (ii). For $k = 0$ we use

$$\widehat{\partial_x^{-1} f}(0) = \left(\partial_x^{-1} f\right)_{\text{avg}} = 0,$$

according to (i). The proof is thus complete. □

Proof of Lemma 3.3.4. We will establish that

$$\left|\frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1} f(x)\overline{g(x)}dx\right| \leq \frac{|\theta|}{2}\left(\|f\|^2 + \|g\|^2\right)$$

from which (3.3.3), (3.3.4) and (3.3.5) all follow. Indeed, from the Cauchy-Schwartz inequality, the Poincaré inequality — which is valid since $(\partial_x^{-1} f)_{\text{avg}} = 0$ — and point (ii) of Lemma 3.3.2 we conclude that

$$\begin{aligned}\left|\frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1} f(x)\overline{g(x)}dx\right| &\leq |\theta|\,\|\partial_x^{-1} f\|\|g\| \leq \frac{|\theta|}{2}\left(\|\partial_x^{-1} f\|^2 + \|g\|^2\right)\\ &\leq \frac{|\theta|}{2}\left(\|\partial_x\left(\partial_x^{-1} f\right)\|^2 + \|g\|^2\right) = \frac{|\theta|}{2}\left(\|f\|^2 + \|g\|^2\right)\,.\end{aligned}$$

The proof is thus completed. □

Bibliography

[1] Achleitner, F., Arnold, A. and Carlen, E.A.: *On linear hypocoercive BGK models.* In: From particle systems to partial differential equations. III, Springer Proc. Math. Stat., vol. 162, 1–37. Springer (2016).

[2] Achleitner, F., Arnold, A. and Carlen, E.A.: *On multi-dimensional hypocoercive BGK models.* Kinetic & Related Models 11 (4) 953–1009 (2018).

[3] F. Achleitner, A. Arnold, B. Signorello.: *On optimal decay estimates for ODEs and PDEs with modal decomposition.* Stochastic Dynamics out of Equilibrium, Springer Proceedings in Mathematics and Statistics vol. 282, 241–264 (2019)

[4] Albi, G., Herty, M., Jörres, C., Pareschi, L.: *Asymptotic preserving time-discretization of optimal control problems for the Goldstein-Taylor model.* Numer. Meth. for PDEs vol. 30 (6), 1770–1784 (2014).

[5] Arnold, A., Carrillo, J.A., Tidriri, M.D.: *Large-time behavior of discrete kinetic equations with non-symmetric interactions.* Math. Models and Meth. in the Appl. Sc. vol. 12 (11), 1555–1564 (2002).

[6] Arnold, A. and Erb, J.: *Sharp Entropy Decay for Hypocoercive and Non-Symmetric Fokker–Planck Equations With Linear Drift.* Preprint, arXiv:1409.5425 (2014).

[7] Arnold, A., Einav, A. and Wöhrer, T.: *On the rates of decay to equilibrium in degenerate and defective Fokker–Planck equations.* J. Differential Equations, vol. 264 (11), 6843–6872 (2018).

[8] Arnold, A., Jin, S. and Wöhrer, T.: *Sharp Decay Estimates in Local Sensitivity Analysis for Evolution Equations with Uncertainties: from ODEs to Linear Kinetic Equations* J. Differential Equations, vol. 268 (3), 1156–1204 (2020).

[9] Bernard, É., Salvarani, F.: *Optimal Estimate of the Spectral Gap for the Degenerate Goldstein-Taylor Model.* J. Stat. Phys. 153, 363–375 (2013);
Erratum. to appear in J. Stat. Phys. (2020).

[10] Bhatnagar, P.L., Gross, E.P. and Krook, M.: *A Model for Collision Processes in Gases. I. Small Amplitude Processes in Charged and Neutral One-Component Systems.* Phys. Rev. vol. 94 (3), 511–525 (1954).

[11] Cercignani, C., Illner, R., Shinbrot, W.: *A boundary value problem for discrete-velocity models.* Duke Math. J. vol. 55 (4), 889–900 (1987).

[12] Desvillettes, L., Salvarani, F.: *Asymptotic behavior of degenerate linear transport equations.* Bull. Sci. Math. vol. 133 (8), 848–858 (2009).

[13] Dolbeault, J., Mouhot, C., Schmeiser, C.: *Hypocoercivity for linear kinetic equations conserving mass.* Trans. Amer. Math. Soc., 3807–3828 (2015).

[14] Goldstein, C.: *On diffusion by discontinuous movements, and on the telegraph equation.* Quart. J. Mech. Appl. Math. vol. 4, 129–156 (1951).

[15] Gosse, L., Toscani, G.: *An asymptotic-preserving well-balanced scheme for the hyperbolic heat equations.* Comptes Rendus Math. vol. 334 (4), 337–342 (2002).

[16] Jin, S.: *Efficient Asymptotic-Preserving (AP) Schemes For Some Multiscale Kinetic Equations.* SIAM J. Sc. Comp. vol. 21 (2), 441–454 (1999).

[17] Kawashima, S.: *Existence and Stability of Stationary Solutions to the Discrete Boltzmann Equation.* Japan J. Indust. Appl. Math. vol. 8, 389–429 (1991).

[18] Lebeau, G.: *Équations des ondes amorties,* Séminaire Équations aux dérivées partielles (École Polytechnique), talk no. 15, 1–14 (1993-1994).

[19] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations,* Springer, 2nd edition, 1992.

[20] Salvarani, F. *Diffusion limits for the initial-boundary value problem of the Goldstein-Taylor model.* Rend. Sem. Mat. Univ. Polit. Torino vol. 57 (3) 211–222 (1999).

[21] Taylor G.I.: *Diffusion by Continuous Movements,* Proc. London Math. Soc., vol. S2-20 (1), 196–212 (1922).

[22] Tran, M.-B.: *Convergence to equilibrium of some kinetic models.* J. Diff. Eq. vol. 255, 405–440 (2013).

[23] Villani, C.: Hypocoercivity, American Mathematical Soc., (2009).