DISSERTATION

On Optimal Decay Estimates for Hypocoercive Evolution Equations

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

Dott.Mag Beatrice Signorello
Matrikelnummer: 11732731

Diese Dissertation haben begutachtet:

1. **Univ.Prof. Dr. Anton Arnold**
   Institut für Analysis und Scientific Computing, TU Wien
2. **Univ.Prof. Dr. Christian Schmeiser**
   Fakultät für Mathematik, Universität Wien
3. **Univ.Prof. Dr. Francesco Salvarani**
   Fakultät für Mathematik F. Casorati, Universität Pavia

Wien, am 27. Oktober 2021

Kurzfassung

Das Hauptthema dieser Arbeit ist das Langzeitverhalten von hypokoerziven Evolutionsoperatoren. In einigen Hilberträumen konvergieren die entsprechenden Lösungen exponentiell zum jeweils eindeutigen stationären Zustand. Der Schwerpunkt dieser Arbeit liegt in der Bestimmung der optimalen Geschwindigkeit dieser Konvergenz. Im Speziellen soll die Zerfallsrate maximiert, und die multiplikative Konstante in den Zerfallsabschätzungen minimiert werden. Es werden insbesondere zwei hypokoerzive partielle Differentialgleichungsmodelle betrachtet: Die linearen degenerierten Fokker-Planck-Gleichungen und das Goldstein-Taylor-System. Die folgenden Ansätze werden für die Untersuchung der beiden Modelle verwendet: die Zerfallsanalyse der zugehörigen Tensorzerlegung und die Konstruktion geeigneter Lyapunov-Funktionale. Darüber hinaus wird die Optimalität von Zerfallsabschätzungen für endlich dimensionale gewöhnliche Differentialgleichungssysteme, die durch positive stabile Matrizen dargestellt werden, im Detail analysiert. Diese Ergebnisse dienen als wichtiges Werkzeug für die Thematisierung wesentlicher Fragestellungen dieser Arbeit. In der Einleitung stellen wir das Konzept der Hypokoerzitivität, den Rahmen der Evolutionsoperatoren, die Motivationen und die wichtigsten Ergebnisse vor. Der Hauptteil der Arbeit ist in vier Kapitel unterteilt. Im ersten Kapitel zeigen wir explizite optimale Zerfallsabschätzungen für endlich dimensionale gewöhnliche Differentialgleichungen. Das Kurz- und Langverhalten der (degenerierten) Fokker-Planck-Gleichung mit linearem Drift ist die zentrale Thematik des zweiten und dritten Kapitels. Das wichtigste Ergebnis ist die Verbindung zwischen solchen partiellen Differentialgleichungen und deren zugehörigen gewöhnlichen Differentialgleichungen mit Drift. In der Tat stimmen ihre Propagator-Normen überein. Dies impliziert, dass optimale Zerfallsabschätzungen auf der gewöhnlichen Differentialgleichungsebene auf einfache Weise auf die Ebene der Fokker-Planck-Gleichung übertragbar sind. Schließlich wird im letzten Kapitel die Konvergenz zum Gleichgewicht des Goldstein-Taylor-Modells auf dem eindimensionalen Torus behandelt. Ziel dieser Analyse ist es, eine allgemeine Methode für den Fall zu entwickeln, dass die Relaxationsfunktion nicht konstant ist, indem ein geeignetes Lyapunov-Funktional pseudodifferentieller Natur definiert wird.

Abstract

The main topic of this thesis is the large-time behaviour of hypocoercive evolution operators. In this setting, the solutions converge exponentially to the unique steady state in some prescribed Hilbert spaces. The optimality of the speed of this convergence is the focus of this work. More precisely, the maximization of the decay rate and the minimization of the multiplicative constant appearing in the decay estimates. In particular, two hypocoercive PDE-models are considered: The *linear degenerate Fokker-Planck equations* and the *Goldstein–Taylor system*. The following approaches are used for the resolution of the two models: the decay analysis of the associated tensor decomposition and the construction of appropriated *Lyapunov functionals*, respectively. Moreover, the optimality of decay estimates for finite dimensional ODE-systems represented by positive stable matrices is studied in detail. These results represent an important tool for the resolution of the main questions of this thesis at the PDE-level. In the introduction we present the concept of hypocoercivity, the setting of the evolution operators, the motivations and the main results. After that, the thesis will be divided into four chapters.

In the first chapter, we shall display explicit optimal decay estimates for finite dimensional ODEs. The short- and large-time behaviour of the (degenerate) Fokker-Planck equation with linear drift will be the main subject of the second and the third chapter. The main result in this regard is the connection between such PDEs and their associated drift-ODEs. In fact, their propagator norms coincide. This implies that optimal decay estimates at the ODE-level carry over to the level of the Fokker-Planck equation in a straightforward way. Finally, the last chapter will focus on the convergence to the equilibrium for the Goldstein-Taylor model on the one-dimensional torus. The goal of this analysis is to provide a general method when the relaxation function is not constant, by defining a suitable Lyapunov functional of pseudodifferential nature.

Acknowledgement

I would like to thank all the people and institutions that foster, in general, the human thirst of knowledge.

In my case, for the realization of this thesis, I feel the need to mention: The Technical University of Vienna (TU), which hosted me for 4 years, The Austrian Science Fund (FWF), which funded the Doctoral Program I was part of, and Prof. Anton Arnold, who trustingly chose me 4 years ago as his Phd-student. Moreover, I want to acknowledge the support of my father, that made the first stages of my education possible and always believed in me, the colleagues and friends that helped me several times during these 4 years and Oliver, because without love we could only stick to the unfulfilling definition of “thinking beings”.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am Datum

Name des Autors

Contents

1 Introduction

This thesis is devoted to the qualitative and quantitative analysis of the long time behaviour of some classes of differential operators that converge to a unique equilibrium. In particular, we specialize to the case of two models: the (degenerate) Fokker-Planck (FP) equation with linear drift and the Goldstein-Taylor (GT) model. Furthermore, an explicit computation of the propagator norm of a simple class of ODEs will be provided. This will prove to be a key tool for the main results regarding the FP-equation. The partial differential equations considered here feature a diffusion/relaxation term that is possibly degenerate. On one hand, the presence of a kernel for this component makes the investigation of the convergence to the steady state more challenging. On the other hand, such degenerate evolutions allow for a wider correspondence in certain physical systems. The common phenomenon that appears throughout this thesis goes under the name of "hypocoercivity". When this occurs, the solution of the evolution operators converges with an exponential decay that is slowed down by a multiplicative constant bigger than one, contrarily to what occurs in case of coercivity where the decay is purely exponential. This thesis is concerned with the optimality of the above mentioned decay estimates. In particular, the results concerning the optimal multiplicative constant appearing in the decay estimates characterize the novelty of the present work.

1.1 Hypocoercivity

In this section, we shall introduce the general setting of hypocoercive dynamics and the related notation. In this regard, we follow the monograph on hypocoercivity [14] (we refer the reader to this text for a comprehensive understanding of the topic).

Let $\mathcal{H}$ be a separable (real or complex) Hilbert space and consider an unbounded operator L with domain $D(L)$ and the evolution equation

$$\partial_t f_t = -L f_t, \qquad f_0 \in \mathcal{H}, \forall\, t > 0. \tag{1.1.1}$$

In this case, we say that the operator $-L$ is the generator of the semigroup $\{e^{-Lt}\}_{t\geq 0}$. Once the well-posedness of the problem (1.1.1) is established, the following questions arise: First of all, does there exist a steady state and is it unique (after normalization)? Does the solution f_t converge to a unique steady state $f_\infty \in \mathcal{H}$ for $t \to +\infty$? If so, can we obtain (possibly optimal) decay estimates? For the evolutions that will be treated in the following chapters the existence and uniqueness of an equilibrium for the solution has been proven with different techniques, according to the specific scenario. For this reason, this property shall be always assumed, although we shall cite the literature for the proofs in each case. A convenient Hilbert space to consider, in order to obtain convergence, is given by $\mathcal{H} = L^2(K^d, f_\infty^{-1})$, i.e. the weighted-$L^2$ space defined as

$$L^2(K^d, f_\infty^{-1}) := \left\{ f : K^d \to K,\ f \text{ measurable such that} \int_{K^d} \left|f(x)\right|^2 f_\infty^{-1}(x)\, dx < \infty \right\}. \tag{1.1.2}$$

Here K denotes $\mathbb{R}$ or $\mathbb{C}$. When the speed of convergence $f_t \to f_\infty$ is purely exponential, we say that the operator L is *coercive*.

Definition 1.1.1. Let L be an unbounded operator on a Hilbert space $\mathscr{H}$, with kernel $\mathscr{K}$, and let $\widetilde{\mathscr{H}} := \mathscr{K}^\perp$ be the orthogonal complement of $\mathscr{K}$ endowed with the scalar product $\langle\cdot,\cdot\rangle_{\widetilde{\mathscr{H}}}$ and the Hilbertian norm $||\cdot||_{\widetilde{\mathscr{H}}}$. The operator L is said to be λ-coercive on $\widetilde{\mathscr{H}}$ if

$$||e^{-Lt} f_0||_{\widetilde{\mathscr{H}}} \leq e^{-\lambda t} ||f_0||_{\widetilde{\mathscr{H}}}, \qquad \forall f_0 \in \widetilde{\mathscr{H}}, \forall t \geq 0. \tag{1.1.3}$$

The operator L is said to be *coercive* on $\widetilde{\mathscr{H}}$ if it is λ-coercive for some $\lambda > 0$.

It is straightforward to prove (see §3.1 in [14]) that Definition 1.1.1 is equivalent to the following definition, which does not require semigroup-regularity, but rather involves the spectral properties of L.

Definition 1.1.2. Let L be an unbounded operator on a Hilbert space $\mathscr{H}$, with kernel $\mathscr{K}$, and let $\widetilde{\mathscr{H}} := \mathscr{K}^\perp$ be the orthogonal complement of $\mathscr{K}$ endowed with the scalar product $\langle\cdot,\cdot\rangle_{\widetilde{\mathscr{H}}}$ and the Hilbertian norm $||\cdot||_{\widetilde{\mathscr{H}}}$.. The operator L is said to be λ-coercive on $\widetilde{\mathscr{H}}$ if

$$\forall f_0 \in \mathscr{K}^\perp \cup D(L), \quad \mathrm{Re}\langle L f_0, f_0\rangle_{\widetilde{\mathscr{H}}} \geq \lambda ||f_0||^2_{\widetilde{\mathscr{H}}}, \tag{1.1.4}$$

where Re denotes the real part of a complex number. The operator L is said to be coercive on $\widetilde{\mathscr{H}}$ if it is λ-coercive on $\widetilde{\mathscr{H}}$ for some $\lambda > 0$.

Keeping in mind the notion of coercivity (determined by the spectral gap property) is useful in order to better understand the definition of hypocoercivity, even if the latter shall be given using a different language/notation. As the etymology (*hypo*=weak) of the name suggests, the property of being hypocoercive is a weaker concept that the one of being coercive, as it appears immediately clear from the following definition.

Definition 1.1.3. Let L be an unbounded operator on a Hilbert space $\mathscr{H}$, with kernel $\mathscr{K}$, and let $\widetilde{\mathscr{H}} := \mathscr{K}^\perp$ be the orthogonal complement of $\mathscr{K}$ endowed with the scalar product $\langle\cdot,\cdot\rangle_{\widetilde{\mathscr{H}}}$ and the Hilbertian norm $||\cdot||_{\widetilde{\mathscr{H}}}$. The operator L is said to be λ-hypocoercive on $\widetilde{\mathscr{H}}$ if there exists a finite constant $c \geq 1$ such that

$$||e^{-Lt} f_0||_{\widetilde{\mathscr{H}}} \leq c e^{-\lambda t} ||f_0||_{\widetilde{\mathscr{H}}}, \qquad \forall f_0 \in \widetilde{\mathscr{H}}, \forall t \geq 0. \tag{1.1.5}$$

The operator L is said to be hypocoercive on $\widetilde{\mathscr{H}}$ if it is λ-hypocoercive on $\widetilde{\mathscr{H}}$ for some $\lambda > 0$.

Remark 1.1.4. The definition of (hypo) coercivity can be generalized by replacing $\mathscr{K}^\perp$ with another Hilbert space continuously and densely embedded in $\mathscr{K}^\perp$. Throughout this thesis the standard choice $\widetilde{\mathscr{H}} = \mathscr{K}^\perp$ will be used all the chapters.

The difference between Definition 1.1.1 and 1.1.3 lies in the constant c appearing as a prefactor in (1.1.5). In the hypocoercive case a constant $c > 1$ is admitted, in contrast with the coercive case that corresponds to $c = 1$. Therefore, every coercive operator is also hypocoercive but not vice versa.

Let us assume the existence and uniqueness of a function $f_\infty \in \mathscr{H}$ such that $\mathscr{K} = \mathrm{span}_{\mathbb{C}}\{f_\infty\}$, i.e. f_∞ is the unique (up to a constant) steady state for L. As we shall see more in detail in the next chapters, the inequality (1.1.5) expresses the property of the solutions $f(t) := e^{-Lt} f_0$ to converge (w.r.t. $||\cdot||_{\widetilde{\mathscr{H}}}$) to the steady state f_∞ with exponential speed. The two constants λ and c appearing in the decay estimates (1.1.5) represent, respectively, the decay rate of the convergence and the multiplicative factor that slows it down in comparison to the pure exponential decay of coercive evolutions. The main purpose of this thesis is to optimize the decay estimates (1.1.5) for three prescribed classes of hypocoercive evolutions, with particular attention to the minimization of the multiplicative constant.

The definition of hypocoercivity can be naturally extended to vectorial operators $L : \mathscr{H} \to \mathscr{H}^m$ for some $m \geq 1$. Nevertheless, throughout this work we shall only consider scalar operators with values in $\mathscr{H}$. In particular we shall focus our attention on linear scalar operators of the form

$$L = A^* A + B, \quad B^* = -B, \tag{1.1.6}$$

where A^* denotes the adjoint operator of A with respect to the scalar product $\langle \cdot, \cdot \rangle_{\mathscr{H}}$ in $\mathscr{H}$. The simplest picture to bear in mind is the one in which the symmetric operator $D := A^* A$ represents the diffusive component of the operator L (note that in the GT-model considered in Chapter 5 D represents a relaxation term instead of a diffusive term). It can be observed that the property of B to be antisymmetric implies $\text{Re}\langle B f_0, f_0 \rangle = 0$. Therefore, operators that are hypocoercive but not coercive (i.e they do not admit a spectral gap) correspond to degenerate diffusions. In absence of a global force driving solutions to the equilibrium, the convergence in this case is guaranteed by the action of B on the directions of the kernel of D. The antisymmetric part acts by “mixing” such directions and making them unstable along the evolution. For instance, B shall play the role of a “drift term” and of a “transport” operator, respectively, in the FP-equations and the GT-model that will be examined in this thesis. The interplay between the degenerate diffusive component and the antisymmetric one is the essence of hypocoercivity.

1.2 Entropy Methods

Plenty of methods are used in the literature to establish hypocoercivity for a given evolution. The standard approach for the linear case goes under the name of *entropy methods* ([9], [5], [8], [12]). We present here a sketch of the main steps of the general strategy. The aim of entropy methods is to find a functional $E : \mathscr{H} \times \mathscr{H} \to \mathscr{H}$ that defines an equivalent norm in $\mathscr{H}$ such that the considered evolution is coercive along the trajectories w.r.t. this norm. More precisely, one first needs to prove that there exist two constants c_1 and c_2 such that

$$c_1 E(f, f) \leq ||f||_{\widetilde{\mathscr{H}}} \leq c_2 E(f, f) \qquad \forall f \in \widetilde{\mathscr{H}}, \tag{1.2.1}$$

and then that the following inequality holds along the trajectories $f(t) = e^{-Lt}$ (1.2.3)

$$E(f(t)) \leq e^{-\lambda t} E(f(0)), \qquad \forall t \geq 0. \tag{1.2.2}$$

Above, for simplicity of notation, we have used $E(f) = E(f, f)$ for any $f \in \mathscr{H}$. The standard way to obtain inequality (1.2.2) is to derive a functional inequality of the form

$$\frac{d}{dt} E(f(t)) \leq -\lambda E(f(t)), \qquad \forall t \geq 0, \tag{1.2.3}$$

and then apply Gronwall's Lemma. By combining (1.2.1) and (1.2.2), the hypocoercive estimates (1.1.5) can be easily achieved. In this work, entropy methods will be applied to ODE-equations in Chapter 2 and the GT-model in Chapter 5. We want to remark that, although this method is useful to prove the existence of a rate of convergence of exponential type, it does not allow to infer any information about the optimality of the obtained decay estimates. In general for the minimization of the decay rate other strategies need to be developed. We shall mention in the next chapters the results of some previous works in this regard. For the GT-model with space-dependent relaxation, however, the obtained convergence rates are, in general, not optimal.

On the other hand, the literature concerning the best multiplicative constant for hypocoercive estimates is still limited. For the finite dimensional ODEs studied in Chapter 2, the best constant will be provided explicitly in the two-dimensional case. This result represents a substantial contribution for the study carried in Chapter 3 and Chapter 4 about the best constant for the FP-equation. The main theorem in Chapter 3 states that the propagator norm of the FP-equation is the same as that of its corresponding ODE. This implies that the hypocoercive estimates at the PDE level can be studied at the level of the finite dimensional ODEs, making the explicit computations in Chapter 2 a key tool for the computation of the best constant for the the FP-equation.

We conclude this section by comparing the concept of hypocoercivity with that of hypoellipticity. While the first refers to the convergence of the solutions to the equilibrium, hypoellipticity regards regularity issues. They often occur together, as for the FP-equation analysed in Chapter 3. For this evolution the concept of *hypocoercivity index* can be defined, [5]. This value reflects the degeneracy structure of the equation and the interplay between the diffusion and the anti-symmetric part. Furthermore, it provides a rate of regularization of the solution for short times from the Hilbert space to some weighted Sobolev space (§7.3, §A.21 for the kinetic FP-equation with $m_{HC} = 1$, [14] and [5, Theorem 4.8]). As we shall see in Chapter 3 and Chapter 4, hypoellipticity is actually a necessary condition in the FP-setting for the hypocoercivity to occur. This is not the case for the GT-Taylor model ([13], [7]), where the solution starting from the Hilbert space $L^2(\mathbb{R}^d, f_\infty^{-1})$ do not benefit in general from any regularization effect for bigger times.

In the sequel of this chapter we shall give an overview of the remaining chapters. We shall split it into three parts, each one corresponding to a hypocoercive evolution that will be studied in this work: The finite dimensional ODEs, the (degenerate) FP-equation with linear drift and the GT-model including x-dependent relaxation. The main results and the corresponding strategies regarding hypocoercivity will be presented for each of these equations.

1.3 Finite dimensional ODEs

In this section we shall review the study of large time behaviour for linear finite dimensional ODEs, carried out in the joint paper [3] with Anton Arnold and Franz Achleitner, corresponding to Chapter 2 of the present thesis. In the sequel $\mathscr{S}^{>0}$ will denote the set of self-adjoint and positive definite matrices in $\mathbb{C}^{d\times d}$.

We consider the ODE system

$$\begin{cases} \frac{d}{dt}x(t) = -Cx(t), \quad t \geq 0\,, \\ x_0 := x(0) \in \mathbb{C}^d\,, \end{cases} \tag{1.3.1}$$

with $C \in \mathbb{C}^{d\times d}$ a (typically non-Hermitian) constant matrix. Following the notation of the previous section, we set $\mathscr{H} = \widetilde{\mathscr{H}} = \mathbb{C}^d$ endowed with the Euclidean norm $||\cdot||_2$. Clearly the origin is a steady state for the operator $L := C$. In order to guarantee convergence to it (and uniqueness of the equilibrium), we shall assume that the spectral gap of C is positive, i.e.

$$\rho(C) := \min\{\mathrm{Re}(\lambda) : \ \lambda \text{ is eigenvalue of } C\} > 0. \tag{1.3.2}$$

A matrix satisfying such condition is called *positive stable*. Since C_s, the symmetric part of C, can admit a non-trivial kernel, C is in general not coercive. Therefore one cannot deduce a decay rate by simply considering the euclidean norm as Lyapunov functional. Indeed

$$\frac{d}{dt}||x(t)||_2^2 = -2\langle x(t), C_s x(t)\rangle_2 \le 0. \tag{1.3.3}$$

On the other hand, the positive stability of C is a necessary and sufficient condition for hypocoercivity. More specifically, there exists $\lambda > 0$ and $c \ge 1$ such that

$$||x(t)||_2 \le c e^{-\lambda t} ||x_0||_2, \qquad \forall x_0 \in \mathbb{C}^d, \forall t \ge 0. \tag{1.3.4}$$

Once hypocoercivity is established, the question about optimality of the decay estimates (1.3.4) can arise. With the additional assumption that C is *non defective* (i.e. for every eigenvalue the algebraic and geometric multiplicity coincide) one can prove that the best admissible decay rate corresponds to the spectral gap of C, i.e.

$$\lambda_{opt}^{(1)} := \max\{\lambda > 0 \text{ such that (1.3.4) holds with } \lambda\} = \rho(C). \tag{1.3.5}$$

In [3] we focused on the two-dimensional case and we constructed Lyapunov functionals that yield both the optimal decay rate and the minimal multiplicative constant

$$c_{\min}^{(1)} := \min\{c \ge 1 \text{such that (1.3.4) holds with } c \text{ and } \lambda = \lambda_{opt}^{(1)}\}. \tag{1.3.6}$$

Following the strategy of entropy methods we defined in [3] a *modified norm* $||x||_P := \sqrt{\langle x, Px\rangle_2}$ in $\mathbb{C}$, represented by a matrix $P \in \mathscr{S}^{>0}$ that satisfies the *Lyapunov inequality*

$$PC + C^* P \ge 2\lambda_{opt}^{(1)} P. \tag{1.3.7}$$

The advantage of the new norm consists in the fact that the evolution (1.3.1) is coercive w.r.t. this norm. Indeed

$$\frac{d}{dt}||x(t)||_P^2 = -\langle x(t), (PC + C^* P)x(t)\rangle_2 \le -2\lambda_{opt}^{(1)} ||x(t)||_P^2. \tag{1.3.8}$$

Therefore,

$$||x(t)||_P \le e^{-\lambda_{opt}^{(1)} t} ||x_0||_P, \qquad \forall x_0, t \ge 0. \tag{1.3.9}$$

Since the norm $||\cdot||_P$ is equivalent to $||\cdot||_2$, one can finally obtain from (1.3.9) the hypocoercive estimates

$$||x(t)||_2 \le \kappa(P) e^{-\lambda_{opt}^{(1)}} ||x_0||_2, \qquad \forall x_0 \in \mathbb{C}^d, \forall t \ge 0. \tag{1.3.10}$$

Here $\kappa(P)$ denotes the condition number of P, i.e. $\kappa(P) := \sigma_{\max}(P)/\sigma_{\min}(P)$ where $\sigma_{\max}$ and $\sigma_{\min}$ are the maximal and minimal singular values of the matrix P.

Considering that the choice of P is not unique, we minimized in [3] the condition number among the set of matrices that satisfy the inequality (1.3.7). This is an easy question in the two-dimensional case, since such matrices can be explicitly classified. However, the minimal condition number does not coincide in general with the best constant $c^{(1)}_{\min}$. It rather represents in general an upper bound for it. One should indeed distinguish three cases, depending on the spectrum of C:

1. the two eigenvalues of C have the same real part, i.e. $\mathrm{Re}(\lambda_1) = \mathrm{Re}(\lambda_2)$;
2. the imaginary part is the same, i.e. $\mathrm{Im}(\lambda_1) = \mathrm{Im}(\lambda_2)$;
3. λ_1 and λ_2 have both distinct real and imaginary part.

In the first case the minimal $\kappa(P)$ coincides with the optimal constant, whose explicit formula is provided in [3] (see Theorem 2.3.7 in Chapter 2). For the second case the strategy of minimizing the conditional number had to be refined, by dividing the plane into sectors and localizing the minimization. However, also for this case $c^{(1)}_{min}$ was explicitly computed (see Theorem 2.4.1 in Chapter 2). For the most general case we only provided an implicit formula for $c^{(1)}_{min}$ as supremum in time of the upper envelope of the norm of the solutions, that was explicitly computed. We refer here to the content of Theorem 2.4.2 and Corollary 2.4.3 in Chapter 2. We end this section by mentioning that the study of large time behaviour of finite dimensional ODEs in [3] was initially motivated by the same question for the GT-model, when subjected to a modal decomposition. A posteriori, this analysis of the hypocoercivity for ODEs turned out to be useful also for the study of optimal decay estimates for the FP-equation, due to Theorem 3.4 in [4] that relates the propagator norm of the two hypocoercive models (the ODE -and the FP-equation).

1.4 The (degenerate) Fokker Plank equation with linear drift

In this section we will review the study about hypocoercive FP-equations with linear drift that was pursued in [4] and [7]. The full text of the two papers can be found in Chapter 3 and Chapter 4 of the present thesis.

1.4.1 Propagator Norm and Sharp Decay Estimates

In [4] we focus on the short and large time behaviour of the (possibly degenerate) Fokker-Planck equation with linear drift, i.e.

$$\partial_t f_t = -L f_t := \mathrm{div}_x(D f_t + C x f_t), \qquad x \in \mathbb{R}^d, t \geq 0, \tag{1.4.1}$$

for some matrices C, $D \in \mathbb{R}^{d\times d}$ constant in x and D positive semi-definite. We shall consider initial values f_0 in the weighted Hilbert space $\mathscr{H} = L^2(\mathbb{R}^d, f_\infty^{-1})$ (see (1.1.2) for its definition). Moreover, without loss of generality, we shall assume that $\int_{\mathbb{R}^d} f_0(x)\,dx = 1$.

In statistical mechanics, the FP-equation is a partial differential equation that describes the time evolution of the probability density function of the velocity of a particle under the influence of drag forces and random forces, as in Brownian motion.

Let K be a positive definite matrix satisfying the *Lyapunov equation*

$$2D = CK + KC^T. \tag{1.4.2}$$

A matrix with such property will be called *covariance matrix*. We shall mention in the following the assumptions that guarantee the existence and uniqueness of the covariance matrix. Then, it is straightforward to verify that the (non-isotropic) Gaussian

$$f_\infty(x) = (2\pi)^{-d/2}(\det K)^{-1/2}\exp\left(-\frac{x^T K^{-1}x}{2}\right), \qquad x \in \mathbb{R}^d, \tag{1.4.3}$$

is a steady state for the operator L, i.e. $Lf_\infty = 0$.

As for the ODE (1.3.1), the assumptions to guarantee the convergence of the solutions of (1.4.1) to the unique steady state f_∞ and to established hypocoercivity coincide. It has to be required:

- C is positive stable;
- there is no non-trivial C^T-invariant subspace of $\ker(D)$ (*hypoellipticity*).

When the above mentioned assumptions are satisfied, we say that *Condition A* holds for the FP-equation (1.4.1). In this setting there exists a unique covariance matrix $K \in \mathcal{S}^{>0}$ (Theorem 3.1, [5]) and hypocoercivity occurs (Theorem 4.9, [5]) i.e. there exists $\lambda > 0$ and $c \geq 1$ such that

$$\|f_t - f_\infty\|_{\mathcal{H}} \leq ce^{-\lambda t}\|f_0 - f_\infty\|_{\mathcal{H}}, \qquad \forall f_0 \in \mathcal{H}, \quad t \geq 0. \tag{1.4.4}$$

.

We shall call a FP-equation *non-defective* if its drift matrix C is not defective. For non-defective FP-equations, the best decay rate for the hypocoercive estimates (1.4.4) $\lambda_{opt}^{(2)}$ coincides with the spectral gap of the drift matrix, i.e. $\lambda_{opt}^{(2)} = \rho(C)$. Via entropy methods the authors in [5] obtained the decay estimates (1.4.4) with sharp decay rate $\lambda_{opt}^{(2)}$ but with a sub-optimal multiplicative constant $c > c_{min}^{(2)}$, with

$$c_{min}^{(2)} := \min\{c \geq 1 \text{ such that (1.4.4) holds with } \lambda_{opt}^{(2)} \text{ and } c\}. \tag{1.4.5}$$

The original motivation of the study carried out in [4] was to determine the best multiplicative constant $c_{min}^{(2)}$. As a matter of fact, the main result of this paper has a wider spectrum of applications. It states that the propagator norm of the FP-equation (1.4.1) coincides with the propagator norm of the ODE represented by its linear drift matrix, after normalization. In order to understand this theorem and its terminology, we shall introduce the definition of operator norms and normalized FP-equations.

Definition 1.4.1. If f_t is the solution of (1.4.1) with $f_0 \in \mathcal{H}$, we define the propagator norm of the FP-equation as

$$\|e^{-Lt}\|_{\mathcal{B}(V_0^\perp)} := \sup_{\substack{0\neq f_0\in\mathcal{H} \\ \int_{\mathbb{R}^d} f_0\,dx=1}} \frac{\|f_t - f_0\|_{\mathcal{H}}}{\|f_0 - f_\infty\|_{\mathcal{H}}}. \tag{1.4.6}$$

If $x(t) \in \mathbb{R}^d$ is the solution of the ODE $\frac{d}{dt}x(t) = -Cx(t)$ with initial datum $x(0) := x_0$, we define the propagator norm as

$$\|e^{-Ct}\|_{\mathcal{B}(\mathbb{R}^d)} = \sup_{0\neq x_0\in\mathbb{R}^d} \frac{\|x(t)\|_2}{\|x_0\|_2}. \tag{1.4.7}$$

Moreover, we associate to any FP-equation (1.4.1) with linear drift C its *normalized* version

$$\partial_t g_t = \widetilde{L} g_t = -\mathrm{div}_y \left(\widetilde{C}_s \nabla_y g_t + \widetilde{C} y g_t \right), \qquad t \geq 0, \tag{1.4.8}$$

where $\widetilde{C} := K^{-1/2} C K^{1/2}$. The normalized FP-equation is obtained from (1.4.1) after the change of coordinates $y = K^{-1/2} x$ and it is characterized by the fact that the diffusion is represented by the symmetric part of the drift matrix . Furthermore, the covariance matrix for the normalized equation (1.4.8) is the identity matrix, implying that the steady state is the isotropic Gaussian $g_\infty(y) = (2\pi)^{-d/2} e^{-|x|^2/2}$. Equivalently to the non normalized case, we set the convergence of the operator $e^{-\widetilde{L}t}$ in the weighted Hilbert space $\widetilde{H} := L^2(\mathbb{R}^d, g_\infty^{-1})$ and we define the subspace $\widetilde{V_0} := \mathrm{span}_{\mathbb{R}}\{g_\infty^{-1}\}$.

The main result in [4] (Theorem 3.4) establishes the equality between the propagator norm of the normalized FP-equation (1.4.8) and its associated ODE-equation $\frac{d}{dt}x(t) = -\widetilde{C}x(t)$. As a consequence, the equality holds also for the propagator norm of the the original FP-equation (1.4.1), since $||g_t||_{\widetilde{\mathcal{H}}} = ||f_t||_{\mathcal{H}}$, $\forall\, t \geq 0$.

Theorem 1.4.2. *Let us consider the FP-equation* (1.4.1) *and let Condition A hold for* (1.4.1). *Then the propagator norm of the normalized FP-equation* (1.4.8) *and its corresponding ODE are equal, i.e.,*

$$||e^{-\widetilde{L}t}||_{\mathcal{B}(\widetilde{V_0}^{\perp})} = ||e^{-\widetilde{C}t}||_{\mathcal{B}(\mathbb{R}^d)}, \qquad \forall\, t \geq 0. \tag{1.4.9}$$

As a consequence,

$$||e^{-Lt}||_{\mathcal{B}(V_0^{\perp})} = ||e^{-\widetilde{C}t}||_{\mathcal{B}(\mathbb{R}^d)}, \qquad \forall\, t \geq 0. \tag{1.4.10}$$

As a application of Theorem 1.4.2 one can observe that the study of the decay estimates for the PDE (1.4.1) can be reduced at the level of a finite dimensional ODE, that clearly requires less effort. In particular, from equality (1.4.10) it follows that the best decay rate and the best constant for the hypocoercivity estimates of the FP-equation (1.4.1) and the ODE $\frac{d}{dt}x(t) = -\widetilde{C}x(t)$ coincide. For example in the two dimensional case the explicit formulas for $c_{min}^{(1)}$ derived in [3] can be used to obtain the optimal multiplicative constant $c_{min}^{(2)}$ for the related FP-equation.

Corollary 1.4.3. Let $C \in \mathbb{R}^d$ be non-defective and satisfy *Condition A*. Then the best constant $c_{min}^{(1)}$ for the ODE $\frac{d}{dt}x(t) = -\widetilde{C}x(t)$ is also the optimal constant $c_{min}^{(2)}$ for the following hypocoercive estimate

$$||f_t - f_\infty||_{\mathcal{H}} \leq c_{min}^{(1)} e^{-\lambda_{opt}^{(2)} t} ||f_0 - f_\infty||_{\mathcal{H}}, \qquad \forall\, t \geq 0, \forall\, f_0 \in \mathcal{H}, \int_{\mathbb{R}^d} f_0\, dx = 1. \tag{1.4.11}$$

However, the applications of Theorem 1.4.2 are not just confined to the analysis of the large time behaviour of the FP-equation. They also concern the short time behaviour of the solutions of (1.4.1), involving the concept of *hypocoercivity index* m_{HC}. This value, defined both for FP-evolutions and for ODEs, represents the degeneracy structures of such evolutions and it corresponds to the polynomial degree in the short time of their propagator norms. In other words, m_{HC} describes how fast the propagator norm decays for short times. The bigger m_{HC}, the slower is the decay. Equation (1.4.10) makes the computation of the m_{HC} for FP-equations straightforward since it can be reduced to the ODE-level. More details about the hypocoercivity index and the application of Theorem 1.4.2 in this regards can be found in Chapter 3.

The proof of Theorem 1.4.2 does not rely on classical entropy methods. We rather decomposed the normalized FP-equation into orthonormal subspaces of $\widetilde{\mathcal{H}}$ that are invariant under the action of the operator $\widetilde{L}$. It turned out that the FP-equation can be depicted as the tensor version of its corresponding drift-ODE. More precisely, we wrote the FP operator L in terms of *second quantization* formalism. This allowed to display the explicit connection between the above mentioned FP-equation and the corresponding ODE and then to prove Theorem 1.4.2 in a straightforward way.

1.4.2 Optimality of convergence of the FP-equation to a given equilibrium

We have so far discussed the optimality of the convergence of a *given* FP-equation, i.e. when the diffusion matrix D and the drift matrix C are fixed. In [7], we adopted instead a different prospective, even though we still focused our attention on the optimality of decay estimates for the FP-equation. Namely, we considered the equilibrium $f_\infty \in \mathcal{H}$ to be fixed and we aimed to find the "optimal" FP-equation among the family of non symmetric FP-equations with linear drift that converge to that given equilibrium. The origin of this question comes from a statistical and probabilistic setting and it involves Markov Chain Monte Carlo (MCMC) algorithms. In this language the question is: For a given exponential probability measure which Markov process that samples that measure converges the fastest to it? Since the speed of this convergence determines the efficiency of (MCMC)-algorithms, one is motivated to find the optimal evolution that provides the fastest process converging to the fixed measure. For more details about the origin of the question posed here we refer the reader to Chapter 4.

We shall here display the setting and the main result achieved in [7] compared to the previous literature. The full version of [7] can be found in Chapter 4.

In the following we will denote with $\mathcal{M}$ the set of real matrices in $\mathbb{R}^{d\times d}$, $\mathcal{S}^{>0}$ (resp. $\mathcal{S}^{\geq 0}$) the set of positive definite (resp. positive semi-definite) symmetric matrices.

Given K in $\in \mathcal{S}^{>0}$, we define the (typically) anisotropic Gaussian

$$f_{\infty,K}(x) := \frac{\det(K)^{-1/2}}{(2\pi^{d/2})} \exp\left(-\frac{x^T K^{-1} x}{2}\right), \qquad x \in \mathbb{R}^d, \tag{1.4.12}$$

and the linear (possibly degenerate) FP-equation

$$\partial_t f_t = -L_{C,D} f_t := \operatorname{div}_x \left(D\nabla_x f t + Cx f_t\right), \qquad \forall t \geq 0, \tag{1.4.13}$$

for arbitrary x-independent matrices $D \in \mathcal{S}^{>0}$ and $C \in \mathcal{M}$. As we already know form the previous sections, the following hypocoercive estimate holds

$$||f_t - f_\infty||_{L^2(\mathbb{R}^d, f_\infty^{-1})} \leq c e^{-\lambda t} ||f_t - f_\infty||_{L^2(\mathbb{R}^d, f_\infty^{-1})}, \qquad t \geq 0, \tag{1.4.14}$$

if C is positive stable and the hypothesis of hypoellipticity is satisfied.

We report here the main questions posed in [7]:

(Q1) Which FP-evolution(s) converge(s) the fastest, i.e. with largest decay rate λ_{opt} to the steady state in the operator norm of $e^{-L_{C,D}t}$ on $V_0^\perp \subset \mathcal{H} := L^2(\mathbb{R}^d, f_{\infty,K}^{-1})$?

(Q2) When the best decay rate is fixed, what is the infimum of the multiplicative constant, c_{inf}, in the decay estimates (1.4.14)?

(Q3) For a fixed $K \in \mathscr{S}^{>0}$ and the corresponding λ_{opt}, and for any $c > c_{inf}$, which pair(s) of matrices $(C_{opt}, D_{opt}) \in \mathscr{M} \times \in \mathscr{S}^{\geq 0}$ are such that $e^{-L_{C_{opt},D_{opt}} t}$ yields the convergence estimate (1.4.14) with the constants (λ_{opt}, c)?

(Q4) For such an optimal pair of matrices, what bound on C_{opt} can be found, and how does this bound grow w.r.t. to the space dimension d?

(Q5) Could something be gained by allowing for time-dependent matrices $C(t)$, $D(t)$?

The question of finding the optimal degenerate FP-equation for a given steady state was the main goal of the study carried out in [10]. The authors gave a complete answer about the best decay rate λ_{opt} (Q1) but they did not achieve an optimal lower bound for the multiplicative constant c_{inf}. These last issues (Q2 and Q3) represent the main improvement of our result compared to the previous literature. Moreover, we improved the estimates of the Frobenius norm of the optimal matrix pair (Q4) and gave a hint of a possible improvement by considering time-dependent matrices $(C(t), D(t))$ (Q5). The results achieved in [7] rely heavily on Theorem 1.4.2 that allows us to consider computations for the FP-equation at the level of its corresponding ODE. We shall present in the following our main result in [7] (see Theorem 4.3.1 in Chapter 4) and mention the main tools used in the proof.

As a first step we shall classify the matrix pairs $(C, D) \in \mathscr{M} \times \mathscr{S}^{\geq 0}$ that are *admissible* for our question (as done in [10]). With this term we refer to the family of matrix pairs (C, D) such that $e^{-L_{C,D} t}$ converges to the prescribed steady state $f_{\infty,K}$, i.e.

$$(C, D) \in \mathscr{I}(K) := \{(C, D) \in \mathscr{M} \times \mathscr{S}^{\geq 0},\ \mathrm{Tr}(D) \leq d : L_{C,D} f_{\infty,K} = 0\}, \tag{1.4.15}$$

with the further assumption of C to be positive stable.

Remark 1.4.4. We observe that, without the request of an upper bound for the trace of the diffusion matrix, the question of finding the optimal decay rate would be ill-posed. Indeed, if f_t converges to $f_{\infty,K}$ then also f_t^{α}, for any $\alpha > 0$ converge to the same equilibrium but α times faster.

Thanks to Lemma 2.1 in [7] the admissible matrix pairs can be characterize by stating that

$$\mathscr{I}(K) := \{((D + J)K^{-1}, D) \in \mathscr{M} \times \mathscr{S}^{\geq 0} :\ J \in \mathscr{A}, \mathrm{Tr}(D) \leq d\}. \tag{1.4.16}$$

By considering that the best decay rate for a *given* FP-equation (1.4.13) corresponds to $\rho(C)$ (excluding the defective case), we are interested in investigating the maximum spectral gap of drift matrices of the form $C = (D + J)K^{-1}$. This value has been already computed in [10], as

$$\max\{\rho(C)\ :\ (C, D) \in \mathscr{I}(K)\} = \max(\sigma(K^{-1})). \tag{1.4.17}$$

In [7] we proved that the infimum of the multiplicative constant for the decay estimates (1.4.14) with best decay rate λ_{opt} is $c_{inf} = 1$. This is due to the fact that the FP-equation (1.4.13) is in general non-symmetric and degenerate. The symmetric case would indeed admit a constant equals to one but a worst decay rate that cannot overcome the value $\lambda = \min(\sigma(K^{-1})$.

Given a constant $c > 1$, in order to answer (Q1)-(Q2)-(Q3) one has to find *one* explicit pair (C_{opt}, D_{opt}) that yields the optimal decay rate $\lambda_{opt} = \max(\sigma(K^{-1}))$ and admits c as multiplicative constant. This is the achievement of the next theorem, corresponding to the main result in [7] (Theorem 3.1), where also estimates of the Frobenius norm of such optimal pair are provided.

Theorem 1.4.5. *Let $K \in \mathscr{S}^{>0}$ be given.*

(a) Then, for any constant $c > 1$ there exists a pair $(C_{opt}, D_{opt}) = (C_{opt}(c), D_{opt}(c)) \in \mathscr{I}(K)$ such that

$$\left|\left|e^{-L_{C_{opt},D_{opt}}t}\right|\right|_{\mathscr{B}(V_0^{\perp})} \leq c e^{-\max(\sigma(K^{-1}))t}, \qquad t \geq 0. \tag{1.4.18}$$

(b) The matrices from part (a) can be estimated as

$$||C_{opt}||_{\mathscr{F}} \leq \lambda_{opt}\left[d + \sqrt{\kappa(K)}\,\frac{2\pi c^2}{\sqrt{3}(c^2-1)}\,\sqrt{d}\,(d-1)\right], \quad ||D_{opt}||_{\mathscr{F}} = d. \tag{1.4.19}$$

There are mainly three directions of improvement of Theorem 1.4.5 with respect to the main result in [10]. First, the infimum for the multiplicative constant, which improved from $c = \sqrt{\kappa(K)e}$ to $c_{min} = 1$. Moreover, the evolution in [10] needed to be split in time, in order to avoid a dramatic growth of the constant with the dimension of order d^{40d^2}. Finally, concerning question (Q4), Part (b) of Theorem 1.4.5 shows that our optimal linear drift grows like $\mathscr{O}(d^{3/2})$, compared to the $\mathscr{O}(d^2)$-growth in [10].

The proof of Theorem 1.4.5 builds upon the proof of Theorem 2.2 in [10]. It differs from it only in the conclusive part. Here the main key tool is represented by equality (1.4.10) between the propagator norm of the FP-equation and that of the corresponding ODE in Theorem 4.3.1. This allows to enhance both the best decay rate and a multiplicative constant arbitrarily close to one in a straightforward way, in contrast to the methods used in [10].

1.5 The Goldstein-Taylor Model with Space-Dependent Relaxation

In this section we shall present the problem studied in [6] that regards the large time convergence of the Goldstein-Taylor (GT) model with x-dependent relaxation on the one-dimensional torus. Given the technicality of this work, we prefer to give a comprehensive overview of the topic, rather than stating the main result. In particular, we shall focus on the setting, on the heuristic interpretation of the GM model, and on the novelty of the above mentioned paper. The work done in [6] is, nevertheless, fully reported in Chapter 5.

Notation. We will denote with $\mathbb{T}$ the one-dimensional torus, i.e. the interval $[0, 2\pi] \in \mathbb{R}$ with periodic boundary conditions, and with $L^2(\mathbb{T})^{\otimes 2}$ the set of real valued vectorial functions $F = (f_1, f_2)$, $f_i \in L^2(\mathbb{T})$ endowed with the standard inner product for each component. Moreover, we define the functional space

$$L^{\infty}_{+}(\mathbb{T}) := \left\{ f \in L^{\infty}(\mathbb{T}) \mid \operatorname{essmin} f > 0 \right\} \tag{1.5.1}$$

Lastly, we introduce the notation

$$h_{avg} := \frac{1}{2\pi}\int_{\mathbb{T}} h(x)\,dx. \tag{1.5.2}$$

The GT-system with space-dependent relaxation in $\mathbb{T} \times (0, +\infty)$ is given by

$$
\begin{aligned}
\partial_t f_+(x,t) + \partial_x f_+(x,t) &= \frac{\sigma(x)}{2}(f_-(x,t) - f_+(x,t)),\\
\partial_t f_-(x,t) - \partial_x f_-(x,t) &= -\frac{\sigma(x)}{2}(f_-(x,t) - f_+(x,t)),\\
f_\pm(x,0) &= f_{\pm,0}(x) \in L^2(\mathbb{T}),
\end{aligned}
\tag{1.5.3}
$$

where the function $\sigma(x) \in L^\infty_+(\mathbb{T})$ denotes the *relaxation* function. For any $t \in (0,\infty), x \in \mathbb{T}$ the functions $f_\pm(x,t)$ represent the probability of finding a particle of the system with velocity ± 1 in the position x at time t, respectively. In fact the GT-model was originally derived as limit of a random migration in 1D, where particles could change direction among a discrete set of values with rate σ. It can also be seen as a simplification of the BGK-model, where the velocities can assume only the two values ± 1. The challenging question that moved the interest of many works in the literature about the GT-model concerns the converge to the equilibrium for the solutions. The difficulty in this regard is represented by the lack of coercivity for the model. The GT-model represents, instead, an exemplifying example of the already mentioned interplay between the two parts that characterize hypocoercive dynamics: the diffusive and the conservative components. In this case the diffusion is represented by the r.h.s, where the relaxation acts as a "vertical force" that tends to spread the mass densities on the space and to reduce the local-in-x mass difference between the two kinds of particles (the ones with positive and the ones with negative velocity). Nevertheless, the lack of coercivity corresponds to the degeneracy of the relaxation force. This translates to the fact that the r.h.s operator alone would only provide a *local* equilibrium, for a fixed $x \in T$. On the other hand, the transport operator in the l.h.s is a conservative force that drives "horizontally" the particles and flattens out the inhomogeneity of the mass density over the torus. The interplay of these two forces of different natures enables the convergence to a unique global equilibrium $F_\infty = (f_\infty, f_\infty)^T \in L^2(\mathbb{T})^{\otimes 2}$ that we shall reveal in the following.

Remark 1.5.1. In contrast to the FP-equation, the GT-Taylor model is not hypoelliptic and it does not benefit from a regularization effect. Initial data in $L^2(\mathbb{T})^{\otimes 2}$ remain in general in this functional space for positive times. Therefore we cannot refer throughout this section to classical solutions to (1.5.3), but rather to *mild* solutions.

In order to better understand the properties of the GT-model it is helpful to recast it in the *macroscopic variables*

$$
u := f_+ + f_-, \quad v := f_+ - f_-, \tag{1.5.4}
$$

which represent, respectively, the spatial mass density and the flux density. In these new variables the GT-model is given as

$$
\begin{aligned}
\partial_t u(x,t) + \partial_x v(x,t) &= 0,\\
\partial_t v(x,t) + \partial_x u(x,t) &= -\sigma(x) v(x,t),\\
u(.,0) = u_0 := f_{+,0} + f_{-,0}, &\quad v(.,0) = v_0 := f_{+,0} - f_{-,0}\,.
\end{aligned}
\tag{1.5.5}
$$

From the equation for u one can deduce the mass conservation of the system, i.e.

$$
u(t)_{\text{avg}} = \frac{1}{2\pi}\int_{\mathbb{T}} f_+(x,t) + f_-(x,t)\,dx \equiv \frac{1}{2\pi}\int_{\mathbb{T}} f_{+,0}(x) + f_{-,0}(x,)\,dx = (u_0)_{\text{avg}} := u_{\text{avg}} \quad \forall t \geq 0. \tag{1.5.6}
$$

On the other hand, the equation for v implies that the difference between the two densities decays to zero for $t \to +\infty$ with a "fast" decay. Considering these two properties the unique steady state for (1.5.5) is given by

$$u_\infty = u_{\text{avg}}, \qquad v_\infty = 0, \tag{1.5.7}$$

and therefore

$$f_{\pm}\infty = \frac{1}{2}(f_{+,0} + f_{-,0})_{\text{avg}} := f_\infty. \tag{1.5.8}$$

The goal of our work [6] was to obtain explicit hypocoercive estimates in $\mathscr{H} = L^2(\mathbb{T})^{\otimes 2}$ for the solutions of (1.5.3) of the from

$$||F(t) - F_\infty||_{\mathscr{H}} \le ce^{-\lambda t}||F_0 - F_\infty||_{\mathscr{H}}, \qquad \forall\, t \ge 0, \forall F_0 \in \mathscr{H}, \tag{1.5.9}$$

with the vectorial notation

$$F(t) := (f_+(\cdot,t), f_-(\cdot,t))^T, \quad F_0 := F(0), \quad F_\infty := (f_\infty, f_\infty)^T,$$

by developing a method that could be generalized for more complex cases. The best decay rate and the optimal multiplicative constant for (1.5.9) will be denoted in the following, respectively, with $\lambda_{opt}^{(4)}$ and $c_{min}^{(4)}$.

Many recent works have been devoted to the study of the convergence of the solutions of (1.5.3) to the equilibrium. Most of the results concern, nevertheless, the model with constant relaxation (see [1]). For this simplified setting the entropy decay estimates can be obtained by considering the modal decomposition of the GT-model in the Fourier space and its spectral analysis. In the case of x-dependent relaxation functions $\sigma(x)$ this approach cannot be used. One example of a study that analysed the GT-model with non constant relaxation is given by the work of Francesco Salvarani [13]. Therein, exponential decay estimates are provided by using the equivalence between the GT-model and the Telegraph's equation. The main result in [13] states that the optimal decay rate for the decay estimates (1.5.9) is

$$\lambda_{opt}^{(4)} = \min\{\sigma_{\text{avg}}, \rho(T)\}, \tag{1.5.10}$$

where $\rho(T)$ denotes the spectral gap of the Telegram's equation. The explicit value $\rho(T)$ is, however, hardly accessible, even for simple non-constant relaxation functions $\sigma(x)$. Moreover, the result strictly requires $H^1(T)$-regularity of the initial data and it cannot be extended to more general settings. This is due to the proof heavily relying on the Telegraph's equation.

In [6] we aimed to solve these issues. As a result we developed a robust method that provides an explicit, quantitative lower bound for the decay rate for the GT-model in $L^2(\mathbb{T})^{\otimes 2}$ with a general bounded non-homogeneous $\sigma(x)$. The provided decay rate is optimal in the case where σ is constant but only suboptimal in the general case. On the other hand, we showed in [6] how this method can be extended to multi-velocities GT-model, proving its robustness. In this regard we provided the explicit decay rate of convergence for a simple 3-velocity evolution, by adapting the original functional (the one used for the 2-velocities model) to this more general case.

Considering that the GT-evolution is not coercive in $\mathscr{H}$, entropy methods have been used in [6] to obtain explicit decay rates of convergence. We constructed a modified norm, equivalent to $||\cdot||_{\mathscr{H}}$ whose time derivative along the trajectories of (1.5.3) decays with pure exponential decay.

Although the standard method to tackle the question of convergence with a constant σ cannot be generalized, the idea of using the Fourier decomposition inspired our methodology. Roughly speaking, we first decomposed the GT-model into infinity many decoupled ODE-systems, governed by degenerate matrices C_k in the Fourier space. For any Fourier mode $k \in \mathbb{Z}$ we derived a *modal entropy* in terms of a positive matrix P_k equivalent to the Euclidean norm that yields a strict exponential decay. Afterwards we translated such modal entropy to a spacial entropy that is modal-independent, by applying the inverse of the Fourier transform operator. Following this strategy, we defined for any given $\theta > 0$ the following *pseudodifferential* Lyapunov functional

$$E_\theta(f,g) := ||f||^2_{L^2(\mathbb{T})} + ||g||^2_{L^2(\mathbb{T})} - \frac{\theta}{2\pi}\int_T \partial_x^{-1} f(x) g(x) dx, \qquad \forall f,g \in L^2(\mathbb{T}). \tag{1.5.11}$$

Here the *anti-derivative* of f is defined as

$$\partial_x^{-1} f(x) := \int_0^x f(y)dy - \left(\int_0^x f(y)dy\right)_{\text{avg}}, \tag{1.5.12}$$

with the normalization constant chosen such that $(\partial_x^{-1} f) = 0$. In Lemma 3.4, [6] we first showed the equivalence between the norm induced by E_θ and the Hilbertian norm $||\cdot||_{\mathcal{H}}$. Moreover, by differentiating in time the entropy along the trajectories we proved the pure exponential decay

$$E_\theta(u(t) - u_{\text{avg}}, v(t)) \le E_\theta(u_0 - u_{\text{avg}}, v_0)e^{-2\lambda t} t. \tag{1.5.13}$$

The parameter θ was chosen in a suitable way as a function of $\sigma_{\min}$ and $\sigma_{\max}$ with the notation

$$0 < \sigma_{\min} := \min_{x\in\mathbb{T}} \sigma(x) \le \max_{x\in\mathbb{T}} \sigma(x) := \sigma_{\max}, \tag{1.5.14}$$

according to the situation. In fact, in order to use the above mentioned entropy methods one has to consider in general (for a x-dependent σ) two subcases depending on the values of $\sigma(x)$. In particular the two cases are distinguished according to whether $\sigma_{\min}$ is larger or smaller than the threshold $4/\sigma_{\max}$. Furthermore, since the function $\sigma(x)$ is representing the strength of the relaxation force we expect also the decay rate to be influenced by the values $\sigma_{\max}$ and $\sigma_{\min}$. This phenomenon can be read in the main result of [6], Theorem 2.2 (see Theorem 5.2.2). For both subcases it provides the explicit decay rate and the multiplicative constant for the estimates (1.5.9) as a function of $(\sigma_{\min}, \sigma_{\max})$. These are, however, not optimal in general, due to the perturbative approach of this strategy. On the contrary, the functional E_θ is optimal (in the sense that it provides the best decay rate) for the constant case, by choosing θ in a suitable way according to the three cases: $\sigma < 2$, $\sigma > 2$ and $\sigma = 2$ (defective case).

1.6 Structure and Authorship

The main body of this thesis is divided into four chapters.

The main object of study in **Chapter** 2 is the optimality of decay estimates for linear ODE systems in $\mathbb{C}^d$ with positive stable matrices. In particular, we shall focus on the minimization of the multiplicative factor appearing in such estimates, which will be time-dependent in one subcase. We shall give a complete answer in the 2-dimensional case by providing an explicit form of the minimal constant and a partial answer for higher dimensional ODE systems. The content of this chapter is the joint work with Anton Arnold and Franz Achleitner ([3]). This paper was published in 2019. The author of this thesis was involved both by working out the mathematical details and throughout the whole drafting process of the paper. The author particularly focused on §2.1, §2.2, §2.3 and contributed to §2.5. The coauthors focused on §2.4 and contributed to §2.5.

Chapter 3 is devoted to the analysis of the short- and large-time behaviour of Fokker-Planck equations with linear drift in the setting of hypocoercivity. Under this assumption there is convergence to the unique steady state with exponential decay. The main theorem of this chapter states that the propagator norm of the FP-operator is equal to the one of its associated drift-ODE system. This simplifies significantly the research of optimal decay estimates for the FP-equation, by reducing it to a question at the ODE level. The proof is based on the decomposition of the FP-evolution on finite dimensional subspaces, in each of which the evolution is governed by a tensor version of the drift ODE. The content of this chapter is a joint work with Anton Arnold and Christian Schmeiser ([4]). This paper was submitted in 2020. The author of this thesis worked out the mathematical details and carried out the drafting of the thesis. The coauthors contributed by providing mathematical ideas, among which the one that motivated the paper, and in the process of drafting and reviewing.

In **Chapter** 4 we shall determine the optimal FP-equation with linear drift that converges the fastest to a prescribed steady state. We shall state that the best decay rate is represented by the maximum eigenvalue of the inverse covariance matrix, and the infimum for the multiplicative constant appearing in the decay estimates is 1. The content of this chapter is a joint work with Anton Arnold ([7]). This paper was submitted in 2021. The author of this thesis was involved both by working out the mathematical details and throughout the whole drafting process of the paper. The author was the main contributor of sections §4.2, §4.3, §4.4.1. The coauthor was both the main contributor of sections §4.2.1, §4.4.2, §4.4.3 and the one that provided the motivating idea for the paper.

Chapter 5 is focused on the decay estimates for the Goldstein Taylor model on the one-dimensional torus. We shall construct a Lyapunov functional that will provides an explicit decay rate when the relaxation coefficient is not constant in space. The advantage of this method lies in the fact that it can be generalized to more general settings. To this regard an explicit application to the multi-velocity Goldstein-Taylor model shall be shown at the end of this chapter. The content of this chapter is a joint work with Anton Arnold, Amit Einav and Tobias Wöhrer [6]. This paper was published in 2021. The author of this thesis contributed to the mathematical details while the drafting was mostly carried out by the coauthors, who also provided the motivating idea for the paper.

Bibliography

[1] Achleitner, F., Arnold, A., 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] F. Achleitner, A. Arnold, E. Carlen, *The Hypocoercivity Index for the short and large time behaviour of ODEs.* https://arxiv.org/abs/2109.10784 (2021).

[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 282 (2019), 241-264, G. Giacomin et al. (eds.), Springer .

[4] A. Arnold, C. Schmeiser, B. Signorello, *Propagator norm and sharp decay estimates for Fokker-Planck equations with linear drift.* Preprint. https://arxiv.org/abs/2003.01405 (2020).

[5] 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.

[6] A. Arnold, A. Einav, T. Wöhrer, B. Signorello, *Large time convergence of the non-homogeneous Goldstein-Taylor Equation.* Journal of Statistical Physics 182.2 (2021), 1-35.

[7] A.Arnold, B.Signorello, *Optimal non-symmetric Fokker-Planck equation for the convergence to a given equilibrium.* Preprint.https://arxiv.org/abs/2106.15742 (2021).

[8] J. Dolbeault, C. Mouhot, C. Schmeiser, *Hypocoercivity for linear kinetic equations conserving mass.* Trans. Amer. Math. Soc. 367 (2015), 3807-3828.

[9] L. Desvillettes and C. Villani, *On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems: the linear Fokker-Planck equation*, Comm. Pure Appl. Math. 54 (2001), 1-42.

[10] A. Guillin, P. Monmarché, *Optimal linear drift for the speed of convergence of an hypoelliptic diffusion.* Electron. Commun. Probab. 21, no. 74 (2016), 1-14; Erratum: Electron. Commun. Probab. 22, no. 15 (2017), 1-2.

[11] T. Lelièvre, F. Nier, G.A. Pavliotis, *Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion.* Journal of Statistical Physics, 152, no. 2 (2013), 237-274.

[12] L. Neumann, C. Schmeiser, *A kinetic reaction model: decay to equilibrium and macroscopic limit*, Kinetic and Related Models 9 (2016), 571-585.

[13] Salvarani, F. *Diffusion limits for the initial-boundary value problem of the Goldstein-Taylor model.* Rend. Sem. Mat. Univ. Polit. Torino vol. 57 (3), (1999), 211-222 .

[14] C. Villani, *Hypocoercivity*, Memoirs of the American Mathematical Society 202 (2009).

2 On Optimal Decay Estimates for ODEs and PDEs with modal decomposition

2.1 Introduction

This note is concerned with optimal decay estimates of hypocoercive evolution equations that allow for a modal decomposition. The notion *hypocoercivity* was introduced by Villani in [15] for equations of the form $\frac{d}{dt}f = -Lf$ on some Hilbert space H, where the generator L is not coercive, but where solutions still exhibit exponential decay in time. More precisely, there should exist constants $\lambda > 0$ and $c \geq 1$, such that

$$\|e^{-Lt} f^I\|_{\widetilde{H}} \leq c\, e^{-\lambda t} \|f^I\|_{\widetilde{H}} \qquad \forall\, f^I \in \widetilde{H}, \tag{2.1.1}$$

where $\widetilde{H}$ is a second Hilbert space, densely embedded in $(\ker L)^\perp \subset H$.

The large-time behaviour of many hypocoercive equations have been studied in recent years, including Fokker-Planck equations [15, 4, 3], kinetic equations [11] and BGK equations [1, 2]. Determining the sharp (i.e. maximal) exponential decay rate λ was an issue in some of these works, in particular [4, 1, 2]. But finding at the same time the smallest multiplicative constant $c \geq 1$, is so far an open problem. And this is the topic of this note. For simple cases we shall describe a procedure to construct the "optimal" Lyapunov functional that will imply (2.1.1) with the sharp constants λ and c.

For illustration purposes we shall focus here only on the following 2-velocity BGK-model (referring to the physicists Bhatnagar, Gross and Krook [7]) for the two functions $f_\pm(x,t) \geq 0$ on the one-dimensional torus $x \in \mathbb{T}$ and for $t \geq 0$. It reads

$$\begin{cases} \partial_t f_+ &= -\partial_x f_+ + \frac{1}{2}(f_- - f_+)\,, \\ \partial_t f_- &= \partial_x f_- - \frac{1}{2}(f_- - f_+)\,. \end{cases} \tag{2.1.2}$$

This system of two transport-reaction equations is also called *Goldstein-Taylor model.*

For initial conditions normalized as $\int_0^{2\pi} \left[f_+^I(x) + f_-^I(x)\right][x] = 2\pi$, the solution $f(t) = (f_+(t), f_-(t))^\top$ converges to its unique (normalized) steady state with $f_+^\infty = f_-^\infty = \frac{1}{2}$. The operator norm of the propagator for (2.1.2) can be computed explicitly from the Fourier modes, see [13]. By contrast, the goal of this paper and of [1, 11] is to refrain from explicit computations of the solution and to use Lyapunov functionals instead. Following this strategy, an explicit exponential decay rate of this two velocity model was shown in [11, §1.4]. The sharp exponential decay estimate was found in [1, §4.1] via a refined functional, yielding the following result:

Theorem 2.1.1 ([1, Th. 6])**.** *Let $f^I \in L^2(0,2\pi;\mathbb{R}^2)$. Then the solution to* (2.1.2) *satisfies*

$$\|f(t) - f^\infty\|_{L^2(0,2\pi;\mathbb{R}^2)} \leq c\, e^{-\lambda t} \|f^I - f^\infty\|_{L^2(0,2\pi;\mathbb{R}^2)}\,, \qquad t \geq 0\,,$$

with the optimal constants $\lambda = \frac{1}{2}$ and $c = \sqrt{3}$.

Remark 2.1.2. a) Actually, the optimal c was not specified in [1], but will be the result of Theorem 2.3.7 below.

b) As we shall illustrate in §2.5, it does *not* make sense to optimize these two constants at the same time. The optimality in Theorem 2.1.1 refers to first maximizing the exponential rate λ, and then to minimize the multiplicative constant c.

The proof of Theorem 2.1.1 is based on the spatial Fourier transform of (2.1.2), cf. [11, 1]. We denote the Fourier modes in the discrete velocity basis $\{\binom{1}{1}, \binom{1}{-1}\}$ by $u_k(t) \in \mathbb{C}^2$, $k \in \mathbb{Z}$. They evolve according to the ODE systems

$$\frac{d}{dt} u_k = -\mathbf{C}_k\, u_k\,, \quad \mathbf{C}_k = \begin{pmatrix} 0 & ik \\ ik & 1 \end{pmatrix}, \quad k \in \mathbb{Z}\,, \tag{2.1.3}$$

and their (normalized) steady states are

$$u_0^\infty = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \qquad u_k^\infty = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad k \neq 0\,.$$

In the main body of this note we shall construct appropriate Lyapunov functionals for such ODEs, in order to obtain sharp decay rates of the form (2.1.1). In the context of the BGK-model (2.1.2), combining such decay estimates for all modes u_k then yields Theorem 2.1.1, as they are uniform in k. We remark that the construction of Lyapunov functionals to reveal optimal decay rates in ODEs was already included in the classical textbook [6, §22.4], but optimality of the multiplicative constant c was not an issue there.

In this article we shall first review, from [1, 2], the construction of Lyapunov functionals for linear first order ODE systems that reveal the sharp decay rate. They are quadratic functionals represented by some Hermitian matrix $\mathbf{P}$. As these functionals are not uniquely determined, we shall then discuss a strategy to find the "best Lyapunov" functional in §2.3—by minimizing the condition number $\kappa(\mathbf{P})$. The method of §2.3 always yields an upper bound for the minimal multiplicative constant c and the sharp constant in certain subcases (see Theorem 2.3.7). The refined method of §2.4 covers another subclass (see Theorem 2.4.1). Overall we shall determine the optimal constant c for 2-dimensional ODE systems, and give estimates for it in higher dimensions. In the final section §2.5 we shall illustrate how to obtain a whole family of decay estimates—with suboptimal decay rates, but improved constant c. For small time this improves the estimate obtained in §2.3.

2.2 Lyapunov Functionals for Hypocoercive ODEs

In this section we review decay estimates for linear ODEs with constant coefficients of the form

$$\begin{cases} \frac{d}{dt} f = -\mathbf{C} f, \quad t \geq 0\,, \\ f(0) = f^I \in \mathbb{C}^n\,, \end{cases} \tag{2.2.1}$$

for some (typically non-Hermitian) matrix $\mathbf{C} \in \mathbb{C}^{n\times n}$. To ensure that the origin is the unique asymptotically stable steady state, we assume that the matrix $\mathbf{C}$ is *hypocoercive* (i.e. positive stable, meaning that all eigenvalues have positive real part). Since we shall *not* require that

$\mathbf{C}$ is coercive (meaning that its Hermitian part would be positive definite), we *cannot* expect that all solutions to (2.2.1) satisfy for the Euclidean norm: $\|f(t)\|_2 \le e^{-\widetilde{\lambda}t}\|f^I\|_2$ for some $\widetilde{\lambda} > 0$. However, such an exponential decay estimate does hold in an adapted norm that can be used as a Lyapunov functional.

The construction of this Lyapunov functional is based on the following lemma:

Lemma 2.2.1 ([1, Lemma 2], [4, Lemma 4.3])**.** *For any fixed matrix* $\mathbf{C} \in \mathbb{C}^{n\times n}$*, let* $\mu := \min\{\mathrm{Re}(\lambda)|\lambda$ *is an eigenvalue of* $\mathbf{C}\}$*. Let* $\{\lambda_j | 1 \le j \le j_0\}$ *be all the eigenvalues of* $\mathbf{C}$ *with* $\mathrm{Re}(\lambda_j) = \mu$*. If all* λ_j *(*$j = 1, \dots, j_0$*) are non-defective*[1]*, then there exists a positive definite Hermitian matrix* $\mathbf{P} \in \mathbb{C}^{n\times n}$ *with*

$$\mathbf{C}^*\mathbf{P} + \mathbf{P}\mathbf{C} \ge 2\mu\mathbf{P}\,, \tag{2.2.2}$$

but $\mathbf{P}$ *is not uniquely determined.*

Moreover, if all eigenvalues of $\mathbf{C}$ *are non-defective, examples of such matrices* $\mathbf{P}$ *satisfying* (2.2.2) *are given by*

$$\mathbf{P} := \sum_{j=1}^{n} b_j\, w_j \otimes w_j^*\,, \tag{2.2.3}$$

where $w_j \in \mathbb{C}^n$ *(*$j = 1, \dots, n$*) denote the (right) normalized eigenvectors of* $\mathbf{C}^*$ *(i.e.* $\mathbf{C}^* w_j = \bar{\lambda}_j w_j$*), and* $b_j \in \mathbb{R}^+$ *(*$j = 1, \dots, n$*) are arbitrary weights.*

For $n = 2$ all positive definite Hermitian matrices $\mathbf{P}$ satisfying (2.2.2) have the form (2.2.3), but for $n \ge 3$ this is not true (see Lemma 2.3.1 and Example 2.3.2, respectively).

In this article, for simplicity, we shall only consider the case when all eigenvalues of $\mathbf{C}$ are non-defective. For the extension of Lemma 2.2.1 and of the corresponding decay estimates to the defective case we refer to [3, Prop. 2.2] and [5].

Due to the positive stability of $\mathbf{C}$, the origin is the unique and asymptotically stable steady state $f^\infty = 0$ of (2.2.1): Due to Lemma 2.2.1, there exists a positive definite Hermitian matrix $\mathbf{P} \in \mathbb{C}^{n\times n}$ such that $\mathbf{C}^*\mathbf{P} + \mathbf{P}\mathbf{C} \ge 2\mu\mathbf{P}$ where $\mu = \min \mathrm{Re}(\lambda_j) > 0$. Thus, the time derivative of the adapted norm $\|f\|_{\mathbf{P}}^2 := \langle f, \mathbf{P}f\rangle$ along solutions of (2.2.1) satisfies

$$\frac{d}{dt}\|f(t)\|_{\mathbf{P}}^2 \le -2\mu\|f(t)\|_{\mathbf{P}}^2\,.$$

Hence the evolution becomes a contraction in the adapted norm:

$$\|f(t)\|_{\mathbf{P}}^2 \le e^{-2\mu t}\|f^I\|_{\mathbf{P}}^2\,, \qquad t \ge 0\,. \tag{2.2.4}$$

Clearly, this procedure can yield the sharp decay rate μ, only if $\mathbf{P}$ satisfies (2.2.2).

Next we translate this decay in $\mathbf{P}$-norm into a decay in the Euclidean norm:

$$\|f(t)\|_2^2 \le (\lambda_{\min}^{\mathbf{P}})^{-1}\|f(t)\|_{\mathbf{P}}^2 \le (\lambda_{\min}^{\mathbf{P}})^{-1} e^{-2\mu t}\|f^I\|_{\mathbf{P}}^2 \le \kappa(\mathbf{P})\, e^{-2\mu t}\|f^I\|_2^2\,, \quad t \ge 0\,, \tag{2.2.5}$$

where $0 < \lambda_{\min}^{\mathbf{P}} \le \lambda_{\max}^{\mathbf{P}}$ are, respectively, the smallest and largest eigenvalues of $\mathbf{P}$, and $\kappa(\mathbf{P}) = \lambda_{\max}^{\mathbf{P}}/\lambda_{\min}^{\mathbf{P}}$ is the (numerical) condition number of $\mathbf{P}$ with respect to the Euclidean norm. While (2.2.4)

[1] An eigenvalue is defective if its geometric multiplicity is strictly less than its algebraic multiplicity.

is sharp, (2.2.5) is not necessarily sharp: Given the spectrum of $\mathbf{C}$, the exponential decay rate in (2.2.5) is optimal, but the multiplicative constant not necessarily. For the optimality of the chain of inequalities in (2.2.5) we have to distinguish two scenarios: Does there exist an initial datum f^I such that each inequality will be (simultaneously) an equality for some *finite* $t_0 \geq 0$? Or is this only possible asymptotically as $t \to \infty$? We shall start the discussion with the former case, which is simpler, and defer the latter case to §2.4. The first scenario allows to find the optimal multiplicative constant for $\mathbf{C} \in \mathbb{R}^{2\times 2}$, based on (2.2.5). But in other cases it may only yield an explicit upper bound for it, as we shall discuss in §2.4.

Concerning the first inequality of (2.2.5), a solution $f(t_0)$ will satisfy $\|f(t_0)\|_2^2 = (\lambda_{\min}^{\mathbf{P}})^{-1}\|f(t_0)\|_{\mathbf{P}}^2$ for some $t_0 \geq 0$ only if $f(t_0)$ is in the eigenspace associated to the eigenvalue $\lambda_{\min}^{\mathbf{P}}$ of $\mathbf{P}$. Moreover, the initial datum f^I satisfies $\|f^I\|_{\mathbf{P}}^2 = \lambda_{\max}^{\mathbf{P}}\|f^I\|_2^2$ if f^I is in the eigenspace associated to the eigenvalue $\lambda_{\max}^{\mathbf{P}}$ of $\mathbf{P}$. Finally we consider the second inequality of (2.2.5): If the matrix $\mathbf{C}$ satisfies, e.g., $\operatorname{Re}\lambda_j = \mu > 0$; $j = 1, \ldots, n$, with all eigenvalues non-defective, then we always have

$$\|f(t)\|_{\mathbf{P}}^2 = e^{-2\mu t}\|f^I\|_{\mathbf{P}}^2 \qquad \forall t \geq 0\,, \tag{2.2.6}$$

since (2.2.2) is an equality then. This is the case for our main example (2.1.3) with $k \neq 0$.

Since the matrix $\mathbf{P}$ is not unique, we shall now discuss the choice of $\mathbf{P}$ as to minimize the multiplicative constant in (2.2.5). To this end we need to find the matrix $\mathbf{P}$ with minimal condition number that satisfies (2.2.2). Clearly, the answer can only be unique up to a positive multiplicative constant, since $\widetilde{\mathbf{P}} := \tau\mathbf{P}$ with $\tau > 0$ would reproduce the estimate (2.2.5).

As we shall prove in §2.3, the answer to this minimization problem is very easy in 2 dimensions: The best $\mathbf{P}$ corresponds to equal weights in (2.2.3), e.g. choosing $b_1 = b_2 = 1$.

2.3 Optimal Constant via Minimization of the Condition Number

In this section, we describe a procedure towards constructing “optimal” Lyapunov functionals: For solutions $f(t)$ of ODE (2.2.1) they will imply

$$\|f(t)\|_2 \leq c\, e^{-\mu t}\|f^I\|_2 \tag{2.3.1}$$

with the sharp constant μ and partly also the sharp constant c.

We shall describe the procedure for ODEs (2.2.1) with positive stable matrices $\mathbf{C}$. For simplicity we confine ourselves to diagonalizable matrices $\mathbf{C}$ (i.e. all eigenvalues are non-defective). In this case, Lemma 2.2.1 states that there exist positive definite Hermitian matrices $\mathbf{P}$ satisfying the matrix inequality (2.2.2). Following (2.2.5), $\sqrt{\kappa(\mathbf{P})}$ is always an upper bound for the constant c in (2.3.1). Our strategy is now to minimize $\kappa(\mathbf{P})$ on the set of all admissible matrices $\mathbf{P}$. We shall prove that this actually yields the minimal constant c in certain cases (see Theorem 2.3.7). In 2 dimensions this minimization problem can be solved very easily thanks to Lemma 2.3.1 and Lemma 2.3.3:

Lemma 2.3.1. *Let* $\mathbf{C} \in \mathbb{C}^{2\times 2}$ *be a diagonalizable, positive stable matrix. Then all matrices* $\mathbf{P}$ *satisfying* (2.2.2) *are of the form* (2.2.3).

Proof. We use again the matrix $\mathbf{W}$ whose columns are the normalized (right) eigenvectors of $\mathbf{C}^*$ such that

$$\mathbf{C}^*\mathbf{W} = \mathbf{W}\mathbf{D}^*\,, \tag{2.3.2}$$

with $\mathbf{D} = \operatorname{diag}(\lambda_1^{\mathbf{C}}, \lambda_2^{\mathbf{C}})$ where $\lambda_j^{\mathbf{C}}$ ($j \in \{1,2\}$) are the eigenvalues of $\mathbf{C}$. Since $\mathbf{W}$ is regular, $\mathbf{P}$ can be written as

$$\mathbf{P} = \mathbf{W}\mathbf{D}\mathbf{W}^*\,,$$

with some positive definite Hermitian matrix $\mathbf{D}$. Then the matrix inequality (2.2.2) can be written as

$$2\mu\mathbf{W}\mathbf{D}\mathbf{W}^* \le \mathbf{C}^*\mathbf{W}\mathbf{D}\mathbf{W}^* + \mathbf{W}\mathbf{D}\mathbf{W}^*\mathbf{C} = \mathbf{W}(\mathbf{D}^*\mathbf{D} + \mathbf{D}\mathbf{D})\mathbf{W}^*\,.$$

This matrix inequality is equivalent to

$$0 \le (\mathbf{D}^* - \mu\,\mathbf{I})\mathbf{D} + \mathbf{D}(\mathbf{D} - \mu\,\mathbf{I})\,. \tag{2.3.3}$$

Next we order the eigenvalues $\lambda_j^{\mathbf{C}}$ ($j \in \{1,2\}$) of $\mathbf{C}$ increasingly with respect to their real parts, such that $\operatorname{Re}(\lambda_1^{\mathbf{C}}) = \mu$. Moreover, we consider

$$\mathbf{D} = \begin{pmatrix} b_1 & \beta \\ \overline{\beta} & b_2 \end{pmatrix}$$

where $b_1, b_2 > 0$ and $\beta \in \mathbb{C}$ with $|\beta|^2 < b_1 b_2$. Then the right hand side of (2.3.3) is

$$(\mathbf{D}^* - \mu\,\mathbf{I})\mathbf{D} + \mathbf{D}(\mathbf{D} - \mu\,\mathbf{I}) = \begin{pmatrix} 0 & (\lambda_2^{\mathbf{C}} - \lambda_1^{\mathbf{C}})\beta \\ \overline{(\lambda_2^{\mathbf{C}} - \lambda_1^{\mathbf{C}})\beta} & 2b_2\operatorname{Re}(\lambda_2^{\mathbf{C}} - \lambda_1^{\mathbf{C}}) \end{pmatrix} \tag{2.3.4}$$

with $\operatorname{Tr}[(\mathbf{D}^* - \mu\,\mathbf{I})\mathbf{D} + \mathbf{D}(\mathbf{D} - \mu\,\mathbf{I})] = 2b_2\operatorname{Re}(\lambda_2^{\mathbf{C}} - \lambda_1^{\mathbf{C}})$ and

$$\det[(\mathbf{D}^* - \mu\,\mathbf{I})\mathbf{D} + \mathbf{D}(\mathbf{D} - \mu\,\mathbf{I})] = -|\lambda_2^{\mathbf{C}} - \lambda_1^{\mathbf{C}}|^2|\beta|^2\,.$$

Condition (2.3.3) is satisfied if and only if $\operatorname{Tr}[(\mathbf{D}^* - \mu\,\mathbf{I})\mathbf{D} + \mathbf{D}(\mathbf{D} - \mu\,\mathbf{I})] \ge 0$ which holds due to our assumptions on $\lambda_2^{\mathbf{C}}$ and b_2, and $\det[(\mathbf{D}^* - \mu\,\mathbf{I})\mathbf{D} + \mathbf{D}(\mathbf{D} - \mu\,\mathbf{I})] \ge 0$. The last condition holds if and only if

$$\lambda_2^{\mathbf{C}} = \lambda_1^{\mathbf{C}} \quad \text{or } \beta = 0\,.$$

In the latter case $\mathbf{D}$ is diagonal and hence $\mathbf{P}$ is of the form (2.2.3). In the former case, (2.3.2) shows that $\mathbf{C} = \lambda_1^{\mathbf{C}}\,\mathbf{I}$, and the inequality (2.2.2) is trivial. Now any positive definite Hermitian matrix $\mathbf{P}$ has a diagonalization $\mathbf{P} = \mathbf{V}\mathbf{E}\mathbf{V}^*$, with a diagonal real matrix $\mathbf{E}$ and an orthogonal matrix $\mathbf{V}$, whose columns are –of course– eigenvectors of $\mathbf{C}$. Thus, $\mathbf{P}$ is again of the form (2.2.3). □ □

In contrast to this 2D result, in dimensions $n \ge 3$ there exist matrices $\mathbf{P}$ satisfying (2.2.2) which are not of form (2.2.3):

Example 2.3.2. Consider the matrix $\mathbf{C} = \operatorname{diag}(1,2,3)$. Then, all matrices

$$\mathbf{P}(b_1, b_2, b_3, \beta) = \begin{pmatrix} b_1 & 0 & 0 \\ 0 & b_2 & \beta \\ 0 & \beta & b_3 \end{pmatrix} \tag{2.3.5}$$

with positive b_j ($j \in \{1,2,3\}$) and $\beta \in \mathbb{R}$ such that $8b_2b_3 - 9\beta^2 \ge 0$, are positive definite Hermitian matrices and satisfy (2.2.2) for $\mathbf{C} = \operatorname{diag}(1,2,3)$ and $\mu = 1$. But the eigenvectors of $\mathbf{C}^*$ are the canonical unit vectors. Hence, matrices of form (2.2.3) would all be diagonal. □

Restricting the minimization problem to admissible matrices **P** of form (2.2.3) we find: Defining a matrix $\mathbf{W} := (w_1|\dots|w_n)$ whose columns are the (right) normalized eigenvectors of $\mathbf{C}^*$ allows to rewrite formula (2.2.3) as

$$\begin{aligned}\mathbf{P} &= \sum_{j=1}^{n} b_j\, w_j \otimes w_j^* = \mathbf{W}\,\mathrm{diag}(b_1, b_2, \dots, b_n)\mathbf{W}^* \\ &= \big(\mathbf{W}\,\mathrm{diag}(\sqrt{b_1}, \sqrt{b_2}, \dots, \sqrt{b_n})\big)\big(\mathbf{W}\,\mathrm{diag}(\sqrt{b_1}, \sqrt{b_2}, \dots, \sqrt{b_n})\big)^*\end{aligned} \tag{2.3.6}$$

with positive constants b_j $(j = 1, \dots, n)$. The identity

$$\mathbf{W}\,\mathrm{diag}(\sqrt{b_1}, \sqrt{b_2}, \dots, \sqrt{b_n}) = (\sqrt{b_1}\, w_1|\dots|\sqrt{b_n}\, w_n)$$

shows that the weights are just rescalings of the eigenvectors. Finally, the condition number of **P** is the squared condition number of $(\mathbf{W}\,\mathrm{diag}(\sqrt{b_1}, \sqrt{b_2}, \dots, \sqrt{b_n}))$. Hence, to find matrices **P** of form (2.3.6) with minimal condition number, is equivalent to identifying (right) precondition matrices among the positive definite diagonal matrices which minimize the condition number of **W**. This minimization problem can be formulated as a convex optimization problem [9] based on the result [14]. Due to [10, Theorem 1], the minimum is attained (i.e. an optimal scaling matrix exists) since our matrix **W** is non-singular. (Note that its column vectors form a basis of $\mathbb{C}^n$.) The convex optimization problem can be solved by standard software providing also the exact scaling matrix which minimizes the condition number of **P**, see the discussion and references in [9]. For more information on convex optimization and numerical solvers, see e.g. [8].

We return to the minimization of $\kappa(\mathbf{P})$ in 2 dimensions:

Lemma 2.3.3. *Let* $\mathbf{C} \in \mathbb{C}^{2\times 2}$ *be a diagonalizable, positive stable matrix. Then the condition number of the associated matrix* **P** *in* (2.2.3) *is minimal by choosing equal weights, e.g.* $b_1 = b_2 = 1$.

Proof. A diagonalizable matrix **C** has only non-defective eigenvalues. Up to a unitary transformation, we can assume w.l.o.g. that the eigenvectors of $\mathbf{C}^*$ are

$$w_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} \alpha \\ \sqrt{1-\alpha^2} \end{pmatrix} \quad \text{for some } \alpha \in [0,1). \tag{2.3.7}$$

This unitary transformation describes the change of the coordinate system. To construct the new basis, we choose one of the normalized eigenvectors w_1 as first basis vector, and recall that the second normalized eigenvector w_2 is only determined up to a scalar factor $\gamma \in \mathbb{C}$ with $|\gamma| = 1$. The right choice for the scalar factor γ allows to fulfill the above restriction on α.

We use the representation of the positive definite matrix **P** in (2.3.6):

$$\mathbf{P} = \mathbf{W}\,\mathrm{diag}(b_1, b_2)\mathbf{W}^* \quad \text{with } \mathbf{W} = \begin{pmatrix} 1 & \alpha \\ 0 & \sqrt{1-\alpha^2} \end{pmatrix}. \tag{2.3.8}$$

Since **P** and $\tau\mathbf{P}$ have the same condition number, we consider w.l.o.g. $b_1 = 1/b$ and $b_2 = b$. Thus, we have to determine the positive parameter $b > 0$ which minimizes the condition number of

$$\mathbf{P}(b) = \mathbf{W}\,\mathrm{diag}(1/b, b)\mathbf{W}^* = \begin{pmatrix} \frac{1}{b} + b\alpha^2 & b\alpha\sqrt{1-\alpha^2} \\ b\alpha\sqrt{1-\alpha^2} & b(1-\alpha^2) \end{pmatrix}. \tag{2.3.9}$$

The condition number of matrix $\mathbf{P}(b)$ is given by

$$\kappa(\mathbf{P}(b)) = \lambda_+^{\mathbf{P}}(b)/\lambda_-^{\mathbf{P}}(b) \geq 1\,,$$

where

$$\lambda_\pm^{\mathbf{P}}(b) = \frac{\operatorname{Tr}\mathbf{P}(b) \pm \sqrt{(\operatorname{Tr}\mathbf{P}(b))^2 - 4\det\mathbf{P}(b)}}{2}$$

are the (positive) eigenvalues of $\mathbf{P}(b)$. We notice that $\operatorname{Tr}\mathbf{P}(b) = b + 1/b$ is independent of α and is a convex function of $b \in (0,\infty)$ which attains its minimum for $b = 1$. Moreover, $\det\mathbf{P}(b) = 1 - \alpha^2$ is independent of b. This implies that the condition number

$$\kappa(\mathbf{P}(b)) = \frac{\lambda_+^{\mathbf{P}}(b)}{\lambda_-^{\mathbf{P}}(b)} = \frac{1 + \sqrt{1 - \frac{4\det\mathbf{P}(b)}{(\operatorname{Tr}\mathbf{P}(b))^2}}}{1 - \sqrt{1 - \frac{4\det\mathbf{P}(b)}{(\operatorname{Tr}\mathbf{P}(b))^2}}}$$

attains its unique minimum at $b = 1$, taking the value

$$\kappa_{\min} = \frac{1+\alpha}{1-\alpha}\,. \tag{2.3.10}$$

□ □

This 2D-result does not generalize to higher dimensions. In dimensions $n \geq 3$ there exist diagonalizable positive stable matrices $\mathbf{C}$, such that the matrix $\mathbf{P}$ with equal weights b_j does not yield the lowest condition number among all matrices of form (2.2.3). We give a counterexample in 3 dimensions:

Example 2.3.4. For some $\mathbf{C}^*$, consider its eigenvector matrix

$$\mathbf{W} := \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \operatorname{diag}\left(1, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{3}}\right), \tag{2.3.11}$$

which has normalized column vectors. We define the matrices $\mathbf{P}(b_1, b_2, b_3) := \mathbf{W}\operatorname{diag}(b_1, b_2, b_3)\mathbf{W}^*$ for positive parameters b_1, b_2 and b_3, which are of form (2.2.3) and hence satisfy the inequality (2.2.2). In case of equal weights $b_1 = b_2 = b_3$ the condition number is $\kappa(\mathbf{P}(b_1, b_1, b_1)) \approx 15.12825876$. But using [12, Theorem 3.3], the minimal condition number $\min_{b_j} \kappa(\mathbf{P}(b_1, b_2, b_3)) \approx 13.92820324$ is attained for the weights $b_1 = 2$, $b_2 = 4$ and $b_3 = 3$. □

Combining Lemma 2.3.1 and Lemma 2.3.3 we have

Corollary 2.3.5. Let $\mathbf{C} \in \mathbb{C}^{2\times 2}$ be a diagonalizable, positive stable matrix. Then the condition number is minimal among all matrices $\mathbf{P}$ satisfying (2.2.2), if $\mathbf{P}$ is of form (2.2.3) with equal weights, e.g. $b_1 = b_2 = 1$.

This 2D-result does not generalize to higher dimensions. Extending the conclusion of Example 2.3.4, we shall now show that $\mathbf{P}$ does not necessarily have to be of form (2.2.3), if its condition number should be minimal:

Example 2.3.6. We consider a special case of Example 2.3.4, with

$$\widetilde{\mathbf{C}} = (\mathbf{W}^*)^{-1}\operatorname{diag}(1,2,3)\mathbf{W}^*$$

with $\mathbf{W}$, the eigenvector matrix of $\widetilde{\mathbf{C}}^*$, given by (2.3.11). Then the matrices $\widetilde{\mathbf{C}}$ and

$$\widetilde{\mathbf{P}}(b_1,b_2,b_3,\beta) := \mathbf{W}\mathbf{P}(b_1,b_2,b_3,\beta)\mathbf{W}^*$$

with matrix $\mathbf{P}(b_1,b_2,b_3,\beta)$ in (2.3.5) satisfy the matrix inequality (2.2.2) with $\mu = 1$. But $\widetilde{\mathbf{P}}$ is not of form (2.2.3) if $\beta \neq 0$. Nevertheless, the condition number $\kappa(\widetilde{\mathbf{P}}(b_1,b_2,b_3,\beta)) \approx 5.82842780720132$ for the weights $b_1 = 2$, $b_2 = 4$, $b_3 = 3$, and $\beta = -2.45$, is much lower than with $\beta = 0$ (i.e. $\kappa(\widetilde{\mathbf{P}}(2,4,3,0)) \approx 13.92820324$, cf. Example 2.3.4). □

Lemma 2.3.3 and inequality (2.2.5) show that $\sqrt{\kappa_{\min}}$ from (2.3.10) is an *upper bound* for the best constant in (2.3.1) for the 2D case. For matrices with eigenvalues that have the same real part it actually yields the minimal multiplicative constant c, as we shall show now. Other cases will be discussed in §2.4.

For a diagonalizable matrix $\mathbf{C} \in \mathbb{C}^{2\times 2}$ with $\lambda_1^{\mathbf{C}} = \lambda_2^{\mathbf{C}}$ it holds that $\|f(t)\|_2 = e^{-\operatorname{Re}\lambda_1^{\mathbf{C}} t}\|f^I\|_2$. And for the general case we have:

Theorem 2.3.7. *Let* $\mathbf{C} \in \mathbb{C}^{2\times 2}$ *be a diagonalizable, positive stable matrix with eigenvalues* $\lambda_1^{\mathbf{C}} \neq \lambda_2^{\mathbf{C}}$*, and associated eigenvectors* v_1 *and* v_2*, resp. If the eigenvalues have identical real parts, i.e.* $\operatorname{Re}\lambda_1^{\mathbf{C}} = \operatorname{Re}\lambda_2^{\mathbf{C}}$*, then the condition number of the associated matrix* $\mathbf{P}$ *in* (2.2.3) *with equal weights, e.g.* $b_1 = b_2 = 1$*, yields the minimal constant in the decay estimate* (2.3.1) *for the ODE* (2.2.1)*:*

$$c = \sqrt{\kappa(\mathbf{P})} = \sqrt{\frac{1+\alpha}{1-\alpha}} \quad \text{where } \alpha := \left|\left\langle \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}\right\rangle\right|. \tag{2.3.12}$$

Proof. With the notation from the proof of Lemma 2.3.3 we have

$$\mathbf{P}(1) = \begin{pmatrix} 1+\alpha^2 & \alpha\sqrt{1-\alpha^2} \\ \alpha\sqrt{1-\alpha^2} & 1-\alpha^2 \end{pmatrix},$$

with the eigenvectors $y_+^{\mathbf{P}} = (\sqrt{1-\alpha^2}, 1-\alpha)^\top$, $y_-^{\mathbf{P}} = (\sqrt{1-\alpha^2}, -1-\alpha)^\top$. According to the discussion after (2.2.5) we choose the initial condition $f^I = y_+^{\mathbf{P}}$. From the diagonalization (2.3.2) of $\mathbf{C}$ we get

$$f(t) = (\mathbf{W}^*)^{-1} e^{-\mathbf{D}t}\mathbf{W}^* f^I.$$

Using (2.3.8) and $\mathbf{W}^* y_\pm^{\mathbf{P}} = \sqrt{1-\alpha^2}\binom{1}{\pm 1}$ we obtain directly that

$$f(t_0) = e^{-\lambda_1^{\mathbf{C}} t_0} y_-^{\mathbf{P}} \quad \text{with } t_0 = \frac{\pi}{|\operatorname{Im}(\lambda_2^{\mathbf{C}} - \lambda_1^{\mathbf{C}})|}.$$

Hence, also the first inequality in (2.2.5) is sharp at t_0. Sharpness of the whole chain of inequalities then follows from (2.2.6), and this finishes the proof. □ □

This theorem now allows us to identify the minimal constant c in Theorem 2.1.1 on the Goldstein-Taylor model: The eigenvalues of the matrices $\mathbf{C}_k$, $k \neq 0$ from (2.1.3) are $\lambda = \frac{1}{2} \pm i\sqrt{k^2 - \frac{1}{4}}$. The corresponding transformation matrices $\mathbf{P}_k$ with $b_1 = b_2 = 1$ are given by $\mathbf{P}_0 = \mathbf{I}$ and

$$\mathbf{P}_k = \begin{pmatrix} 1 & -\frac{i}{2k} \\ \frac{i}{2k} & 1 \end{pmatrix}, \quad \text{with} \quad \kappa(\mathbf{P}_k) = \frac{2|k|+1}{2|k|-1}, \qquad k \neq 0.$$

Combining the decay estimates for all Fourier modes $u_k(t)$ shows that the minimal multiplicative constant in Theorem 2.1.1 is given by $c = \sqrt{\kappa(\mathbf{P}_{\pm 1})} = \sqrt{3}$. For a more detailed presentation how to recombine the modal estimates we refer to §4.1 in [1].

2.4 Optimal Constant for 2D Systems

The optimal constant c in (2.3.1) for $\mathbf{C} \in \mathbb{C}^{2\times 2}$ with $\operatorname{Re}\lambda_1^{\mathbf{C}} = \operatorname{Re}\lambda_2^{\mathbf{C}}$ was determined in Theorem 2.3.7. In this section we shall discuss the remaining 2D cases. We start to derive the minimal multiplicative constant c for matrices $\mathbf{C}$ with eigenvalues that have distinct real parts but identical imaginary parts.

Theorem 2.4.1. *Let $\mathbf{C} \in \mathbb{C}^{2\times 2}$ be a diagonalizable, positive stable matrix with eigenvalues $\lambda_1^{\mathbf{C}}$ and $\lambda_2^{\mathbf{C}}$, and associated eigenvectors v_1 and v_2, resp. If the eigenvalues have distinct real parts $\operatorname{Re}\lambda_1^{\mathbf{C}} < \operatorname{Re}\lambda_2^{\mathbf{C}}$ and identical imaginary parts $\operatorname{Im}\lambda_1^{\mathbf{C}} = \operatorname{Im}\lambda_2^{\mathbf{C}}$, then the minimal multiplicative constant c in* (2.3.1) *for the ODE* (2.2.1) *is given by*

$$c = \frac{1}{\sqrt{1-\alpha^2}} \quad \textit{where } \alpha := \left|\left\langle \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}\right\rangle\right|. \tag{2.4.1}$$

Proof. We use again the unitary transformation as in the proof of Lemma 2.3.3, such that the eigenvectors w_1 and w_2 of $\mathbf{C}^*$ are given in (2.3.7). If $f(t)$ is a solution of (2.2.1), then $\tilde f(t) := e^{i \operatorname{Im}\lambda_1^{\mathbf{C}} t} f(t)$ satisfies

$$\frac{d}{dt}\tilde f(t) = -\widetilde{\mathbf{C}}\tilde f(t), \quad \tilde f(0) = f^I, \tag{2.4.2}$$

with

$$\widetilde{\mathbf{C}} := (\mathbf{C} - i\operatorname{Im}\lambda_1^{\mathbf{C}}\,\mathbf{I}) = (\mathbf{W}^*)^{-1}\begin{pmatrix}\operatorname{Re}\lambda_1^{\mathbf{C}} & 0\\ 0 & \operatorname{Re}\lambda_2^{\mathbf{C}}\end{pmatrix}\mathbf{W}^*.$$

The multiplication with $e^{i\operatorname{Im}\lambda_1^{\mathbf{C}} t}$ is another unitary transformation and does not change the norm, i.e. $\|f(t)\|_2 = \|\tilde f(t)\|_2$. Therefore, we can assume w.l.o.g. that matrix $\mathbf{C}$ has real coefficients <u>and</u> distinct real eigenvalues. Then, the solution $f(t)$ of the ODE (2.2.1) satisfies $\operatorname{Re} f(t) = f_{re}(t)$ and $\operatorname{Im} f(t) = f_{im}(t)$ where $f_{re}(t)$ and $f_{im}(t)$ are the solutions of the ODE (2.2.1) with initial data $\operatorname{Re} f^I$ and $\operatorname{Im} f^I$, resp. Altogether, we can assume w.l.o.g. that <u>all</u> quantities are real valued:
Considering a matrix $\mathbf{C} \in \mathbb{R}^{2\times 2}$ with two distinct real eigenvalues $\lambda_1 < \lambda_2$ and real eigenvectors v_1 and v_2, then the associated eigenspaces $\operatorname{span}\{v_1\}$ and $\operatorname{span}\{v_2\}$ dissect the plane into four sectors

$$\mathscr{S}^{\pm\mp} := \{z_1 v_1 + z_2 v_2 \mid z_1 \in \mathbb{R}^{\pm},\ z_2 \in \mathbb{R}^{\mp}\}, \tag{2.4.3}$$

see Fig. 2.1. A solution $f(t)$ of ODE (2.2.1) starting in an eigenspace will approach the origin in a straight line, such that

$$\|f(t)\|_2^2 = e^{-2\lambda_j^{\mathbf{C}} t}\|f^I\|_2^2 \qquad \forall t \geq 0. \tag{2.4.4}$$

If a solution starts instead in one of the four (open) sectors $\mathscr{S}^{\pm\mp}$, it will remain in that sector while approaching the origin. In fact, since $\lambda_1^{\mathbf{C}} < \lambda_2^{\mathbf{C}}$, if $f^I = z_1(v_1 + \gamma v_2)$ for some $z_1 \in \mathbb{R}\setminus\{0\}$ and $\gamma \in \mathbb{R}$, then the solution

$$f(t) = z_1\big(e^{-\lambda_1^{\mathbf{C}} t} v_1 + \gamma e^{-\lambda_2^{\mathbf{C}} t} v_2\big) = z_1 e^{-\lambda_1^{\mathbf{C}} t}\big(v_1 + \gamma e^{-(\lambda_2^{\mathbf{C}}-\lambda_1^{\mathbf{C}})t} v_2\big)$$

of the ODE (2.2.1) will remain in the sector

$$\mathscr{S}_\gamma^\pm := \{z_1(v_1 + z_2 v_2) \mid z_1 \in \mathbb{R}^\pm,\ z_2 \in [\min(0,\gamma), \max(0,\gamma)]\}, \tag{2.4.5}$$

see Fig. 2.1. For a fixed $f^I = z_1(v_1 + \gamma v_2)$, let $\mathscr{S}$ be the corresponding sector $\mathscr{S}_\gamma^\pm$. Then esti-

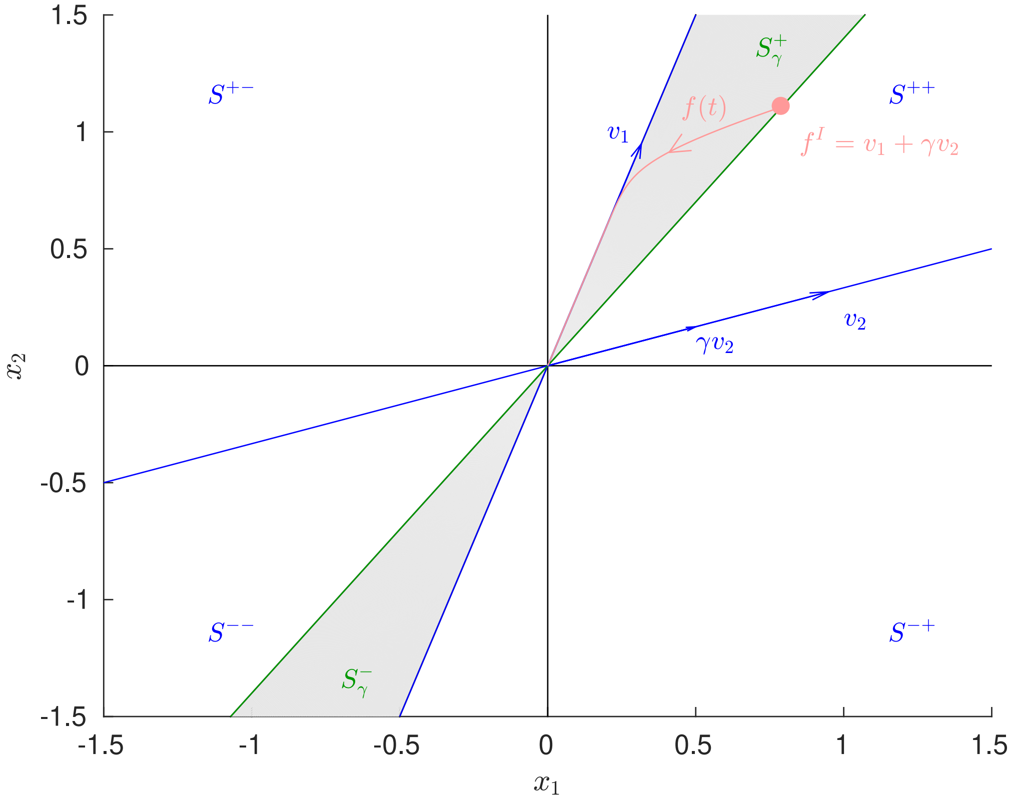

Figure 2.1: The blue (black) lines are the eigenspaces span$\{v_1\}$ and span$\{v_2\}$ of matrix $\mathbf{C}$. The red (grey) curve is a solution $f(t)$ of the ODE (2.2.1) with initial datum f^I. The *shaded regions* are the sectors S_γ^+, S_γ^- with the choice $\gamma = 1/2$. Note: The curves are colored only in the electronic version of this article.

mate (2.2.5) can be improved as follows

$$\|f(t)\|_2^2 \le \frac{1}{\lambda_{\min,\mathscr{S}}^{\mathbf{P}}} \|f(t)\|_{\mathbf{P}}^2 \le \frac{e^{-2\mu t}}{\lambda_{\min,\mathscr{S}}^{\mathbf{P}}} \|f^I\|_{\mathbf{P}}^2 \le c_{\mathscr{S}}(\mathbf{P})\, e^{-2\mu t} \|f^I\|_2^2, \quad t \ge 0, \tag{2.4.6}$$

where

$$\lambda_{\min,\mathscr{S}}^{\mathbf{P}} := \inf_{x \in \mathscr{S}} \frac{\langle x, \mathbf{P}x\rangle}{\langle x, x\rangle}, \quad \lambda_{init,\mathscr{S}}^{\mathbf{P}} := \frac{\langle f^I, \mathbf{P} f^I\rangle}{\langle f^I, f^I\rangle}, \quad c_{\mathscr{S}}(\mathbf{P}) := \frac{\lambda_{init,\mathscr{S}}^{\mathbf{P}}}{\lambda_{\min,\mathscr{S}}^{\mathbf{P}}}. \tag{2.4.7}$$

Note that, in the definition of $\lambda_{init,\mathscr{S}}^{\mathbf{P}}$ the sector $\mathscr{S} \in \left\{\mathscr{S}_\gamma^\pm \middle| \gamma \in \mathbb{R}\right\}$ also determines corresponding initial conditions $f^I \in \partial\mathscr{S}$ via $f^I = z_1(v_1 + \gamma v_2)$ (up to the constant $z_1 \ne 0$ which drops out in $\lambda_{init,\mathscr{S}}^{\mathbf{P}}$).

For (2.4.6) to hold for all trajectories and one fixed constant on the right hand side, we have to take the supremum over all initial conditions or, equivalently, over all sectors $\mathscr{S} \in \left\{\mathscr{S}_\gamma^\pm | \gamma \in \mathbb{R}\right\}$. Although $f^I = z_2 v_2$ is not included in any sector $\mathscr{S}_\gamma^+$, its corresponding multiplicative constant 1 (see (2.4.4)) is still covered. Then, the minimal multiplicative constant in (2.3.1) using (2.4.6) is

$$\widetilde{c} = \sqrt{\inf_{\mathbf{P}} \sup_{\mathscr{S}} c_{\mathscr{S}}(\mathbf{P})}, \tag{2.4.8}$$

where $\mathbf{P}$ ranges over all matrices of the form (2.2.3).

<u>Step 1</u> (computation of $\lambda^{\mathbf{P}}_{\min,\mathscr{S}_\gamma^+}$ for γ fixed): To find an explicit expression for this minimal constant c, we first determine $c_{\mathscr{S}}(\mathbf{P})$ for a given admissible matrix $\mathbf{P}$. As an example of sectors, we consider only $\mathscr{S}_\gamma^+$ for fixed $\gamma \le 0$ and compute

$$\begin{aligned}
\lambda^{\mathbf{P}}_{\min,\mathscr{S}_\gamma^+} &= \inf_{x\in\mathscr{S}_\gamma^+} \frac{\langle x, \mathbf{P}x\rangle}{\|x\|^2} = \inf_{z_1\in\mathbb{R}^+,\, z_2\in[\gamma,0]} \frac{\langle z_1(v_1+z_2v_2), \mathbf{P}(z_1(v_1+z_2v_2))\rangle}{\|z_1(v_1+z_2v_2)\|^2} \\
&= \inf_{z_2\in[\gamma,0]} \frac{\langle v_1+z_2v_2, \mathbf{P}(v_1+z_2v_2)\rangle}{\|v_1+z_2v_2\|^2}.
\end{aligned}$$

This also shows that $\lambda^{\mathbf{P}}_{\min,\mathscr{S}_\gamma^+} = \lambda^{\mathbf{P}}_{\min,\mathscr{S}_\gamma^-}$ for any fixed $\gamma\in\mathbb{R}$. Next, we use the result of Lemma 2.3.1 and (2.3.6), stating that the only admissible matrices are $\mathbf{P} = \mathbf{W}\operatorname{diag}(b_1,b_2)\mathbf{W}^*$ for $b_1, b_2 > 0$. Since $c_{\mathscr{S}}(b\mathbf{P}) = c_{\mathscr{S}}(\mathbf{P})$ for all $b>0$, we consider w.l.o.g. $b_1 = 1/b$ and $b_2 = b$ for $b>0$. Then, we deduce

$$\begin{aligned}
\lambda^{\mathbf{P}}_{\min,\mathscr{S}_\gamma^+} &= \inf_{z\in[\gamma,0]} \frac{\langle v_1+zv_2, \mathbf{P}(v_1+zv_2)\rangle}{\|v_1+zv_2\|^2} \\
&= \inf_{z\in[\gamma,0]} \frac{\langle \mathbf{W}^*(v_1+zv_2), \operatorname{diag}(1/b,b)\mathbf{W}^*(v_1+zv_2)\rangle}{\|v_1+zv_2\|^2}.
\end{aligned}$$

In our case of a real matrix $\mathbf{C}$ with distinct real eigenvalues, the left and right eigenvectors are related as follows: Up to a change of orientation, $\langle w_j, v_k\rangle = \delta_{jk}$ $(j,k\in\{1,2\})$. Considering $\langle w_j, v_j\rangle = 1$ for $j=1,2$, implies that the vectors w_j and v_j can be normalized simultaneously only if matrix $\mathbf{C}$ is symmetric. Therefore, using a coordinate system such that the normalized eigenvectors of $\mathbf{C}^*$ are given as (2.3.7) and $\mathbf{V} := (v_1|v_2) = (\mathbf{W}^*)^{-1}$ yields

$$v_1 = \frac{1}{\sqrt{1-\alpha^2}}\begin{pmatrix}\sqrt{1-\alpha^2}\\ -\alpha\end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{1-\alpha^2}}\begin{pmatrix}0\\1\end{pmatrix} \quad \text{for } \alpha \text{ in (2.3.7).}$$

Finally, we obtain

$$\lambda^{\mathbf{P}}_{\min,\mathscr{S}_\gamma^+} = \inf_{z\in[\gamma,0]} \frac{\langle \mathbf{W}^*(v_1+zv_2), \operatorname{diag}(1/b,b)\mathbf{W}^*(v_1+zv_2)\rangle}{\|v_1+zv_2\|^2} = \inf_{z\in[\gamma,0]} g(z)$$

and $\lambda^{\mathbf{P}}_{init,\mathscr{S}_\gamma^+} = g(\gamma)$ with

$$g(z) := \frac{(1-\alpha^2)\,(\frac{1}{b}+bz^2)}{1-2\alpha z+z^2}. \tag{2.4.9}$$

Step 2 (extrema of the function g): The function g has local extrema at

$$z_\pm = \frac{1}{2\alpha b}\Big(b - \frac{1}{b} \pm \sqrt{\big(b-\frac{1}{b}\big)^2 + 4\alpha^2}\Big)$$

which satisfy $z_- < 0 < z_+$. Writing $g'(z) = h_1(z)/h_2(z)$ with $h_1(z) := \big(-2\alpha b z^2 + 2\big(b-\frac{1}{b}\big)z + \frac{2}{b}\alpha\big)$ and $h_2(z) := (1-2\alpha z + z^2)^2/(1-\alpha^2) > 0$, we derive

$$g''(z_\pm) = \frac{h_1'(z_\pm)}{h_2(z_\pm)} = \mp 2\frac{1}{h_2(z_\pm)}\sqrt{\big(b-\frac{1}{b}\big)^2 + 4\alpha^2}\,.$$

In fact, the function g attains its global minimum on $\mathbb{R}$ (and on $\mathbb{R}_0^-$) at z_-, and its global maximum on $\mathbb{R}$ at z_+. The global supremum of $g(z)$ on $\mathbb{R}^-$ exists and satisfies

$$\sup_{z\in\mathbb{R}^-} g(z) = \begin{cases} g(0) = (1-\alpha^2)/b & \text{if } b \in (0,1)\,,\\ g(0) = \lim_{z\to-\infty} g(z) = 1-\alpha^2 & \text{if } b = 1\,,\\ \lim_{z\to-\infty} g(z) = (1-\alpha^2)b & \text{if } b\in(1,\infty)\,. \end{cases}$$

Step 3 (optimization of $c_{\mathscr{S}_\gamma^\pm}(\mathbf{P})$ w.r.t. γ): We obtain

$$c_{\mathscr{S}_\gamma^\pm}(\mathbf{P}(b)) = \frac{g(\gamma)}{\lambda^{\mathbf{P}(b)}_{\min,\mathscr{S}_\gamma^+}} = \begin{cases} 1 & \text{if } z_- \le \gamma < 0\,,\\ g(\gamma)/g(z_-) & \text{if } \gamma \le z_-\,. \end{cases}$$

Finally, we derive

$$\sup_{\gamma\in\mathbb{R}^-} c_{\mathscr{S}_\gamma^\pm}(\mathbf{P}(b)) = \lim_{\gamma\to-\infty}\frac{g(\gamma)}{g(z_-)} = \frac{(1-\alpha^2)b}{g(z_-)}\,, \tag{2.4.10}$$

and in a similar way,

$$\sup_{\gamma\in\mathbb{R}^+} c_{\mathscr{S}_\gamma^\pm}(\mathbf{P}(b)) = \frac{g(z_+)}{g(0)} = \frac{b g(z_+)}{1-\alpha^2}\,. \tag{2.4.11}$$

To finish this analysis we note that $c_{\mathscr{S}_0^\pm}(\mathbf{P}(b)) = 1$, due to (2.4.4) and $f^I = z_1 v_1$.
Step 4 (minimization of $\sup_{\mathscr{S}} c_{\mathscr{S}}(\mathbf{P})$ w.r.t. $\mathbf{P}$): We obtain

$$\inf_{\mathbf{P}}\sup_{\mathscr{S}} c_{\mathscr{S}}(\mathbf{P}) = \inf_{b\in(0,\infty)}\sup_{\gamma\in\mathbb{R}} c_{\mathscr{S}_\gamma^\pm}(\mathbf{P}(b)) = \inf_{b\in(0,\infty)}\max\Big\{\frac{(1-\alpha^2)b}{g(z_-)}, 1, \frac{bg(z_+)}{1-\alpha^2}\Big\}.$$

Taking into account the b-dependence of $z_\pm$, the functions $\frac{(1-\alpha^2)b}{g(z_-)}$ and $\frac{bg(z_+)}{1-\alpha^2}$ are monotone increasing in b, since

$$\frac{\partial}{\partial b}\frac{(1-\alpha^2)b}{g(z_-)} > 0\,, \qquad \frac{\partial}{\partial b}\frac{bg(z_+)}{1-\alpha^2} > 0\,.$$

Therefore we have to study their limits as $b\to 0$: We derive

$$\begin{aligned} \lim_{b\to0}\frac{(1-\alpha^2)b}{g(z_-)} &= 1 \qquad \text{using} \quad \lim_{b\to0} z_-(b) = -\infty\,,\\ \lim_{b\to0}\frac{bg(z_+)}{1-\alpha^2} &= \frac{1}{1-\alpha^2} > 1 \qquad \text{using} \quad \lim_{b\to0} z_+(b) = \alpha\,. \end{aligned} \tag{2.4.12}$$

Hence, $\inf_{b\in(0,\infty)}\sup_{\gamma\in\mathbb{R}} c_{\mathscr{S}^\pm_\gamma}(\mathbf{P}(b))$ is realized by the sector $\mathscr{S}^\pm_\gamma$ with $\gamma = z_+(b) > 0$ and in the limit $b \to 0$. Altogether we obtain

$$\widetilde{c} = \sqrt{\inf_{\mathbf{P}} \sup_{\mathscr{S}} c_{\mathscr{S}}(\mathbf{P})} = \frac{1}{\sqrt{1-\alpha^2}},$$

where the first equality holds since we discussed all solutions. This finishes the proof.
Step 5: Finally we have to verify that $\widetilde{c}$ is minimal in (2.3.1). We shall show that it is attained asymptotically (as $t\to\infty$) for a concrete trajectory: For fixed $b\in(0,\infty)$, the minimal multiplicative constant in (2.4.6) is attained for the solution with initial datum $f^I = v_1 + z_+(b)v_2 = y_+^{\mathbf{P}(b)}$, which is the eigenvector pertaining to the largest eigenvalue of $\mathbf{P}(b)$ (cp. to the proof of Theorem 2.3.7). The formula for f^I holds since $\sup_{\mathscr{S}} c_{\mathscr{S}}(\mathbf{P}(b)) = bg(z_+(b))/(1-\alpha^2)$. This can be verified by a direct comparison of (2.4.10) and (2.4.11). For b small it also follows from (2.4.12). In the limit $b\to 0$, $\mathbf{P}(b)$ in (2.3.9) approaches a multiple of $w_1\otimes w_1^*$ and

$$f^I = v_1 + z_+(b)v_2 \longrightarrow v_1 + \alpha v_2 = w_1 .$$

The solution $f(t)$ of the ODE (2.2.1) with $f^I = w_1$ satisfies

$$f(t) = e^{-\mathbf{C}t}w_1 = \mathbf{V}\begin{pmatrix} e^{-\lambda_1 t} & 0 \\ 0 & e^{-\lambda_2 t}\end{pmatrix}\mathbf{W}^* w_1 = e^{-\lambda_1 t}v_1 + \alpha e^{-\lambda_2 t}v_2 . \tag{2.4.13}$$

This implies

$$e^{\operatorname{Re}\lambda_1 t}\frac{\|f(t)\|_2}{\|f^I\|_2} \le \|v_1 + \alpha e^{-\operatorname{Re}(\lambda_2-\lambda_1)t}v_2\|_2 \overset{t\to\infty}{\longrightarrow} \|v_1\|_2 = \frac{1}{\sqrt{1-\alpha^2}}$$

and it finishes the proof. □ □

After the analysis in Theorems 2.3.7 and 2.4.1, we are left with the case of a matrix $\mathbf{C}\in\mathbb{C}^{2\times2}$ with eigenvalues λ_1 and λ_2 such that the real and imaginary parts are distinct. This case can not occur for real matrices $\mathbf{C}$. The proof of Lemma 2.3.3 gives an upper bound $\sqrt{\frac{1+\alpha}{1-\alpha}}$ for the multiplicative constant in (2.3.1). On the other hand, the solution $f(t)$ of the ODE (2.2.1) with $f^I = w_1$ satisfies (2.4.13), hence,

$$\begin{aligned}\|f(t)\|_2^2 &= e^{-2\operatorname{Re}\lambda_1 t}\|v_1 + \alpha e^{-(\lambda_2-\lambda_1)t}v_2\|_2^2 \\ &= \tfrac{1}{1-\alpha^2}e^{-2\operatorname{Re}\lambda_1 t}\Big(1 - 2\alpha^2 e^{-\operatorname{Re}(\lambda_2-\lambda_1)t}\cos\big(\operatorname{Im}(\lambda_2-\lambda_1)t\big) + \alpha^2 e^{-2\operatorname{Re}(\lambda_2-\lambda_1)t}\Big).\end{aligned}$$

The expression in the bracket is bigger than 1, e.g. at time $t = \pi/\operatorname{Im}(\lambda_2-\lambda_1)$. Thus the minimal multiplicative constant c is definitely bigger than $\frac{1}{\sqrt{1-\alpha^2}}$, which is the best constant for $\operatorname{Im}\lambda_1 = \operatorname{Im}\lambda_2$ (see Theorem 2.4.1).

Next, we derive the upper and lower envelopes for the norm of solutions $f(t)$ of ODE (2.2.1) in order to determine the sharp constant c. For a diagonalizable matrix $\mathbf{C}\in\mathbb{C}^{2\times2}$ with $\lambda_1^{\mathbf{C}} = \lambda_2^{\mathbf{C}}$ it holds that $\|f(t)\|_2 = e^{-\operatorname{Re}\lambda_1^{\mathbf{C}} t}\|f^I\|_2$. And for the general case we have:

Proposition 2.4.2. Let $\mathbf{C}\in\mathbb{C}^{2\times2}$ be a diagonalizable, positive stable matrix with eigenvalues $\lambda_1^{\mathbf{C}} \neq \lambda_2^{\mathbf{C}}$, and associated eigenvectors v_1 and v_2, resp. Then the norm of solutions $f(t)$ of ODE (2.2.1) satisfies

$$h_-(t)\|f^I\|_2^2 \le \|f(t)\|_2^2 \le h_+(t)\|f^I\|_2^2, \qquad \forall t\ge 0,$$

where the envelopes $h_\pm(t)$ are given by

$$h_\pm(t) := e^{-2\operatorname{Re}\lambda_1^{\mathbf{C}} t} m_\pm(t)$$

with

$$m_\pm(t) := \pm e^{-\gamma t}\Big(\sqrt{\frac{\big(\cosh(\gamma t)-\alpha^2\cos(\delta t)\big)^2}{(1-\alpha^2)^2}-1} \pm \frac{\big(\cosh(\gamma t)-\alpha^2\cos(\delta t)\big)}{1-\alpha^2}\Big),$$

where $\gamma := \operatorname{Re}(\lambda_2^{\mathbf{C}}-\lambda_1^{\mathbf{C}})$, $\delta := \operatorname{Im}(\lambda_2^{\mathbf{C}}-\lambda_1^{\mathbf{C}})$, $\alpha := \left|\left\langle \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}\right\rangle\right|$ and $\alpha \in [0,1)$.

While the rest of the article is based on estimating Lyapunov functionals, the following proof will use the explicit solution formula of the ODE.

Proof. We use again the unitary transformation as in the proof of Lemma 2.3.3, such that the eigenvectors w_1 and w_2 of $\mathbf{C}^*$ are given in (2.3.7). If $f(t)$ is a solution of (2.2.1), then $\tilde f(t) = e^{\lambda_1^{\mathbf{C}} t} f(t)$ satisfies

$$\frac{d}{dt}\tilde f(t) = -\widetilde{\mathbf{C}}\tilde f(t), \quad \tilde f(0) = f^I, \tag{2.4.14}$$

with

$$\widetilde{\mathbf{C}} = (\mathbf{C}-\lambda_1^{\mathbf{C}}\,\mathbf{I}) = (\mathbf{W}^*)^{-1}\begin{pmatrix}0 & 0\\ 0 & \lambda_2^{\mathbf{C}}-\lambda_1^{\mathbf{C}}\end{pmatrix}\mathbf{W}^*.$$

The explicit solution $\tilde f(t)$ of (2.4.14) is

$$\tilde f(t) = (\mathbf{W}^*)^{-1}\begin{pmatrix}1 & 0\\ 0 & e^{-(\gamma+i\delta)t}\end{pmatrix}\mathbf{W}^* f^I = \begin{pmatrix} f_1^I \\ \frac{\alpha}{\sqrt{1-\alpha^2}}(e^{-(\gamma+i\delta)t}-1)f_1^I + e^{-(\gamma+i\delta)t} f_2^I\end{pmatrix},$$

where $\gamma = \operatorname{Re}(\lambda_2^{\mathbf{C}}-\lambda_1^{\mathbf{C}})$ and $\delta = \operatorname{Im}(\lambda_2^{\mathbf{C}}-\lambda_1^{\mathbf{C}})$. If the initial data f^I lies in $\mathbb{R}\times\mathbb{C}$ then the solution will satisfy $\tilde f(t) \in \mathbb{R}\times\mathbb{C}$ for all $t \ge 0$. The multiplication with $\overline{f_1^I}/|f_1^I|$ is another unitary transformation and does not change the norm. Therefore, to compute the envelope for the norm of solutions $\tilde f(t)$ of ODE (2.4.14) we assume w.l.o.g. that

$$f_{\phi,\theta}^I = \begin{pmatrix}\cos(\phi)\\ \sin(\phi)e^{i\theta}\end{pmatrix} \in \mathbb{R}\times\mathbb{C}, \qquad \text{where } \phi,\theta \in [0,2\pi), \tag{2.4.15}$$

such that $\|f_{\phi,\theta}^I\| = 1$. We consider the solution $\tilde f_{\phi,\theta}(t)$ for (2.4.14) with $f^I = f_{\phi,\theta}^I$. To compute the envelopes (for fixed t), we solve $\partial_\phi\|\tilde f_{\phi,\theta}\|^2 = 0$ and $\partial_\theta\|\tilde f_{\phi,\theta}\|^2 = 0$ in terms of ϕ and θ. Evaluating $\|\tilde f_{\phi,\theta}(t)\|^2$ at $\phi = \phi(t)$ and $\theta = \theta(t)$ yields the envelopes for the norm of solutions $\tilde f(t)$ of ODE (2.4.14). Consequently, we derive the envelopes $h_\pm(t)\|f^I\|^2$ for the original problem, since $\|f(t)\|_2 = e^{-\operatorname{Re}\lambda_1^{\mathbf{C}} t}\|\tilde f(t)\|_2$. □ □

Corollary 2.4.3. Let $\mathbf{C} \in \mathbb{C}^{2\times 2}$ be a diagonalizable, positive stable matrix. Then the minimal multiplicative constant c in (2.3.1) for the ODE (2.2.1) is given by

$$c = \sqrt{\sup_{t\ge 0} m_+(t)}, \tag{2.4.16}$$

where $m_+(t)$ is the function given in Proposition 2.4.2.

In general we could not find an explicit formula for $\sup_{t\ge0} m_+(t)$.

2.5 A Family of Decay Estimates for Hypocoercive ODEs

In this section we shall illustrate the interdependence of maximizing the decay rate λ and minimizing the multiplicative constant c in estimates like (2.3.1). For the ODE-system (2.2.1), the procedure described in Remark 2.1.2(b) yields the optimal bound for large time, with the sharp decay rate $\mu := \min\{\mathrm{Re}(\lambda)|\lambda \text{ is an eigenvalue of } \mathbf{C}\}$. But for non-coercive $\mathbf{C}$ we must have $c > 1$. Hence, such a bound cannot be sharp for short time. As a counterexample we consider the simple energy estimate (obtained by premultiplying (2.2.1) with f^*)

$$\|f(t)\|_2 \le e^{-\mu_s t}\|f^I\|_2\,, \quad t \ge 0\,,$$

with $\mathbf{C}_s := \frac{1}{2}(\mathbf{C}+\mathbf{C}^*)$ and $\mu_s := \min\{\lambda|\lambda \text{ is an eigenvalue of } \mathbf{C}_s\}$.

The goal of this section is to derive decay estimates for (2.2.1) with rates in between this weakest rate μ_s and the optimal rate μ from (2.2.5). It holds that $\mu_s \le \mu$. At the same time we shall also present *lower bounds* on $\|f(t)\|_2$. The energy method again provides the simplest example of it, in the form

$$\|f(t)\|_2 \ge e^{-\nu_s t}\|f^I\|_2\,, \quad t \ge 0\,,$$

with $\nu_s := \max\{\lambda|\lambda \text{ is an eigenvalue of } \mathbf{C}_s\}$. Clearly, estimates with decay rates outside of $[\mu_s, \nu_s]$ are irrelevant.

We present our main result only for the two-dimensional case, as the best multiplicative constant is not yet known explicitly in higher dimensions (cf. §2.3):

Proposition 2.5.1. Let $\mathbf{C} \in \mathbb{C}^{2\times 2}$ be a diagonalizable positive stable matrix with spectral gap $\mu := \min\{\mathrm{Re}(\lambda_j^{\mathbf{C}})|\, j = 1,2\}$. Then, all solutions to (2.2.1) satisfy the following upper and lower bounds:

a)
$$\|f(t)\|_2 \le c_1(\tilde\mu)\, e^{-\tilde\mu t}\|f^I\|_2\,, \quad t \ge 0\,, \quad \mu_s \le \tilde\mu \le \mu\,, \tag{2.5.1}$$
with
$$c_1^2(\tilde\mu) = \kappa_{\min}(\beta(\tilde\mu))$$
given explicitly in (2.5.8) below. There, $\alpha \in [0,1)$ is the cos of the (minimal) angle of the eigenvectors of $\mathbf{C}^*$ (cf. the proof of Lemma 2.3.3), and $\beta(\tilde\mu) = \max(-\alpha, -\beta_0)$, with β_0 defined in (2.5.6), (2.5.7) below.

b)
$$\|f(t)\|_2 \ge c_2(\tilde\mu)\, e^{-\tilde\mu t}\|f^I\|_2\,, \quad t \ge 0\,, \quad \nu \le \tilde\mu \le \nu_s\,, \tag{2.5.2}$$
with $\nu := \max\{\mathrm{Re}(\lambda_j^{\mathbf{C}})|\, j = 1,2\}$. The maximal constant
$$c_2^2(\tilde\mu) = \kappa_{\min}(\beta(\tilde\mu))^{-1}$$
is given again by (2.5.8), with α, $\beta(\tilde\mu)$ defined as in Part (a).

Proof. Part (a): For a fixed $\tilde\mu \in [\mu_s, \mu]$ we have to determine the smallest constant c_1 for the estimate (2.5.1), following the strategy of proof from §2.3. To this end, we use a unitary transformation of the coordinate system and write $\mathbf{P}(\tilde\mu) = \mathbf{W}\mathbf{D}_u\mathbf{W}^*$ with

$$\mathbf{W} = \begin{pmatrix} 1 & \alpha \\ 0 & \sqrt{1-\alpha^2} \end{pmatrix}, \quad \mathbf{D}_u = \begin{pmatrix} 1/b & \beta(\tilde\mu) \\ \bar\beta(\tilde\mu) & b \end{pmatrix}, \tag{2.5.3}$$

where we set w.l.o.g. $b_1 = 1/b$, $b_2 = b$ with $b > 0$. Moreover, $|\beta|^2 < 1$ has to hold. Now, we have to find the positive definite Hermitian matrix $\mathbf{D}_u$, such that the analog of (2.3.3), (2.3.4) holds, i.e.:

$$\mathring{A} := \begin{pmatrix} 2(\mathrm{Re}(\lambda_1^{\mathbf{C}}) - \tilde{\mu})/b & (\bar{\lambda}_1^{\mathbf{C}} + \lambda_2^{\mathbf{C}} - 2\tilde{\mu})\beta \\ (\lambda_1^{\mathbf{C}} + \bar{\lambda}_2^{\mathbf{C}} - 2\tilde{\mu})\bar{\beta} & 2(\mathrm{Re}(\lambda_2^{\mathbf{C}}) - \tilde{\mu})b \end{pmatrix} \geq 0\,, \tag{2.5.4}$$

As in the proof of Lemma 2.3.1, we assume that the eigenvalues of $\mathbf{C}$ are ordered as $\mathrm{Re}(\lambda_2^{\mathbf{C}}) \geq \mathrm{Re}(\lambda_1^{\mathbf{C}}) = \mu \geq \tilde{\mu}$. Hence, $\mathrm{Tr}\mathring{A} \geq 0$. For the non-negativity of the determinant to hold, i.e.

$$\det \mathring{A} = 4\big(\mathrm{Re}(\lambda_1^{\mathbf{C}}) - \tilde{\mu}\big)\big(\mathrm{Re}(\lambda_2^{\mathbf{C}}) - \tilde{\mu}\big) - |\lambda_1^{\mathbf{C}} + \bar{\lambda}_2^{\mathbf{C}} - 2\tilde{\mu}|^2 |\beta|^2 \geq 0\,, \tag{2.5.5}$$

we have the following restriction on β:

$$|\beta|^2 \leq \beta_0^2 := \frac{4\big(\mathrm{Re}(\lambda_1^{\mathbf{C}}) - \tilde{\mu}\big)\big(\mathrm{Re}(\lambda_2^{\mathbf{C}}) - \tilde{\mu}\big)}{|\lambda_1^{\mathbf{C}} + \bar{\lambda}_2^{\mathbf{C}} - 2\tilde{\mu}|^2}\,. \tag{2.5.6}$$

If $\lambda_1^{\mathbf{C}} + \bar{\lambda}_2^{\mathbf{C}} - 2\tilde{\mu} = 0$, we conclude $\lambda_1^{\mathbf{C}} = \lambda_2^{\mathbf{C}}$ and that we have chosen the sharp decay rate $\tilde{\mu} = \mu$. As the associated, minimal condition number $\kappa(\mathbf{P})$ was already determined in Lemma 2.3.3, we shall not rediscuss this case here. But to include this case into the statement of the theorem, we set

$$\beta_0 := 1\,, \quad \text{if } \lambda_1^{\mathbf{C}} = \lambda_2^{\mathbf{C}} \text{ and } \tilde{\mu} = \mu\,. \tag{2.5.7}$$

From (2.5.6) we conclude that $\beta_0 \in [0,1]$. Note that $\beta_0 = 1$ is only possible for $\tilde{\mu} = \mu$ and $\lambda_1^{\mathbf{C}} = \lambda_2^{\mathbf{C}}$, i.e. the case that we just sorted out. For the rest of the proof we hence assume that condition (2.5.6) holds with $\beta_0 \in [0,1)$.

For admissible matrices $\mathbf{D}_u$ (i.e. with $b > 0$ and $|\beta| \leq \beta_0$) it remains to determine the matrix

$$\mathbf{P}(b,\beta) = \mathbf{W}\mathbf{D}_u\mathbf{W}^* = \begin{pmatrix} \frac{1}{b} + 2\alpha\,\mathrm{Re}\,\beta + b\alpha^2 & (\beta + b\alpha)\sqrt{1-\alpha^2} \\ (\bar{\beta} + b\alpha)\sqrt{1-\alpha^2} & b(1-\alpha^2) \end{pmatrix},$$

(with $\mathbf{W}$ and $\mathbf{D}_u$ given in (2.5.3)), having the minimal condition number $\kappa\big(\mathbf{P}(b,\beta)\big) = \lambda_+^{\mathbf{P}}(b,\beta)/\lambda_-^{\mathbf{P}}(b,\beta)$. Here

$$\lambda_\pm^{\mathbf{P}}(b,\beta) = \frac{\mathrm{Tr}\,\mathbf{P}(b,\beta) \pm \sqrt{(\mathrm{Tr}\,\mathbf{P}(b,\beta))^2 - 4\det\mathbf{P}(b,\beta)}}{2}$$

are the (positive) eigenvalues of $\mathbf{P}(b,\beta)$.

As a first step we shall minimize $\kappa\big(\mathbf{P}(b,\beta)\big)$ w.r.t. b (and for β fixed), since $\mathrm{argmin}_{b>0}\,\kappa\big(\mathbf{P}(b,\beta)\big)$ will turn out to be independent of β. We notice that $\mathrm{Tr}\,\mathbf{P}(b,\beta) = b + 2\alpha\,\mathrm{Re}\,\beta + 1/b$ is a convex function of $b \in (0,\infty)$ which attains its minimum for $b = 1$. Moreover, $\det\mathbf{P}(b,\beta) = (1-\alpha^2)(1-|\beta|^2) > 0$ is independent of b. This yields the condition number

$$\kappa_{\min}(\beta) = \frac{\lambda_+^{\mathbf{P}}(1,\beta)}{\lambda_-^{\mathbf{P}}(1,\beta)} = \frac{1 + \sqrt{1 - \frac{(1-\alpha^2)(1-|\beta|^2)}{(1+\alpha\,\mathrm{Re}\,\beta)^2}}}{1 - \sqrt{1 - \frac{(1-\alpha^2)(1-|\beta|^2)}{(1+\alpha\,\mathrm{Re}\,\beta)^2}}}\,.$$

As a second step we minimize $\kappa_{\min}(\beta)$ on the disk $|\beta| \leq \beta_0$. To this end, the quotient $\frac{(1-\alpha^2)(1-|\beta|^2)}{(1+\alpha\,\mathrm{Re}\,\beta)^2}$ should be as large as possible. For any fixed $|\beta| \leq \beta_0$, this happens by choosing $\beta = -|\beta|$, since

$\alpha \in [0,1)$. Hence it remains to maximize the function $g(\beta) := \frac{1-\beta^2}{(1+\alpha\beta)^2}$ on the interval $[-\beta_0, 0]$. It is elementary to verify that g is maximal at $\tilde{\beta} := \max(-\alpha, -\beta_0)$. Then, the minimal condition number is

$$\kappa_{\min}(\tilde{\beta}) = \kappa\big(\mathbf{P}(1,\tilde{\beta})\big) = \frac{1+\sqrt{1-\frac{(1-\alpha^2)(1-\tilde{\beta}^2)}{(1+\alpha\tilde{\beta})^2}}}{1-\sqrt{1-\frac{(1-\alpha^2)(1-\tilde{\beta}^2)}{(1+\alpha\tilde{\beta})^2}}}. \tag{2.5.8}$$

Part (b): Since the proof of the lower bound is very similar to Part (a), we shall just sketch it. For a fixed $\tilde{\mu} \in [\nu, \nu_s]$ we have to determine the largest constant c_2 for the estimate (2.5.2). To this end we need to satisfy the inequality

$$\mathbf{C}^*\mathbf{P} + \mathbf{P}\mathbf{C} \le 2\tilde{\mu}\mathbf{P}$$

with a positive definite Hermitian matrix $\mathbf{P}$ with minimal condition number $\kappa(\mathbf{P})$. In analogy to §2.2 this would imply

$$\frac{d}{dt}\|f(t)\|_{\mathbf{P}}^2 \ge -2\tilde{\mu}\|f(t)\|_{\mathbf{P}}^2,$$

and hence the desired lower bound

$$\|f(t)\|_2^2 \ge (\lambda_{\max}^{\mathbf{P}})^{-1}\|f(t)\|_{\mathbf{P}}^2 \ge (\lambda_{\max}^{\mathbf{P}})^{-1} e^{-2\tilde{\mu}t}\|f^I\|_{\mathbf{P}}^2 \ge (\kappa(\mathbf{P}))^{-1} e^{-2\tilde{\mu}t}\|f^I\|_2^2.$$

For minimizing $\kappa(\mathbf{P})$, we again use a unitary transformation of the coordinate system and write $\mathbf{P}$ as $\mathbf{P}(\tilde{\mu}) = \mathbf{W}\mathbf{D}_l\mathbf{W}^*$, with $\mathbf{W}$ from (2.5.3) and the positive definite Hermitian matrix

$$\mathbf{D}_l = \begin{pmatrix} 1/b & \beta(\tilde{\mu}) \\ \bar{\beta}(\tilde{\mu}) & b \end{pmatrix},$$

with $b > 0$ and $|\beta|^2 < 1$. Then, the matrix Å from (2.5.4) has to satisfy Å ≤ 0. Since we chose the eigenvalues of $\mathbf{C}$ to be ordered as $\mathrm{Re}(\lambda_1^{\mathbf{C}}) \le \mathrm{Re}(\lambda_2^{\mathbf{C}}) = \nu \le \tilde{\mu}$, we have TrÅ ≤ 0. The necessary non-negativity of its determinant again reads as (2.5.5).

In the special case $\lambda_1^{\mathbf{C}} + \bar{\lambda}_2^{\mathbf{C}} - 2\tilde{\mu} = 0$, we conclude again $\lambda_1^{\mathbf{C}} = \lambda_2^{\mathbf{C}}$ and $\tilde{\mu} = \nu$. Hence Å $= 0$. Since β is then only restricted by $|\beta| < 1$, we can again set $\beta_0 = 1$ and obtain the minimal $\kappa(\mathbf{P})$ for $\tilde{\beta}(\nu) = -\alpha$, as in Part (a).

In the generic case, the minimal $\kappa(\mathbf{P})$ is obtained for $\tilde{\beta} = \max(-\alpha, -\beta_0)$ with β_0 given in (2.5.6). Hence, the maximal constant in the lower bound (2.5.2) is $c_2^2(\tilde{\mu}) = \kappa_{\min}(\tilde{\beta})^{-1}$ where $\kappa_{\min}$ is given by (2.5.8). This finishes the proof. □ □

We illustrate the results of Proposition 2.5.1 with two examples.

Example 2.5.2. We consider ODE (2.2.1) with the matrix

$$\mathbf{C} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$$

which has eigenvalues $\lambda_\pm = (1 \pm i\sqrt{3})/2$, and some normalized eigenvectors of $\mathbf{C}^*$ are, e.g.

$$w_+ = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ \lambda_- \end{pmatrix}, \qquad w_- = \frac{1}{\sqrt{2}}\begin{pmatrix} -\lambda_- \\ 1 \end{pmatrix}. \tag{2.5.9}$$

The optimal decay rate is $\mu = 1/2$, whereas the minimal and maximal eigenvalues of $\mathbf{C}_s$ are $\mu_s = 0$ and $\nu_s = 1$, respectively. To bring the eigenvectors of $\mathbf{C}^*$ in the canonical form used in the proof of Proposition 2.5.1, we fix the eigenvector w_+, and choose the unitary multiplicative factor for the second eigenvector w_- as in (2.5.9) such that $\langle w_+, w_- \rangle$ is a real number. Finally, we use the Gram-Schmidt process to obtain a new orthonormal basis such that the eigenvectors of $\mathbf{C}^*$ in the new orthonormal basis are of the form (2.3.7) with $\alpha = 1/2$. Then, the upper and lower bounds for the Euclidean norm of a solution of (2.2.1) are plotted in Fig. 2.2 and Fig. 2.3. For

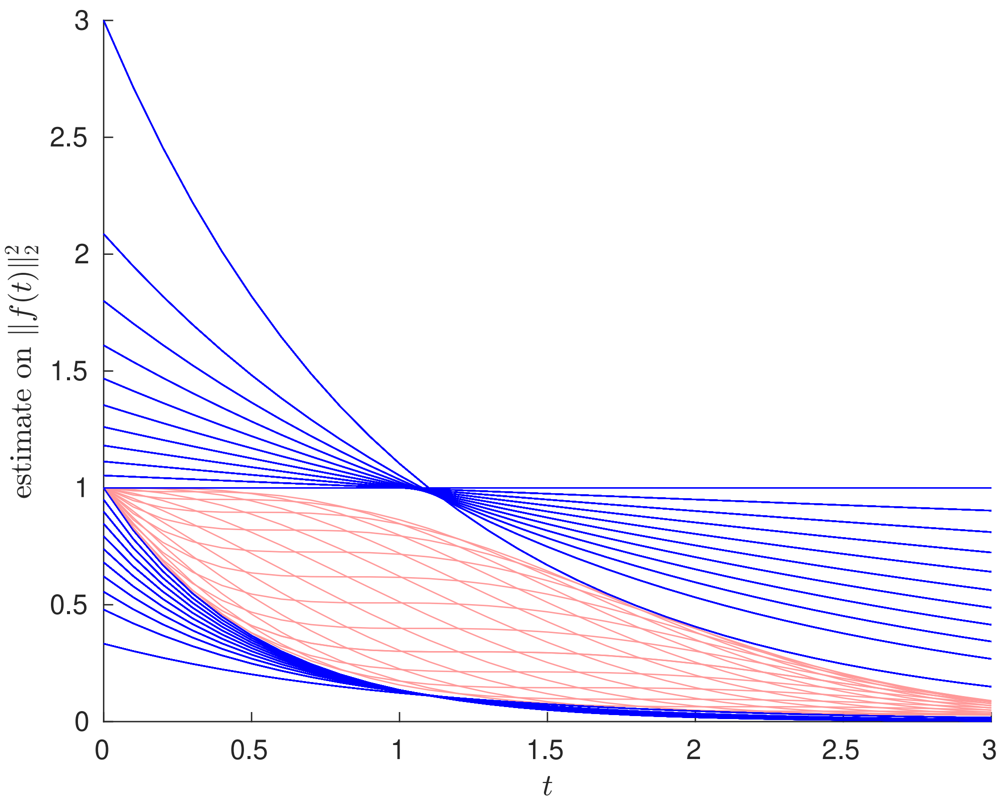

Figure 2.2: The red (grey) curves are the squared norm of solutions $f(t)$ for ODE (2.2.1) with matrix $\mathbf{C} = [1, -1; 1, 0]$ and various initial data f^I with norm 1. The blue (black) curves are the lower and upper bounds for the squared norm of solutions. Note: The curves are colored only in the electronic version of this article.

both the upper and lower bounds, the respective family of decay curves does *not* intersect in a single point (see Fig. 2.3). Hence, the whole family of estimates provides a (slightly) better estimate on $\|f(t)\|_2$ than if just considering the two extremal decay rates. For the upper bound this means

$$\|f(t)\|_2 \le \min_{\tilde{\mu} \in [\mu_s, \mu]} c_1(\tilde{\mu})\, e^{-\tilde{\mu} t}\, \|f^I\|_2 \le \min\{1, c_1(\mu)\, e^{-\mu t}\}\, \|f^I\|_2, \quad t \ge 0,$$

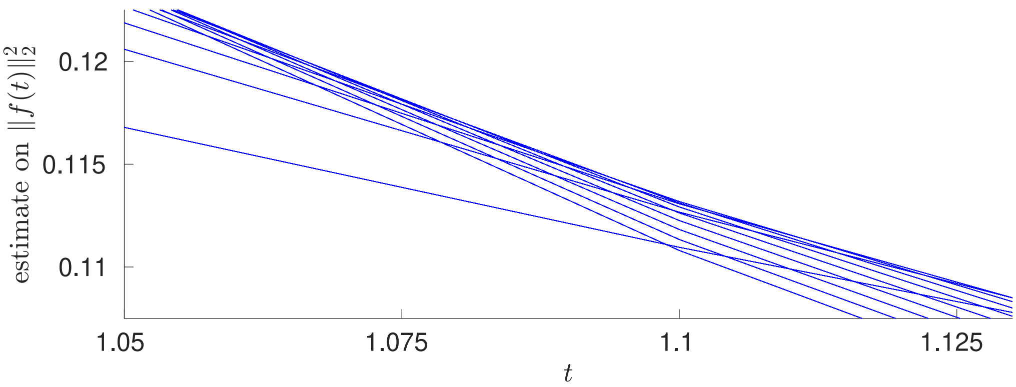

Figure 2.3: Zoom of Fig. 2.2: The curves are the lower bounds for the squared norm of solutions for ODE (2.2.1) with matrix $\mathbf{C} = [1, -1; 1, 0]$ and various initial data f^I with norm 1. This plot shows that these lower bounds do not intersect in a single point.

and for the lower bound

$$\|f(t)\|_2 \geq \max_{\tilde{\nu}\in[\nu,\nu_s]} c_2(\tilde{\nu})\, e^{-\tilde{\nu}t}\, \|f^I\|_2 \geq \max\{c_2(\nu)\, e^{-\nu t}, c_2(\nu_s)\, e^{-\nu_s t}\}\, \|f^I\|_2\,, \quad t \geq 0\,.$$

Note that the upper bound $\sqrt{3}e^{-t/2}$ with the sharp decay rate $\mu = \frac{1}{2}$ carries the optimal multiplicative constant $c = \sqrt{3}$, as it touches the set of solutions (see Fig. 2.2). But this is not true for the estimates with smaller decay rates (except of $\tilde{\mu} = 0$). The reason for this lack of sharpness is the fact that the inequality $\|f(t)\|_{\mathbf{P}}^2 \leq e^{-2\tilde{\mu}t}\|f^I\|_{\mathbf{P}}^2$ used in the proof of Proposition 2.5.1 is, in general, not an equality (in contrast to (2.2.6)). □

In the next example we consider a matrix $\mathbf{C} \in \mathbb{R}^{2\times 2}$ with $\operatorname{Re}\lambda_1 \neq \operatorname{Re}\lambda_2$, which corresponds to the case analysed in Theorem 2.4.1. For such cases the strategy of Proposition 2.5.1 (based on minimizing $\kappa(\mathbf{P})$) could be improved in the spirit of Theorem 2.4.1, but we shall not carry this out here. Hence, the estimates of the following example will not be sharp, see Fig. 2.4.

Example 2.5.3. We consider ODE (2.2.1) with the matrix

$$\mathbf{C} = \begin{pmatrix} 19/20 & -3/10 \\ 3/10 & -1/20 \end{pmatrix}$$

which has the eigenvalues $\lambda_1 = 1/20$ and $\lambda_2 = 17/20$, and some normalized eigenvectors of $\mathbf{C}^*$ are, e.g.

$$w_1 = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ -3 \end{pmatrix}, \qquad w_2 = \frac{1}{\sqrt{10}}\begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

The optimal decay rate is $\mu = 1/20$, whereas the minimal and maximal eigenvalues of $\mathbf{C}_s$ are $\mu_s = -1/20$ and $\nu_s = 19/20$, respectively. Since the matrix $\mathbf{C}$ and its eigenvalues are real valued, the eigenvectors of $\mathbf{C}^*$ are already in the canonical form used in the Gram-Schmidt process to obtain a new orthogonal basis such that the eigenvectors of $\mathbf{C}^*$ in the new basis are of the form (2.3.7) with $\alpha = 3/5$. Then, the upper and lower bounds for the Euclidean norm of a solution of (2.2.1)

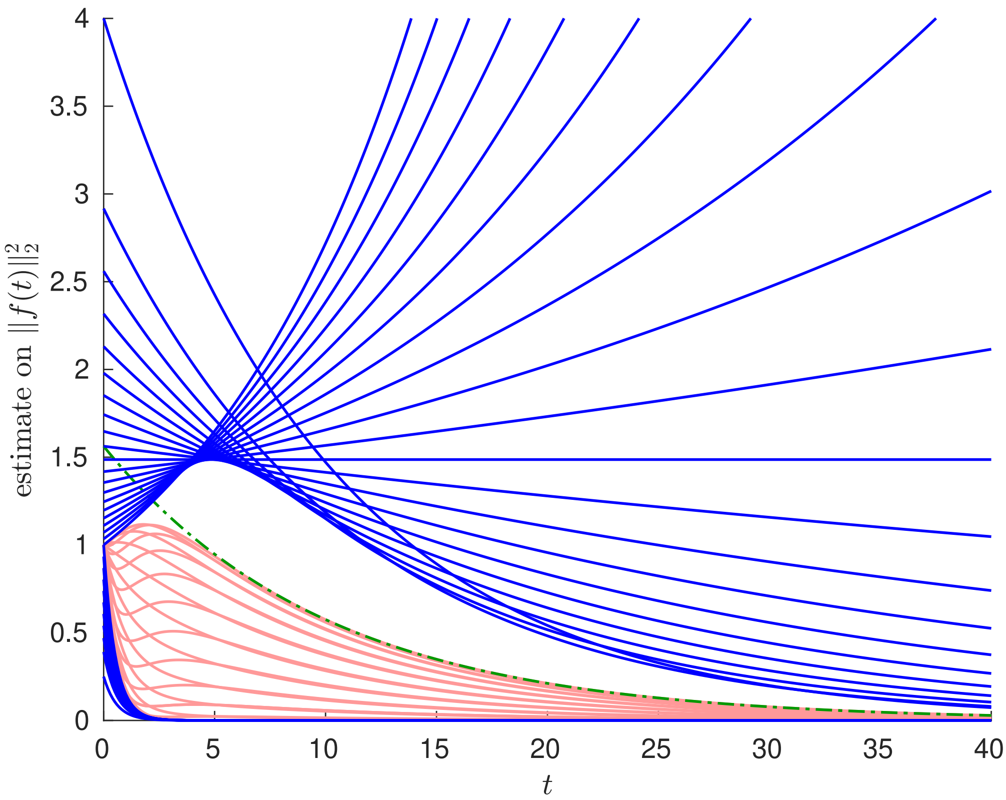

Figure 2.4: The red (grey) curves are the squared norm of solutions $f(t)$ for ODE (2.2.1) with matrix $\mathbf{C} = [19/20, -3/10; 3/10, -1/20]$ and various initial data f^I with norm 1. The blue (black) curves are the lower and upper bounds for the squared norm of solutions derived from Proposition 2.5.1. The green (black) *dash-dotted curve* is the upper bound for the squared norm of solutions derived from Theorem 2.4.1. Note: The curves are colored only in the electronic version of this article.

are plotted in Fig. 2.4. Since $\mu_s < 0$, solutions $f(t)$ to this example may initially increase in norm. □

Acknowledgments. All authors were supported by the FWF-funded SFB #F65. The second author was partially supported by the FWF-doctoral school W1245 “Dissipation and dispersion in nonlinear partial differential equations”. We are grateful to the anonymous referee who led us to better distinguish the different cases studied in §3 and §4.

Bibliography

[1] Achleitner, F., Arnold, A., Carlen, E.A.: On linear hypocoercive BGK models. Gonçalves P., Soares A. (eds) From Particle Systems to Partial Differential Equations III. Springer Proc. Math. Stat., vol 162, pp. 1–37. Springer, Cham (2016)

[2] Achleitner, F., Arnold, A., Carlen, E.A.: On multi-dimensional hypocoercive BGK models. Kinet. Relat. Models **11**, 953–1009 (2018)

[3] Achleitner, F., Arnold, A., Stürzer, D.: Large-Time behaviour in Non-Symmetric Fokker-Planck Equations. Riv. Math. Univ. Parma (N.S.) **6**, 1–68 (2015)

[4] Arnold, A., Erb, J.: Sharp entropy decay for hypocoercive and non-symmetric Fokker-Planck equations with linear drift. arXiv preprint, arXiv:1409.5425 (2014)

[5] Arnold, A., Jin, S., Wöhrer, T.: Sharp Decay Estimates in Defective Evolution Equations: from ODEs to Kinetic BGK Equations. preprint, (2018)

[6] Arnold, V.I.: Ordinary differential equations. MIT Press, Cambridge, Mass.-London (1978)

[7] Bhatnagar, P.L., Gross, E.P., Krook, M.: A Model for Collision Processes in Gases. I. Small Amplitude Processes in Charged and Neutral One-Component Systems. Phys. Rev. **94**, 511–525 (1954)

[8] Boyd, S.P., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)

[9] Braatz, R.D., Morari, M.: Minimizing the Euclidean condition number. SIAM J. Control Optim. **32**, 1763–1768 (1994)

[10] Businger, P.A.: Matrices which can be optimally scaled. Numer. Math. **12**, 346–348 (1968)

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

[12] Kolotilina, L.Yu.: Solution of the problem of optimal diagonal scaling for quasireal Hermitian positive definite 3×3 matrices. J. Math. Sci. (N.Y.) **132**, 190–213 (2006)

[13] Miclo, L., Monmarché, P.: Étude spectrale minutieuse de processus moins indécis que les autres, In: Donati-Martin C., Lejay A., Rouault A. (eds) Séminaire de Probabilités XLV, Lecture Notes in Mathematics, vol 2078, pp. 459–481. Springer, Heidelberg (2013). English summary available at https://www.ljll.math.upmc.fr/~monmarche

[14] Sezginer, R.S., Overton, M.L.: The largest singular value of $e^X A_0 e^{-X}$ is convex on convex sets of commuting matrices, IEEE Trans. Automat. Control **35**, 229–230 (1990)

[15] Villani, C.: Hypocoercivity. Mem. Amer. Math. Soc., **202** (2009)

3 Propagator Norm and Sharp Decay Estimates for Fokker-Planck Equations with Linear Drift

3.1 Introduction

We are going to study the large-time and short-time behaviour of the solution of Fokker-Planck (FP) equations with linear drift and possibly degenerate diffusion for $g = g(t, y)$:

$$\partial_t g = -\widetilde{L}g := \operatorname{div}_y(\widetilde{D}\nabla_y g + \widetilde{C}yg), \qquad y \in \mathbb{R}^d,\ t \in (0,\infty), \tag{3.1.1}$$

$$g(t=0) = g_0 \in L^1_+(\mathbb{R}^d), \tag{3.1.2}$$

$$\int_{\mathbb{R}^d} g_0(y)dy = 1. \tag{3.1.3}$$

We assume that

- $\widetilde{D} \in \mathbb{R}^{d\times d}$ is non-zero, positive semi-definite, symmetric, and constant in y,
- $\widetilde{C} \in \mathbb{R}^{d\times d}$ is positive stable, (typically non-symmetric,) and constant in y.

The goal of this study is to investigate the qualitative and quantitative large time behaviour of the solution of (3.1.1). Several authors (see, e.g., [6], [7], [27], [5]) have addressed the following questions: Under which conditions is there a non trivial steady state g_∞? In the affirmative case, does the solution $g(t)$ converge to the steady state for $t \to \infty$ in a suitable norm? Is the convergence exponential?

In particular, the large-time behaviour of FP-equations has been treated in [34] via spectral methods. Instead, entropy methods are used in [7]. From these previous studies it is well known that (under some assumptions that will be defined in the next section) the solution $g(t)$ converges to the steady state g_∞ with an exponential decay rate, up to a multiplicative constant greater than one. In the degenerate case, where the diffusion matrix $\widetilde{D}$ is non-invertible, this property of the solution is known as *hypocoercivity*, as introduced in [36].

Optimal exponential decay estimates for the convergence of the solution to the steady state in both the degenerate and the non-degenerate cases have been shown in [6]. Special care is required when the eigenvalues of $\widetilde{C}$ with smallest real part are *defective*. This situation is covered in [5] and [25]. In both cases, the sharpness of the estimate refers only to the exponential decay rate of the convergence of the solution. The issue of finding the best multiplicative constant in the decay estimate for FP-equations (3.1.1) is still open. This is one of the topics of this paper. Even for linear ODEs there are only partial results on this best constant, as for example in [24] and [3]. In particular, [3] gives the explicit best multiplicative constant in the two-dimensional case for $\dot{x} = -Cx$, where C is a positive stable matrix. A very complete solution has been derived in [17] for a special case, the kinetic FP-equation with quadratic confining potential. There the propagator norm is computed explicitly. The result can be written as an exponential decay estimate with time dependent multiplicative constant, whose maximal value is the result we are looking for. A related result based on Phi-entropies can be found in [15], where improved time dependent decay rates are derived.

The main result of this paper (Theorem 3.3.4) is equality of the propagator norms of the PDE on the orthogonal complement of the space of equilibria and of its associated drift ODE. The underlying norms are the L^2-norm weighted by the inverse of the equilibrium distribution for the PDE, and the Euclidian norm for the ODE. This has two main consequences: First, the sharp (exponential) decay of the PDE is reduced to the same, but much easier question on the ODE level. The second consequence is that the hypocoercivity index (see [6, 1, 2]) of the drift matrix determines the short-time behaviour (in the sense of a Taylor series expansion) both of the drift ODE and the FP-equation. As a further consequence for solutions of the FP-equation we determine the short-time regularization from the weighted L^2-space to a weighted H^1-space. This result can be seen as an illustration of the fact that for the FP-equation hypocoercivity is equivalent to hypoellipticity. Finally, it is shown that the FP-equation can be considered as the second quantization of the drift ODE. This follows from the proof of the main theorem, where the FP-evolution is decomposed on invariant subspaces, in each of which the evolution is governed by a tensorized version of the drift ODE.

The paper is organized as follows: In Section 2 we transform the FP-operator $\tilde{L}$ to an equivalent version L such that $D = C_S$, the symmetric part of the drift matrix. The conditions for the existence of a unique positive steady state and for hypocoercivity are also set up. The main theorem is formulated in Section 3 together with the main consequences. The proof of the main theorem requires a long preparation that is split into Sections 4 and 5. In Section 4 we derive a spectral decomposition for the FP-operator into finite-dimensional invariant subspaces. This allows to see an explicit link with the drift ODE $\dot{x} = -Cx$. In order to make this link more evident, we work with the space of symmetric tensors, presented in Section 5. In Section 6 we give the proof of the main theorem as a corollary of the fact that the propagator norm on each subspace is an integer power of the propagator norm of the ODE evolution. Finally, in Section 7 the FP-operator is rewritten in the second quantization formalism.

3.2 Preliminaries and main result

3.2.1 Equilibria – normalized Fokker-Planck equation

The following theorem (from [6], Theorem 3.1 or [23], p. 41) states under which conditions on the matrices $\widetilde{D}$ and $\widetilde{C}$ there exists a unique steady state g_∞ for (3.1.1) and it provides its explicit form. We denote the *spectral gap of* $\widetilde{C}$ by $\mu(\widetilde{C}) := \min\{\mathrm{Re}(\lambda) : \lambda \text{ is an eigenvalue of } \widetilde{C}\}$.

Definition 3.2.1. We say that *Condition* $\widetilde{A}$ holds for the Equation (3.1.1), iff

1. the matrix $\widetilde{D}$ is symmetric, positive semi-definite,
2. there is no non-trivial $\widetilde{C}^T$-invariant subspace of $\ker\widetilde{D}$,
3. the matrix $\widetilde{C}$ is positive stable, i.e. $\mu(\widetilde{C}) > 0$.

Note that condition (2) is known as Kawashima's degeneracy condition [20] in the theory for systems of hyperbolic conservation laws. It also appears in [19] as a condition for hypoellipticity of FP-equations (see [36, Section 3.3] for the connection to hypocoercivity).

Theorem 3.2.2 (Steady state)**.** *There exist a unique (L^1-normalized) steady state $g_\infty \in L^1(\mathbb{R}^d)$ of (3.1.1), iff Condition $\widetilde{A}$ holds. It is given by the (non-isotropic) Gaussian*

$$g_\infty(y) = c_K \exp\left(-\frac{y^T K^{-1} y}{2}\right), \tag{3.2.1}$$

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

$$2\widetilde{D} = \widetilde{C}K + K\widetilde{C}^T, \tag{3.2.2}$$

and $c_K = (2\pi)^{-d/2}(\det K)^{-1/2}$ is the normalization constant.

In the above theorem, the matrix K can be represented analytically as

$$K = 2\int_0^\infty e^{-\widetilde{C}\tau}\widetilde{D}e^{-\widetilde{C}^T\tau}\,d\tau$$

(see [23], p. 41), and the numerical solution of (3.2.2) can be obtained with the Matlab routine *lyap*.

Under Condition $\widetilde{A}$ the FP-equation (3.1.1) can be rewritten (see Theorem 3.5, [6]) as

$$\partial_t g = \operatorname{div}_y\left(g_\infty(\widetilde{D}+\widetilde{R})\nabla_y\left(\frac{g}{g_\infty}\right)\right), \qquad y \in \mathbb{R}^d,\ t \in (0,\infty), \tag{3.2.3}$$

where $\widetilde{R} \in \mathbb{R}^{d\times d}$ is the anti-symmetric matrix $\widetilde{R} = \frac{1}{2}\left(\widetilde{C}K - K\widetilde{C}^T\right)$. The natural setting for the evolution equation (3.1.1) is the weighted L^2-space $\widetilde{\mathcal{H}} := L^2(\mathbb{R}^d, g_\infty^{-1})$ with the inner product

$$\langle g_1, g_2\rangle_{\widetilde{\mathcal{H}}} := \int_{\mathbb{R}^d} g_1(y)g_2(y)\frac{dy}{g_\infty(y)}.$$

Using the notations $\widetilde{V}_0 := \operatorname{span}_{\mathbb{R}}\{g_\infty\} \subset \widetilde{\mathcal{H}}$ and $C := K^{-1/2}\widetilde{C}K^{1/2}$ we can now formulate the main result of this paper: [1]

Theorem 3.2.3. *Let Condition $\widetilde{A}$ hold for the FP-equation* (3.1.1)*. Then the propagator norms of the FP-equation* (3.1.1) *and its corresponding drift ODE $\frac{d}{dt}x = -Cx$ are equal, i.e.,*

$$\left\|e^{-\widetilde{L}t}\right\|_{\mathcal{B}(\widetilde{V}_0^\perp)} = \left\|e^{-Ct}\right\|_{\mathcal{B}(\mathbb{R}^d)}, \qquad \forall t \geq 0, \tag{3.2.4}$$

where $\mathcal{B}(.)$ denotes the operator and Euclidean matrix norms (for more details see Definition 3.3.3 below).

[1] Note added in print: In the follow-up paper [9], Theorem 3.2.3 was recently extended to FP-equations with time dependent coefficient matrices $\widetilde{D}(t)$, $\widetilde{C}(t)$, provided that all these FP-operators with fixed t have the same steady state, i.e. if (3.2.2) holds for all t with a constant matrix K. In this extension the two propagators in (3.2.4) are replaced by the propagation operators that map the solution at time t_1 to the solution at time $t_2 \geq t_1$, both for the FP-equation and for the corresponding drift ODE $\frac{d}{dt}x = -C(t)x$. K being constant in time implies that the FP-normatization to (3.2.5), the spaces $\mathcal{H}$ and $\widetilde{\mathcal{H}}$, as well as the subspace decomposition in §3.4.1 are all time independent.

The fact that (3.2.4) involves the matrix C (and not $\widetilde{C}$), motivates to introduce the following coordinate transformation. Using $x := K^{-1/2}y,\ f(x) := (\det K)^{1/2} g(K^{1/2}x)$ transforms (3.1.1) into

$$\partial_t f = -Lf := \operatorname{div}_x(D\nabla_x f + Cxf) = \operatorname{div}_x\left(f_\infty C\nabla_x\left(\frac{f}{f_\infty}\right)\right), \tag{3.2.5}$$

where $D := K^{-1/2}\widetilde{D}K^{-1/2}$, and the steady state is the normalized Gaussian

$$f_\infty(x) = (2\pi)^{-d/2}e^{-|x|^2/2}. \tag{3.2.6}$$

This is due to the property

$$D = C_S := \frac{1}{2}\left(C + C^T\right), \tag{3.2.7}$$

which is a simple consequence of (3.2.2). We shall call a FP-equation *normalized,* if the diffusion and drift matrices satisfy (3.2.7).

For later reference we rewrite Condition $\tilde{A}$ in terms of the matrix C:

Definition 3.2.4. We say that *Condition A* holds for the Equation (3.2.5), iff

1. the matrix C_S is positive semi-definite,
2. there is no non-trivial C^T-invariant subspace of $\ker C_S$.

Proposition 3.2.5. *The Equation* (3.1.1) *satisfies Condition $\widetilde{A}$ iff its normalized version* (3.2.5) *satisfies Condition A. Moreover, Condition A implies that the matrix C is positive stable, i.e.* $\mu(C) > 0$.

Proof. Equivalence of the items (1) in Definitions 3.2.1 and 3.2.4 follows from $C_S = K^{-\frac{1}{2}}\widetilde{D}K^{-\frac{1}{2}}$. For the second item, let us assume that (2) in Definition 3.2.4 does not hold. Then, there exist $v \in \ker C_S,\ v \neq 0 \in \mathbb{R}^d$ such that

$$0 = C_S C^T v = (K^{-1/2}\widetilde{D}K^{-1/2})(K^{1/2}\widetilde{C}^T K^{-1/2})v = K^{-1/2}\widetilde{D}\widetilde{C^T}(K^{-1/2}v).$$

This implies $\widetilde{D}\widetilde{C}^T(K^{-1/2}v) = 0$, since $K^{-1/2} > 0$. But this is a contradiction to (2) in Condition $\widetilde{A}$ since it holds that $v \in \ker C_S$ iff $K^{-1/2}v \in \ker\widetilde{D}$. With a similar argument the reverse implication can be proven.

For the proof that Condition A implies positive stability of C we refer to Proposition 1 and Lemma 2.4 in [1]. □

From now on we shall study the normalized equation (3.2.5) on the normalized version $\mathcal{H} := L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$ of the Hilbert space $\widetilde{\mathcal{H}}$. It is easily checked that

$$\|g(t)\|_{\widetilde{\mathcal{H}}} = \|f(t)\|_{\mathcal{H}}, \qquad \forall t \geq 0, \tag{3.2.8}$$

holds for the solutions g and f of (3.1.1) and, respectively, (3.2.5). This implies that the propagator norms for $\widetilde{L}$ and L are the same, and that the Theorems 3.2.3 and 3.3.4 are equivalent.

3.2.2 Convergence to the equilibrium: hypocoercivity

In [6], a hypocoercive entropy method was developed to prove the exponential convergence to f_∞, for the solution to (3.2.5) with any initial datum $f_0 \in \mathscr{H}$. It employed a family of relative entropies w.r.t. the steady state, i.e. $e_\psi(f(t)|f_\infty) := \int_{\mathbb{R}^d} \psi\left(\frac{f(t)}{f_\infty}\right) f_\infty dx$, where the convex functions ψ are admissible entropy generators (as in [7] and [11]).

Definition 3.2.6. Given $\mu(C) := \min\{\mathrm{Re}(\lambda) : \lambda \text{ is an eigenvalue of } C\}$.

1. We call the matrix C *non-defective* if all the eigenvalues λ with $\mathrm{Re}(\lambda) = \mu(C)$ are *non-defective,* i.e., their algebraic and geometric multiplicities coincide.
2. We call a FP-equation (3.1.1) (non-)defective if its drift-matrix $\widetilde{C}$ is (non-)defective, or equivalently, if the matrix C in the normalized version (3.2.5) is (non-)defective.

For non-defective FP-equations, the decay result from [6] provides on the one hand the sharp exponential decay rate $\mu > 0$, but, on the other hand, only a sub-optimal multiplicative constant $c > 1$. We give a slightly modified version of it:

Theorem 3.2.7 (Exponential decay of the relative entropy, Theorem 4.9, [6])**.** *Let ψ generate an admissible entropy and let f be the solution of (3.2.5) with normalized initial state $f_0 \in L^1_+(\mathbb{R}^d)$ such that $e_\psi(f_0|f_\infty) < \infty$. Let C satisfy Condition A. Then, if the FP-equation is non-defective, there exists a constant $c \geq 1$ such that*

$$e_\psi(f(t)|f_\infty) \leq c^2 e^{-2\mu t} e_\psi(f_0|f_\infty), \quad t \geq 0. \tag{3.2.9}$$

Choosing the admissible quadratic function $\psi(\sigma) = (\sigma - 1)^2$ yields the exponential decay of the $\mathscr{H}$-norm. For this particular choice of ψ, Theorem 3.2.7 holds also for $f_0 \in L^1(\mathbb{R}^d) \cap \mathscr{H}$, i.e. the positivity of the initial datum f_0 is not necessary.

Corollary 3.2.8 (Hypocoercivity)**.** Under the assumptions of Theorem 3.2.7 the following estimate holds with the same $\mu > 0$, $c \geq 1$:

$$\|f(t) - f_\infty\|_{\mathscr{H}} \leq c e^{-\mu t} \|f_0 - f_\infty\|_{\mathscr{H}}, \quad t \geq 0. \tag{3.2.10}$$

The hypocoercivity approach in [6] provides the optimal (i.e. maximal) value for μ and a computable value for c, which is however not sharp, i.e. $c > c_{\min}$ with

$$c_{\min} := \min\left\{ c \geq 1 : \text{ (3.2.10) holds for all } f_0 \in \mathscr{H} \text{ with } \int_{\mathbb{R}^d} f_0 \, dx = 1 \right\}. \tag{3.2.11}$$

One central goal of this paper is the determination of $c_{\min}$. But, actually, we shall go much beyond this: The main result of this paper, Theorem 3.3.4, states that the $\mathscr{H}$-propagator norm of each (stable) FP-equation is *equal* to the (spectral) propagator norm of its corresponding drift ODE $\dot{x}(t) = -Cx(t)$. Hence, all decay properties of the FP-equation (3.1.1) can be obtained from a simple linear ODE, and sharp exponential decay estimates of this ODE carry over to the corresponding FP-equation. So, for quantifying the decay behaviour of FP-equations with linear drift, an infinite dimensional PDE problem can be replaced by a (small) finite dimensional ODE problem.

3.2.3 The best multiplicative constant for the ODE-decay

In [3] we analysed the best decay constants for the (of course easier) finite dimensional problem

$$\dot{x}(t) = -Cx(t), \quad t > 0, \qquad x(0) = x_0 \in \mathbb{C}^n, \tag{3.2.12}$$

where $C \in \mathbb{C}^{n\times n}$ is a positive stable and non-defective matrix. In this case we constructed a problem adapted norm as a Lyapunov functional. This allowed to derive a hypocoercive estimate for the Euclidean norm $\|\cdot\|_2$ of the solution:

$$\|x(t)\|_2 \le c e^{-\mu t} \|x_0\|_2, \qquad t \ge 0. \tag{3.2.13}$$

Here $\mu > 0$ is the spectral gap of the matrix C (and the sharp decay rate of the ODE (3.2.12)), and $c \ge 1$ is some constant.

In [3] we investigated, in the two dimensional case, the sharpness of the constant c. By analogy with (3.2.11), we define the best multiplicative constant for the hypocoercivity estimate of the ODE as

$$c_1 := c_1(C) := \min\left\{c \ge 1 : \text{ (3.2.13) holds for all } x_0 \in \mathbb{C}^n\right\}.$$

The explicit expression for the best constant c_1 depends on the spectrum of C. In [3] we treated all the cases for matrices in $\mathbb{C}^{2\times 2}$. In particular, denoting by λ_1, λ_2 the two eigenvalues of C, we distinguish three cases:

1. $\mathrm{Re}(\lambda_1) = \mathrm{Re}(\lambda_2) = \mu$;
2. $\mu = \mathrm{Re}(\lambda_1) < \mathrm{Re}(\lambda_2)$, $\mathrm{Im}(\lambda_1) = \mathrm{Im}(\lambda_2)$;
3. $\mu = \mathrm{Re}(\lambda_1) < \mathrm{Re}(\lambda_2)$, $\mathrm{Im}(\lambda_1) \ne \mathrm{Im}(\lambda_2)$.

The corresponding explicit form of c_1 in the cases (1) and (2) is described in the next theorem (see Theorem 3.7 and Theorem 4.1 in [3]). For the case (3) we have, instead, an implicit form, see Proposition 4.2 and Corollary 4.3 in [3].

Theorem 3.2.9. *Let $C \in \mathbb{C}^{2\times 2}$ be positive stable and non-defective with eigenvalues λ_1, λ_2. Denoting by $\alpha \in [0,1)$ the cosine of the angle between the two eigenvectors of C^T, the best constant for* (3.2.13) *in the cases* (1) *and* (2) *is*

$$c_1 = \sqrt{\frac{1+\alpha}{1-\alpha}} \qquad \textit{and, respectively,} \qquad c_1 = \frac{1}{\sqrt{1-\alpha^2}}.$$

For dimension $n \ge 3$, explicit expressions for the best constant c_1 seem to be unknown in general.

The defective case

So far we have discussed non-defective matrices $C \in \mathbb{R}^{d\times d}$. The remaining case has to be treated apart since we cannot obtain both the optimality of the multiplicative constant and the sharpness of the exponential decay at the same time if C is defective. Nevertheless, hypocoercive estimates do hold (see Chapter 1.8 in [29] and Theorem 2.8 in [8]) with either reduced exponential decay rates (see Theorem 4.9 in [6]) or with the best decay rate μ, but augmented with a time-polynomial coefficient (see Theorem 2.8 in [8]), as the following theorem claims.

Theorem 3.2.10. *Let $C \in \mathbb{C}^{d\times d}$ be a positive stable (possibly defective) matrix with spectral gap $\mu > 0$. Let M be the maximal size of a Jordan block associated to μ. Let $x(t)$ be the solution of the ODE $\frac{d}{dt}x(t) = -Cx(t)$ with initial datum $x_0 \in \mathbb{C}^d$. Then, for each $\epsilon > 0$ there exist a constant $c_\epsilon \geq 1$ such that*

$$\|x(t)\|_2 \leq c_\epsilon e^{-(\mu-\epsilon)t}\|x_0\|_2, \qquad \forall\, t \geq 0, x_0 \in \mathbb{C}^d. \tag{3.2.14}$$

Moreover, there exists a polynomial $p(t)$ of degree $M-1$ such that

$$\|x(t)\|_2 \leq p(t)e^{-\mu t}\|x_0\|_2, \qquad \forall\, t \geq 0, x_0 \in \mathbb{C}^d. \tag{3.2.15}$$

As we did for the non-defective case, we define the best constant $c_{1,\epsilon}$ for the estimate (3.2.14) with rate $\mu-\epsilon$ as

$$c_{1,\epsilon} := \min\left\{c_\epsilon \geq 1 : \text{ (3.2.14) holds for all } x_0 \in \mathbb{C}^d\right\}.$$

We do not attempt to define an "optimal polynomial" $p(t)$ in (3.2.15). In the next section it is shown that these ODE-results carry over to the corresponding FP-equation (3.2.5).

3.3 Main result for normalized FP-equations and applications

In Theorem 3.2.3 we anticipated the main result of this paper for the *non-normalized* FP-equation (3.1.1). In the sequel we shall deal with its equivalent formulation for *normalized* FP-equations, since this will simplify the proof. With the above review of ODE results we can now state an essential aspect of this main result: The best decay constants in (3.2.10) for the FP-equation (3.2.5) (and therefore also for (3.1.1)) coincide with the best constants for the ODE (3.2.12). This result is a corollary of the main theorem of this paper, namely Theorem 3.3.4. It claims that the propagator norm of the FP-equation coincides with the propagator norm of its corresponding ODE (w.r.t. the Euclidean vector norm). With *propagator norm* we refer to the following notion for linear ODEs or PDEs: If A is their infinitesimal generator on some Banach space X and e^{At}, $t \geq 0$ their *propagator*, forming a $\mathscr{C}_0$-semigroup of bounded operators (cf. [28]), the propagator norm is the operator norm of e^{At} on X, see Definition 3.3.3 below.

First we define the projection operator Π_0 that maps a function in $\mathscr{H}$ into the subspace generated by the steady state f_∞.

Definition 3.3.1. Let $f \in \mathscr{H} = L^2\left(\mathbb{R}^d, f_\infty^{-1}\right)$ and f_∞ the normalized Gaussian (3.2.6). We define the operator $\Pi_0 : \mathscr{H} \longrightarrow \mathscr{H}$ as

$$\Pi_0 f := \langle f, f_\infty\rangle_{\mathscr{H}} f_\infty,$$

i.e., Π_0 projects f onto $V_0 := \operatorname{span}_{\mathbb{R}}\{f_\infty\} = \mathscr{N}(L)$.

Remark 3.3.2. Let $f \in \mathscr{H}$. Then, the coefficient $\langle f, f_\infty\rangle_{\mathscr{H}}$ is equal to $\int_{\mathbb{R}^d} f(x)dx$, by definition. Moreover, it is obvious from the divergence form of (3.2.5) that the "total mass" $\int_{\mathbb{R}^d} f(t,x)dx$ remains constant in time under the flow of the equation. Hence, $(\Pi_0 f)(t)$ is independent of t, if $f(t)$ solves (3.2.5). This implies $e^{-Lt}(\mathbb{1} - \Pi_0) = e^{-Lt} - \Pi_0$.

We introduce the standard definitions of operator norms.

Definition 3.3.3. Let $A : \mathscr{H} \to \mathscr{H}$ and $B : \mathbb{R}^d \to \mathbb{R}^d$ be linear operators. Then

$$\|A\|_{\mathscr{B}(\mathscr{H})} := \sup_{0\neq f\in\mathscr{H}} \frac{\|Af\|_{\mathscr{H}}}{\|f\|_{\mathscr{H}}}, \qquad \|B\|_{\mathscr{B}(\mathbb{R}^d)} := \sup_{0\neq x\in\mathbb{R}^d} \frac{\|Bx\|_2}{\|x\|_2}.$$

If $f(t)$ is the solution of the FP-equation (3.2.5) with $f(0) = f_0 \in \mathcal{H}$, then

$$\left\| e^{-Lt}(\mathbb{1}-\Pi_0)\right\|_{\mathscr{B}(\mathcal{H})} = \left\| e^{-Lt}\right\|_{\mathscr{B}(V_0^\perp)} = \sup_{0\neq f_0\in\mathcal{H}} \frac{\|f(t)-\Pi_0 f_0\|_{\mathcal{H}}}{\|f_0\|_{\mathcal{H}}}.$$

If $x(t) \in \mathbb{R}^d$ is the solution of the ODE $\frac{d}{dt}x = -Cx$ with initial datum $x(0) := x_0$, then

$$\left\| e^{-Ct}\right\|_{\mathscr{B}(\mathbb{R}^d)} = \sup_{0\neq x_0\in\mathbb{R}^d} \frac{\|x(t)\|_2}{\|x_0\|_2}.$$

With these notations we can state the main result of this paper.

Theorem 3.3.4. *Let Condition A hold for the FP-equation* (3.2.5). *Then the propagator norms of the FP-equation* (3.2.5) *and its corresponding ODE* $\frac{d}{dt}x = -Cx$ *are equal, i.e.,*

$$\left\| e^{-Lt}\right\|_{\mathscr{B}(V_0^\perp)} = \left\| e^{-Ct}\right\|_{\mathscr{B}(\mathbb{R}^d)}, \qquad \forall t \geq 0. \tag{3.3.1}$$

The proof of Theorem 3.3.4 will be prepared in the following two sections and finally completed in Section 3.6.

Theorem 3.3.4 can be seen as a generalization of a result in [17], where the propagator norm for the following kinetic FP-equation (the L^2-adjoint equation of (2) in [17])

$$\begin{aligned} \partial_t g &= -\widetilde{L}_a g := -v\partial_x g + \partial_v(\partial_v g + (ax+v)g) \\ &= \operatorname{div}_{(x,v)}\left(\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\nabla_{(x,v)} g + \begin{pmatrix}0 & -1\\ a & 1\end{pmatrix}\begin{pmatrix}x\\ v\end{pmatrix} g\right), \end{aligned} \tag{3.3.2}$$

with $(x, v) \in \mathbb{R}^2$ and the parameter $a > 0$, has been computed explicitly.

Theorem 3.3.5. *[17, Theorem 1.2] For any* $a > 0$ *and* $t \geq 0$*, it holds:*

$$\left\| e^{-\widetilde{L}_a t}\right\|_{\mathscr{B}(V_0^\perp)} = c_a(t)\exp\left(-\frac{1-\sqrt{(1-4a)_+}}{2}\,t\right), \tag{3.3.3}$$

where the non-negative factor $c_a(t)$ *is given for* $0 < a < 1/4$ *by*

$$c_a(t) := \sqrt{e^{-2\theta t} + \frac{1-\theta^2}{2\theta^2}(1-e^{-\theta t})^2 + \frac{1-e^{-2\theta t}}{2}\left(1+\frac{1}{\theta}\sqrt{1+(\theta^{-2}-1)\left(\frac{e^{\theta t}-1}{e^{\theta t}+1}\right)^2}\right)}, \tag{3.3.4}$$

with $\theta = \sqrt{1-4a}$*, for* $a > 1/4$ *by*

$$c_a(t) := \sqrt{1 + \frac{|e^{\theta t}-1|}{2|\theta|^2}\left(|e^{\theta t}-1| + \sqrt{|e^{\theta t}-1|^2 + 4|\theta|^2}\right)}, \tag{3.3.5}$$

with $\theta := \sqrt{4a-1}\,i$*, and for* $a = 1/4$ *by*

$$c_a(t) := \sqrt{1 + \frac{t^2}{2} + t\sqrt{1+\left(\frac{t}{2}\right)^2}}. \tag{3.3.6}$$

Note that there is a small typo in the formula for $c_a(t)$, $a < 1/4$ in [17] that corresponds to (3.3.4).

After normalization of the FP-equation (3.3.2), the corresponding drift matrix is given by

$$C_a := \begin{pmatrix} 0 & -\sqrt{a} \\ \sqrt{a} & 1 \end{pmatrix}. \tag{3.3.7}$$

Its eigenvalues are $\lambda_{1,2} := \frac{1}{2}(1 \pm \theta)$, with θ as in Theorem 3.3.5, and the corresponding eigenvectors are $v_{1,2} = (\sqrt{a}, -\lambda_{1,2})^T$. This shows that the spectral gap is given by $\mu = \frac{1}{2}\left(1 - \sqrt{(1-4a)_+}\right)$. It is easy to check that C_a satisfies Condition A for each $a > 0$. We observe that the value $a = 1/4$ is critical in the sense that $C_{1/4}$ is defective.

With the approach of this work we can employ the results of Section 3.2.3 for obtaining the best possible constant c_1 in

$$\left\| e^{-\widetilde{L}_a t} \right\|_{\mathscr{B}(V_0^\perp)} = \left\| e^{-C_a t} \right\|_{\mathscr{B}(\mathbb{R}^d)} \le c_1 e^{-\mu t}.$$

For $a \neq 1/4$ we apply Theorem 3.2.9 and note that for $0 < a < 1/4$ we are in case (2). We compute $\alpha = 2\sqrt{a}$, giving the optimal constant

$$c_1 = (1-4a)^{-1/2},$$

which can also be obtained from (3.3.4) in the limit $t \to \infty$. For $a > 1/4$ we are in case (1) and obtain $\alpha = (2\sqrt{a})^{-1}$ and

$$c_1 = \frac{2\sqrt{a}+1}{\sqrt{4a-1}}.$$

The same is obtained as the maximal value of $c_a(t)$ in (3.3.5), taken whenever $\left|e^{\theta t} - 1\right| = 2$.

Finally, for $a = 1/4$ the results of Theorems 3.2.10 and 3.3.5 agree with $c_a(t) \approx t$ as $t \to \infty$, since the best approximation for the function in (3.3.6), i.e. the smallest affine linear upper bound to (3.3.6), is the polynomial $p(t) = 1 + t$.

The plot in Figure 3.1 shows the right-hand side of (3.3.3) as a function of time for 3 values of a ($a = 1/5$, $a = 1/4$, $a = 2$). Note the non-smooth behaviour in the case $a = 2$.

3.3.1 Applications of Theorem 3.3.4

Long time behaviour

One consequence of Theorem 3.3.4 is that all the estimates about the decay of the solutions of the ODE carry over to the corresponding FP-equation. In particular, it follows that the hypocoercive ODE estimates (3.2.13) and (3.2.14) hold also for solutions of the corresponding FP-equation. Moreover, the best constants in the estimates are the same both for the FP-case and for its corresponding drift ODE.

Theorem 3.3.6. *Let $C \in \mathbb{R}^{d\times d}$ be non-defective and satisfy Condition A. Let c_1 be the best constant in the estimate* (3.2.13) *for the ODE* (3.2.12). *Then it is also the optimal constant $c_{\min}$ in the following hypocoercive estimate*

$$\|f(t) - f_\infty\|_{\mathscr{H}} \le c_1 e^{-\mu t} \|f_0 - f_\infty\|_{\mathscr{H}}, \quad \forall t \ge 0, \forall f_0 \in \mathscr{H}, \int_{\mathbb{R}^d} f_0(x)\,dx = 1 \tag{3.3.8}$$

for the solution of the FP-equation (3.2.5).

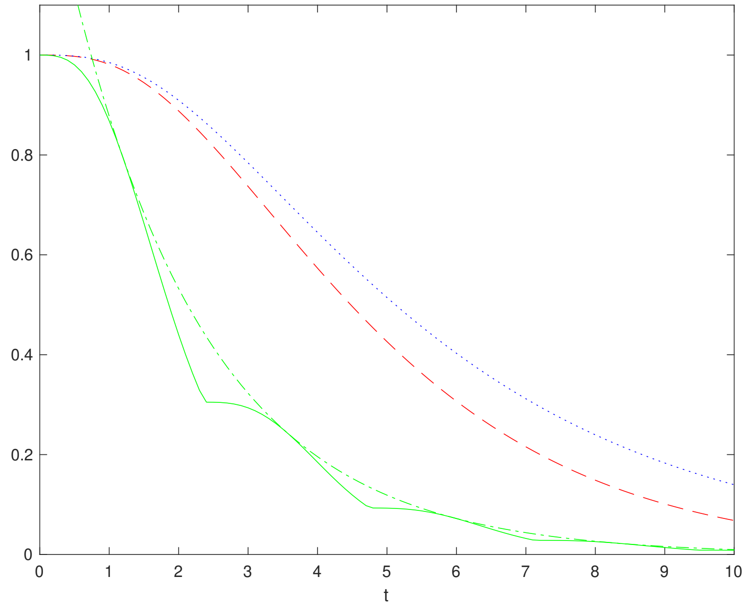

Figure 3.1: The propagator norm for equation (3.3.2) for 3 values of the parameter a. Solid (green) curve for $a = 2$, dashed (red) curve for $a = 1/4$, dotted (blue) curve for $a = 1/5$. The dash-dotted (green) curve, gives the best exponential bound of the form $c_1 e^{-t/2}$ for the case $a = 2$. Note: The curves are colored only in the electronic version of this article.

Theorem 3.3.7. *Let $C \in \mathbb{R}^{d\times d}$ be defective and satisfy Condition A. Let M be the maximal size of a Jordan block associated to μ. Let $\epsilon > 0$ be fixed and $c_{1,\epsilon}$ be the best constant in the estimate* (3.2.14) *for the ODE* (3.2.12)*. Then the following hypocoercive estimate holds*

$$\|f(t) - f_\infty\|_{\mathcal{H}} \le c_{1,\epsilon} e^{-(\mu-\epsilon)t} \|f_0 - f_\infty\|_{\mathcal{H}}, \quad \forall t \ge 0, \forall f_0 \in \mathcal{H}, \int_{\mathbb{R}^d} f_0(x)\,dx = 1 \tag{3.3.9}$$

for the solution of the FP-equation (3.2.5)*, and $c_{1,\epsilon}$ is the optimal multiplicative constant. Moreover,*

$$\|f(t) - f_\infty\|_{\mathcal{H}} \le p(t) e^{-\mu t} \|f_0 - f_\infty\|_{\mathcal{H}}, \quad \forall t \ge 0, \forall f_0 \in \mathcal{H}, \int_{\mathbb{R}^d} f_0(x)\,dx = 1, \tag{3.3.10}$$

where $p(t)$ is the polynomial of degree $M - 1$ appearing in (3.2.15).

We remind that the quest to obtain the best decay for (3.1.1) is thus reduced to the knowledge of the best decay constants for the corresponding drift ODE.

Short time behaviour

The second application of Theorem 3.3.4 concerns the short time behaviour of the propagator norm of the FP-operator. It is linked to the concept of *hypocoercivity index*, which describes the

”structural complexity” of the matrix C and, more precisely, the intertwining of its symmetric and anti-symmetric parts. For the FP-equation, the hypocoercivity index reflects its degeneracy structure. As we are going to illustrate in this section, this index represents the polynomial degree in the short time behaviour of the propagator norm, both in the FP-equation and in the ODE case. Moreover it describes the rate of regularization of the FP-solution from $\mathscr{H}$ to a weighted Sobolev space H^1.

Next we recall the definition of *hypocoercivity index* both for FP-equations and ODEs, respectively, from [6] and [1, 2]. We will see that these two concepts coincide when we consider the drift ODE associated to the FP-equation. We first give the definition for the normalized FP-equation and then it will be illustrated that the index is invariant for the general ($D \neq C_S$) equation (3.1.1).

Definition 3.3.8. We define m_{HC}, the *hypocoercivity index* for the normalized FP-equation (3.2.5) as the minimum $m \in \mathbb{N}_0$ such that

$$T_m := \sum_{j=0}^{m} C_{AS}^j C_S (C_{AS}^T)^j > 0\,. \tag{3.3.11}$$

Here $C_{AS} := \frac{1}{2}(C - C^T)$ denotes the anti-symmetric part of C.

Remark 3.3.9. Lemma 2.3 in [6] states that the condition $m_{HC} < \infty$ is equivalent to the FP-equation being hypoelliptic. This index can be seen as a measure of “how much” the drift matrix has to mix the directions of the kernel of the diffusion matrix with its orthogonal space in order to guarantee convergence to the steady state. For example, $m_{HC} = 0$ means, by definition, that the diffusion matrix $D = C_S$ is positive definite, and hence coercive. In general, m_{HC} is finite when we are assuming Condition A (see Lemma 2.3, [6]).

For completeness, we include the definition of hypocoercivity index also for the non-normalized case. For simplicity we will denote it as well with m_{HC}. This is actually allowed since the next proposition will prove that these two definitions are unchanged under normalization.

Definition 3.3.10. We define m_{HC} the *hypocoercivity index* for the FP-equation (3.1.1) as the minimum $m \in \mathbb{N}_0$ such that

$$\widetilde{T}_m := \sum_{j=0}^{m} \widetilde{C}^j \widetilde{D} (\widetilde{C}^T)^j > 0, \tag{3.3.12}$$

and $m_{HC} = \infty$ if this minimum does not exist.

Proposition 3.3.11. *Let us consider the FP-equation* (3.1.1) *and its normalized version* (3.2.5). *Let Condition $\widetilde{A}$ (or, equivalently, Condition A) be satisfied. Then, the hypocoercivity indices of the two equations coincide, i.e., for any $m \in \mathbb{N}_0$*

$$T_m > 0 \quad \textit{if and only if} \quad \widetilde{T}_m > 0. \tag{3.3.13}$$

Proof. First we recall from Lemma 3.4, [2] that

$$\sum_{j=0}^{m} C_{AS}^j C_S (C_{AS}^T)^j > 0 \quad \text{if and only if} \quad \sum_{j=0}^{m} C^j C_S (C^T)^j > 0. \tag{3.3.14}$$

The second step consists in proving that $\widetilde{T}_m > 0$ iff

$$\hat{T}_m := \sum_{j=0}^{m} C^j D (C^T)^j > 0,$$

where $C = K^{-1/2}\widetilde{C}K^{1/2}$ and $D = K^{-1/2}\widetilde{D}K^{-1/2} = C_S$ are the matrices appearing in the normalized equation and K from (3.2.2). By substituting we get

$$\begin{aligned}\hat{T}_m =& \sum_{j=0}^{m} (K^{-1/2}\widetilde{C}K^{1/2})^j K^{-1/2}\widetilde{D}K^{-1/2}(K^{1/2}\widetilde{C}^T K^{-1/2})^j \\ =& K^{-1/2}\sum_{j=0}^{m} \widetilde{C}^j \widetilde{D}(\widetilde{C}^T)^j K^{-1/2} \\ =& K^{-1/2}\widetilde{T}_m K^{-1/2}.\end{aligned}$$

Then, it is immediate to conclude that the positivity of the two matrices is equivalent since $K > 0$.

Combining this last equivalence with (3.3.14) yields (3.3.13). □

Remark 3.3.12. We shall now compare the hypocoercivity index m_{HC} of the normalized FP-equation (3.2.5) to the commutator condition (3.5) in [36]. To this end we rewrite (3.2.5) for $h(x,t) := f(x,t)/f_\infty(x)$. In Hörmander form it reads

$$\partial_t h = \operatorname{div}(C\nabla h) - x^T C\nabla h = -(A^*A + B)h, \tag{3.3.15}$$

where the adjoint A^* is taken w.r.t. $L^2(\mathbb{R}^d, f_\infty)$. Here, the vector valued operator A and the scalar operator B are given by

$$A := \sqrt{D}\cdot\nabla, \qquad B := x^T \cdot C_{AS}\cdot\nabla.$$

Following §3.3 in [36] we define the iterated commutators

$$C_0 := A, \qquad C_k := [C_{k-1}, B].$$

They are vector valued operators mapping from $L^2(\mathbb{R}^d, f_\infty)$ to $(L^2(\mathbb{R}^d, f_\infty))^d$. Hence, the nabla operator in B can be either the gradient or the Jacobian, depending on the dimensionality of the argument of B. By induction one easily verifies that $C_k = \sqrt{D}\cdot C_{AS}^k \cdot \nabla$, $k \in \mathbb{N}_0$.

We recall condition (3.5) from [36]: "There exists $N_c \in \mathbb{N}_0$ such that

$$\sum_{k=0}^{N_c} C_k^* C_k \quad \text{is coercive on } \ker(A^*A + B)^\perp.\text{"} \tag{3.3.16}$$

Note that $\ker(A^*A + B)$ consists of the constant functions, and its orthogonal complement is $\{h \in L^2(\mathbb{R}^d, f_\infty) : \int_{\mathbb{R}^d} h f_\infty dx = 0\}$. The coercivity in (3.3.16) reads

$$\int_{\mathbb{R}^d} \nabla^T h \cdot T_{N_c} \cdot \nabla h f_\infty dx \ge \kappa \int_{\mathbb{R}^d} h^2 f_\infty dx \tag{3.3.17}$$

for some $\kappa > 0$ and all $h \in \ker(A^*A + B)^\perp$, where $T_{N_c} := \sum_{k=0}^{N_c} (C_{AS}^T)^k D C_{AS}^k$. Clearly, the weighted Poincaré inequality (3.3.17) holds iff $T_{N_c} > 0$, see §3.2 in [7], e.g. Hence, the minimum N_c for condition (3.3.16) to hold equals the hypocoercivity index m_{HC} from Definition 3.3.8 above.

Next we shall link the hypocoercivity index of the FP-equation with the hypocoercivity index m_{HC} of its associated ODE $\dot{x}(t) = -Cx(t)$, which is defined in the same way. At the ODE level, this index describes the short time decay of the propagator norm $\|e^{-Ct}\|_{\mathscr{B}(\mathbb{R}^d)}$ as it is shown in the following theorem (see Theorem 3.2, [2]).

Remark 3.3.13. We note that our hypocoercivity index m_{HC} also coincides with the index appearing in the characterization of the *singular space S* of the FP-operator, i.e. the smallest integer k_0 such that

$$\bigcap_{j=0}^{k_0} \ker[C_S(C_{AS})^j] = S = \{0\}$$

(see (2.9) in [4], (3.22) in [26]). The equivalence of these two indices follows since they are both equivalent to the smallest integer τ in the *Kalman rank condition*, i.e.

$$\operatorname{rank}\{\sqrt{C_S}, C_{AS}\sqrt{C_S}, ..., C_{AS}^{\tau}\sqrt{C_S}\} = d\ .$$

This was established in Proposition 1 of [1] and, respectively, on pages 705/706 of [26]. The latter proof uses the version (3.3.15) of the FP-equation.

Theorem 3.3.14. *Let C satisfy Condition A. Then its hypocoercivity index is* $m_{HC} \in \mathbb{N}_0$ *(and hence finite) if and only if*

$$\|e^{-Ct}\|_{\mathscr{B}(\mathbb{R}^d)} = 1 - ct^{\alpha} + \mathscr{O}(t^{\alpha+1}), \qquad \text{as } t \to 0+, \tag{3.3.18}$$

for some $c > 0$*, where* $\alpha := 2m_{HC} + 1$.

Remark 3.3.15. We observe that, in the coercive case (i.e., $m_{HC} = 0$), the propagator norm satisfies an estimate of the form

$$\|e^{-Ct}\|_{\mathscr{B}(\mathbb{R}^d)} \le e^{-\lambda t}, \quad t \ge 0, \text{ for some } \lambda > 0. \tag{3.3.19}$$

In that case ($\alpha = 1$) Theorem 3.3.14 states that the propagator norm $\|e^{-Ct}\|_{\mathscr{B}(\mathbb{R}^d)}$ behaves as $g(t) := 1 - ct$ for short times. With $c = \lambda$, this is the (initial part of the) Taylor expansion of the exponential function in (3.3.19).

Next we shall use this result to derive information about the short time behaviour of the Fokker-Planck propagator norm $\|e^{-Lt}\|_{\mathscr{B}(V_0^{\perp})}$. By Theorem 3.3.4 the propagator norms of the FP-equation and the corresponding ODE coincide.

Theorem 3.3.16. *Let L be the Fokker-Planck operator defined in* (3.2.5)*. Let C satisfy Condition A. Then the hypocoercivity index of* (3.2.5) *is* $m_{HC} \in \mathbb{N}_0$ *(and hence finite) if and only if*

$$\|e^{-Lt}\|_{\mathscr{B}(V_0^{\perp})} = 1 - ct^{\alpha} + \mathscr{O}(t^{\alpha+1}), \quad t \to 0+, \tag{3.3.20}$$

where $\alpha = 2m_{HC} + 1$*, for some* $c > 0$.

Proof. This result is an immediate corollary of Theorem 3.3.4 and Theorem 3.3.14, by recalling that the FP-equation and its associated ODE have the same hypocoercivity index. □

Remark 3.3.17. As for the ODE case, the equality (3.3.20) shows that the index m_{HC} describes how fast the propagator norm decays for short times. This is consistent with the fact that the coercive case ($m_{HC} = 0$) corresponds to the fastest behaviour, i.e., with an exponential decay ($\alpha = 1$). In general, the bigger the index, the slower is the decay of the norm for short times.

Example 3.3.18. In Theorem 1.2 of [17] the authors derive the explicit expression for the propagator norm of the FP-equation associated to the matrix (3.3.7), see Theorem 3.3.5. With it they also estimate the short time behaviour of this norm, depending on the parameter a. In the case $a > 0$, equality (2) in [17] implies

$$\left\| e^{-\tilde{L}_a t} \right\|_{\mathscr{B}(V_0^\perp)} = 1 - \frac{a}{6} t^3 + o(t^3).$$

We note that this result is consistent with the equality (3.3.20). Indeed, it is easy to verify that for $a > 0$ the matrix C_a has hypocoercivity index $m_{HC} = 1$. Hence the exponent in the polynomial short time behaviour turns out to be $\alpha = 3$, as above. □

It is known that the hypocoercivity index also has a second implication on the qualitative behaviour of FP-equations, namely the rate of regularization from some weighted L^2-space into a weighted H^1-space (like in non-degenerate parabolic equations). The following proposition was proven in [36] (see §7.3, §A.21 for the kinetic FP-equation with $m_{HC} = 1$. The extension from Theorem A.12 is given without proof and includes a small typo.) and in [6, Theorem 4.8]. The following result can also be seen as a special case of (2.21) as well as of Theorem 2.6 in [4].

Proposition 3.3.19. *Let $f(t)$ be the solution of* (3.2.5). *Let C satisfy Condition A and m_{HC} be its associated hypocoercivity index. Then, there exist $\tilde{c}, \delta > 0$, such that*

$$\left\| f_\infty \nabla \left(\frac{f(t)}{f_\infty} \right) \right\|_{\mathscr{H}} \le \tilde{c}\, t^{-\alpha/2} \| f_0 \|_{\mathscr{H}}, \qquad 0 < t \le \delta, \tag{3.3.21}$$

with $\alpha := 2m_{HC} + 1$ for all $f_0 \in \mathscr{H}$.

So far we have seen that the hypocoercivity index of a FP-equation determines both the short time decay and its regularization rate. An obvious question is now to understand the relation of these two qualitative properties. The following proposition shows that they are essentially equivalent for the family (3.2.5) of FP-equations:

Proposition 3.3.20. *Let the matrix C satisfy Condition A (see Definition 3.2.4), and let $f(t)$ be the solution of* (3.2.5). *We denote its propagator norm by $\| e^{-Lt} \|_{\mathscr{B}(V_0^\perp)} =: \tilde{h}(t)$, $t \ge 0$.*

(a) Assume that $\tilde{h}(t) = 1 - ct^\alpha + o(t^\alpha)$ as $t \to 0^+$ for some $c > 0$ and $\alpha > 0$. Then the regularization estimate (3.3.21) *follows with the same α, and for all $f_0 \in \mathscr{H}$. Moreover, this α in* (3.3.21) *is optimal (i.e. minimal).*

(b) Let there exist some $\tilde{c}, \delta > 0$ and $\alpha > 0$ (not necessarily integer) such that (3.3.21) *holds for all $f_0 \in \mathscr{H}$. Then, there are $\delta_2 > 0$ and $c_2 > 0$, such that $\tilde{h}(t) \le 1 - c_2 t^\alpha$ on $0 \le t \le \delta_2$. Moreover, if α is minimal in the assumed regularization estimate* (3.3.21)*, then it is also minimal in the concluded decay estimate $\tilde{h}(t) \le 1 - c_2 t^\alpha$.*

The proof of Proposition 3.3.20 can be found in the §5.7, since it requires results that will be presented in the next sections.

Remark 3.3.21. We note that the statements (3.3.20) and (3.3.21) are different in nature: While the equality (3.3.20) *characterizes* the short-time decay of e^{-Lt}, the inequality (3.3.21) only provides an *upper bound* for the short time regularization of e^{-Lt}. Hence, since Proposition 3.3.19 is based on (3.3.21), it can only yield the conclusion $\tilde{h}(t) \leq 1 - c_2 t^{\alpha}$, which is also just an upper bound for the short time behaviour, rather than the dominant part of the Taylor expansion of $\tilde{h}(t)$. But if α is known to be minimal in (3.3.21), then it is also minimal for (3.3.20).

Remark 3.3.22. Proposition 3.3.19 provides an *isotropic* regularization rate. We note that this result can be improved for degenerate, hypocoercive FP-equations, and it gives rise to anisotropic smoothing: There the regularization is faster in the diffusive directions of $(\ker C_S)^{\perp}$ than in the non-diffusive directions of $\ker C_S$. "Faster" corresponds here to a smaller exponent in (3.3.21).

An example of different speeds of regularization is given in [32, Section 11] for the solution $f(t,x,v)$ of a kinetic FP-equation in $\mathbb{T}^d \times \mathbb{R}^d$ without confinement potential. In that case the short-time regularization estimate for the v-derivatives is the same as for the heat equation, since the operator is elliptic in v. But the regularization in x has an exponent 3 times as large; this corresponds, respectively, to the two cases $m_{HC} = 0, 1$ in (3.3.21). A more general result about anisotropic regularity estimates can be found in [36, Section A.21.2]. In an alternative description one can fix a uniform regularization rate in time, by considering different regularization orders (i.e. higher order derivatives) in different spatial directions in the setting of *anisotropic* Sobolev spaces. A definition of these functional spaces and an example of this behaviour is provided in [26], regarding the solution of a degenerate Ornstein-Uhlenbeck equation.

3.4 Solution of the FP-equation by spectral decomposition

In order to link the evolution in (3.2.5) to the corresponding drift ODE $\dot{x} = -Cx$ we shall project the solution $f(t) \in \mathscr{H}$ of (3.2.5) to finite dimensional subspaces $\{V^{(m)}\}_{m\in\mathbb{N}_0} \subset \mathscr{H}$ with $LV^{(m)} \subseteq V^{(m)}$. Then we shall show that, surprisingly, the evolution in each subspace can be based on the single ODE $\dot{x} = -Cx$. Moreover, the solution component in the subspace $V^{(1)}$ will turn out to decay the slowest, and it is hence the dominant part.

3.4.1 Spectral decomposition of the Fokker Planck operator

First we define the finite dimensional, L-invariant subspaces $V^{(m)} \subset \mathscr{H}$. Let the dimension $d \geq 1$ be fixed. We recall that the (normalized) steady state of (3.2.5) is given by $g_0(x) := f_\infty(x) = \prod_{i=1}^d g(x_i)$, $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, where $g(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$ is the one-dimensional (normalized) Gaussian. The construction and results about the spectral decomposition of L that we are going to summarize can be found in [6, Section 5].

Definition 3.4.1. Let $\alpha = (\alpha_i) \in \mathbb{N}_0^d$ be a multi-index. Its order is denoted by $|\alpha| = \sum_{i=1}^d \alpha_i$. For a fixed $\alpha \in \mathbb{N}_0^d$ we define

$$g_\alpha(x) := (-1)^{|\alpha|} \nabla_x^\alpha g_0(x), \tag{3.4.1}$$

or, equivalently,

$$g_\alpha(x) := \prod_{i=1}^{d} H_{\alpha_i}(x_i) g(x_i), \quad \forall x = (x_i) \in \mathbb{R}^d, \tag{3.4.2}$$

where, for any $n \in \mathbb{N}_0$, H_n is the *probabilists' Hermite polynomial* of order n defined as

$$H_n(y) := (-1)^n e^{\frac{y^2}{2}} \frac{d^n}{dy^n} e^{-\frac{y^2}{2}}, \quad \forall y \in \mathbb{R}.$$

Lemma 3.4.2. *Let $\alpha = (\alpha_i) \in \mathbb{N}_0^d$. Then,*

$$\|g_\alpha\|_{\mathcal{H}} = \sqrt{\alpha!} = \sqrt{\alpha_1! \cdots \alpha_d!}\,. \tag{3.4.3}$$

Proof. We compute

$$\|g_\alpha\|_{\mathcal{H}}^2 = \int_{\mathbb{R}^d} \prod_{i=1}^{d} H_{\alpha_i}(x_i)^2 g(x_i)^2 g(x_i)^{-1} dx = \prod_{i=1}^{d} \int_{\mathbb{R}} H_{\alpha_i}(x_i)^2 g(x_i) dx_i = \prod_{i=1}^{d} \alpha_i!\,,$$

where we have used the following weighted L^2-norm of H_n:

$$\int_{\mathbb{R}} H_n(y)^2 g(y)\, dy = n!. \tag{3.4.4}$$

□

Definition 3.4.3. We define the index sets $S^{(m)} := \{\alpha \in \mathbb{N}_0^d : |\alpha| = m\}$, $m \in \mathbb{N}_0$. For any $m \in \mathbb{N}_0$, the subspace $V^{(m)}$ of $\mathcal{H}$ is defined as

$$V^{(m)} := \operatorname{span}_{\mathbb{R}} \left\{ g_\alpha : \ \alpha \in S^{(m)} \right\}. \tag{3.4.5}$$

Remark 3.4.4. $V^{(m)}$ has dimension

$$\Gamma_m := |S^{(m)}| = \binom{d+m-1}{m} < \infty. \tag{3.4.6}$$

Let us consider some examples. If $d = 2$ we have

1. $V^{(0)} = \{\beta_1 g_0(x), \beta_1 \in \mathbb{R}\}$;
2. $V^{(1)} = \operatorname{span}\{g_{(1,0)}, g_{(0,1)}\} = \operatorname{span}\left\{x_1 e^{-|x|^2/2},\ x_2 e^{-|x|^2/2}\right\}$
$= \{(\beta_1 x_1 + \beta_2 x_2) g_0(x),\ \beta_1, \beta_2 \in \mathbb{R}\}$;
3. $V^{(2)} = \operatorname{span}\{g_{(2,0)}, g_{(1,1)}, g_{(0,2)}\}$
$= \left\{\left[\beta_1(x_1^2 - 1) + \beta_2 x_1 x_2 + \beta_3(x_2^2 - 1)\right] g_0(x),\ \beta_i \in \mathbb{R},\ i = 1,2,3\right\}$;
4. $V^{(3)} = \operatorname{span}\{g_{(3,0)}, g_{(2,1)}, g_{(1,2)}, g_{(0,3)}\}$
$= \left\{ \left[\beta_1(x_1^3 - 3x_1) + \beta_2(x_1^2 x_2 - x_2) + \beta_3(x_2^2 x_1 - x_1) + \beta_4(x_2^3 - 3x_2)\right] g_0(x),\right.$
$\left.\beta_1, \ldots, \beta_4 \in \mathbb{R}\right\}$.

It is well known that $\{g_\alpha\}_{\alpha\in\mathbb{N}_0^d}$ forms an orthogonal basis of $\mathcal{H} = L^2(\mathbb{R}^d, g_0^{-1})$. Hence, also the subspaces $V^{(m)}$ are mutually orthogonal. This yields an orthogonal decomposition of the Hilbert space

$$\mathcal{H} = \bigoplus_{m\in\mathbb{N}_0}^{\perp} V^{(m)}. \tag{3.4.7}$$

Remark 3.4.5. In [21, §5] an alternative block diagonal decomposition of the FP-propagator (when considered in the flat $L^2(\mathbb{R}^d)$) into finite-dimensional subspaces is derived by using *Wick quantization.*

We also consider the normalized version of the basis elements of the subspaces $V^{(m)}$:

Definition 3.4.6 (Normalized basis)**.** For each fixed $\alpha \in \mathbb{N}_0^d$, we denote with $\tilde{g}_\alpha$ the normalized function

$$\tilde{g}_\alpha := \frac{g_\alpha}{\|g_\alpha\|_{\mathcal{H}}}.$$

The reason why we need both g_α and $\tilde{g}_\alpha$ is that we can obtain a "nicer evolution" evolution of $f(t)$ projected into $V^{(m)}$ in terms of the matrix C with the first ones. Instead, the functions $\tilde{g}_\alpha$ can be used to express the equivalence of norms by Plancherel's equality in the Hilbert space $\mathcal{H}$.

The orthogonal decomposition (3.4.7) allows to express $f(t) \in L^2(\mathbb{R}^2, f_\infty^{-1})$, for a fixed $t \geq 0$, in the form

$$f(t,x) = \sum_{\alpha\in\mathbb{N}_0^d} \frac{\langle f(t), g_\alpha\rangle_{\mathcal{H}}}{\|g_\alpha\|_{\mathcal{H}}^2} g_\alpha(x) =: \sum_{\alpha\in\mathbb{N}_0^d} d_\alpha(t) g_\alpha(x), \tag{3.4.8}$$

or in terms of the normalized basis,

$$f(t,x) = \sum_{\alpha\in\mathbb{N}_0^d} \langle f(t), \tilde{g}_\alpha\rangle_{\mathcal{H}} \tilde{g}_\alpha(x) =: \sum_{\alpha\in\mathbb{N}_0^d} \tilde{d}_\alpha(t) \tilde{g}_\alpha(x). \tag{3.4.9}$$

The Fourier coefficients corresponding to a subspace $V^{(m)}$ can be grouped into vectors in $\mathbb{R}^{\Gamma_m}$:

$$d^{(m)} := (d_\alpha)_{\alpha\in S^{(m)}}, \text{ and } \tilde{d}^{(m)} := \left(\tilde{d}_\alpha\right)_{\alpha\in S^{(m)}}.$$

By the completeness of the Hilbert orthonormal basis $\{\tilde{g}_\alpha\}_{\alpha\in\mathbb{N}_0^d}$ in $\mathcal{H}$, Plancherel's Theorem then yields

$$\|f\|_{\mathcal{H}}^2 = \sum_{m\geq 0} \left\|\tilde{d}^{(m)}\right\|_2^2 = \sum_{m\geq 0}\sum_{\alpha\in S^{(m)}} |\tilde{d}_\alpha|^2 = \sum_{m\geq 0}\sum_{\alpha\in S^{(m)}} |d_\alpha|^2 \|g_\alpha\|_{\mathcal{H}}^2, \tag{3.4.10}$$

where we have used the relation $\tilde{d}_\alpha = \|g_\alpha\|_{\mathcal{H}} d_\alpha$.

Moreover, we denote by $(\Pi_m f) \in V^{(m)}$ the orthogonal *projection of f into $V^{(m)}$*. It is given by

$$(\Pi_m f) = \sum_{\alpha\in S^{(m)}} d_\alpha g_\alpha = \sum_{\alpha\in S^{(m)}} \tilde{d}_\alpha \tilde{g}_\alpha.$$

It follows that

$$\left\|\Pi_m f\right\|_{\mathcal{H}} = \left\|\tilde{d}^{(m)}\right\|_2. \tag{3.4.11}$$

In the next proposition we shall see that the subspaces $V^{(m)}$ are invariant under the action of the operator L, by giving the explicit action of L on each basis element g_α. For this purpose we introduce a notation for shifted multi-indices.

Definition 3.4.7. Given $\alpha = (\alpha_i) \in \mathbb{N}_0^d$ and $l \in \langle d \rangle := \{1, ..., d\}$, we define the components of the multi-indices $\alpha^{(l-)}, \alpha^{(l+)} \in \mathbb{N}_0^d$ as

$$\alpha_j^{(l\pm)} := \alpha_j \quad \text{for } j \neq l, \qquad \alpha_l^{(l\pm)} := (\alpha_l \pm 1)_+ .$$

So, for instance, if $g_\alpha \in V^{(m)}$ and $\alpha_l > 0$, then $g_{\alpha^{(l-)}} \in V^{(m-1)}$ and $g_{(\alpha^{(l-)})^{(j+)}} \in V^{(m)}$. Note that cutting off negative values guarantees that $\alpha^{(l-)}$ is always an admissible multi-index. This part of the definition will, however, not influence the following.

The action of the operator L on $V^{(m)}$ can be taken from [6, Proposition 5.1 and its proof]:

Proposition 3.4.8. *For every $m \in \mathbb{N}_0$, the subspace $V^{(m)}$ is invariant under L, its adjoint L^* and, hence, the solution operator e^{-Lt}, $t \geq 0$. Moreover, for each g_α,*

$$L g_\alpha = - \sum_{j,l=1}^{d} \alpha_l C_{jl} g_{(\alpha^{(l-)})^{(j+)}} , \tag{3.4.12}$$

where C_{jl} are the matrix elements of C.

3.4.2 Evolution of the Fourier coefficients

In this section we shall derive the evolution of $\Pi_m f$ in terms of the Fourier coefficients $d^{(m)}$:

Proposition 3.4.9. *Let f satisfy the FP-equation* (3.2.5)*. Then the coefficients in the expansion* (3.4.8) *satisfy*

$$\frac{d}{dt} d_\alpha = - \sum_{j,l=1}^{d} \mathbb{1}_{\alpha_j \geq 1} (\alpha^{(j-)})_l^{(l+)} C_{jl} d_{(\alpha^{(j-)})^{(l+)}} , \qquad \alpha \in \mathbb{N}_0^d . \tag{3.4.13}$$

Proof. We substitute (3.4.8) into (3.2.5) and use (3.4.12):

$$\sum_{\alpha \in \mathbb{N}_0^d} \frac{d}{dt} d_\alpha g_\alpha = - \sum_{j,l=1}^{d} \sum_{\alpha : \alpha_l \geq 1} d_\alpha \alpha_l C_{jl} g_{(\alpha^{(l-)})^{(j+)}} .$$

In the sum over α on the right hand side we substitute

$$(\alpha^{(l-)})^{(j+)} = \beta \quad \Longleftrightarrow \quad \alpha = (\beta^{(j-)})^{(l+)} ,$$

leading to

$$\begin{aligned} \sum_{\alpha \in \mathbb{N}_0^d} \frac{d}{dt} d_\alpha g_\alpha &= - \sum_{j,l=1}^{d} \sum_{\beta : \beta_j \geq 1} d_{(\beta^{(j-)})^{(l+)}} (\beta^{(j-)})_l^{(l+)} C_{jl} g_\beta \\ &= \sum_{\beta \in \mathbb{N}_0^d} \left(- \sum_{j,l=1}^{d} \mathbb{1}_{\beta_j \geq 1} (\beta^{(j-)})_l^{(l+)} C_{jl} d_{(\beta^{(j-)})^{(l+)}} \right) g_\beta , \end{aligned}$$

completing the proof. □

Remark 3.4.10. From the family of equations (3.4.13) we can deduce: The vector $d^{(m)} = (d_\alpha)_{\alpha\in S^{(m)}} \in \mathbb{R}^{\Gamma_m}$ satisfies the ODE $\frac{d}{dt}d^{(m)} = -C^{(m)}d^{(m)}$ for some matrix $C^{(m)} \in \mathbb{R}^{\Gamma_m\times\Gamma_m}$. Actually, we shall not write down the matrix $C^{(m)}$ explicitly, as we shall not need it.

As the simplest example we shall first consider the evolution in $V^{(1)}$. We use the notation $S^{(1)} = \{\alpha(1),\dots,\alpha(d)\}$ with $\alpha(k)_j = \delta_{jk}$, $j,k = 1,\dots,d$. In the right hand side of (3.4.13) with $\alpha = \alpha(k)$ obviously only the terms with $j = k$ are nonzero, $(\alpha(k)^{(k-)})^{(l+)} = \alpha(l)$ and, thus, $(\alpha(k)^{(k-)})_l^{(l+)} = 1$. This implies

$$\frac{d}{dt}d_\alpha = -\sum_{l=1}^{d} C_{kl}d_{\alpha(l)}$$

and therefore

$$\frac{d}{dt}d^{(1)} = -Cd^{(1)} \qquad \text{for } d^{(1)} = \left(d_{\alpha(1)},\dots,d_{\alpha(d)}\right). \tag{3.4.14}$$

We define $h(t) := \|e^{-Ct}\|_{\mathscr{B}(\mathbb{R}^d)}$. Then (3.4.14) implies

$$h(t) = \sup_{0\neq\tilde{d}^{(1)}(0)\in\mathbb{R}^{\Gamma_1}} \frac{\|\tilde{d}^{(1)}(t)\|_2}{\|\tilde{d}^{(1)}(0)\|_2}, \qquad t\geq 0. \tag{3.4.15}$$

To analyze the evolution in $V^{(m)}$, $m\geq 2$, it turns out that the representation of $d^{(m)}$ as a vector is not convenient. In the next section we shall rather represent it as a tensor. Not as a tensor of order d, as the number of components of α would indicate, but as a symmetric tensor of order m over $\mathbb{R}^d$. This way it will be easier to characterize its evolution – in fact as a tensored version of (3.4.14).

3.5 Subspace evolution in terms of tensors

3.5.1 Order-m tensors

In this subsection we briefly review some notations and basic results on tensors that will be needed. Most of their elementary proofs are deferred to §5.7. For more details we refer the reader to [13] and [22].

Let $m\in\mathbb{N}$ be fixed. We note that along the paper the convention $\mathbb{N} = \{1,2,...\}$, excluding zero, is used.

Definition 3.5.1. For $n_1,...,n_m\in\mathbb{N}$, a function $h : \langle n_1\rangle\times\cdots\times\langle n_m\rangle\to\mathbb{R}$ is a (real valued) *hypermatrix*, also called *order-m tensor* or *m-tensor*, where $\langle n_k\rangle := \{1,...,n_k\}$, $\forall 1\leq k\leq m$. We denote the set of values of h by an m-dimensional table of values, calling it $A = (A_{i_1\dots i_m})_{i_1,\dots,i_m=1}^{n_1,\dots,n_m}$, or just $A = (A_{i_1\dots i_m})$. The set of order-m hypermatrices (with domain $\langle n_1\rangle\times\cdots\times\langle n_m\rangle$) is denoted by $T^{n_1\times\cdots\times n_m}$.

We will consider only the case in which $n_1 = \cdots = n_m = d$, i.e., $A = (A_{i_1\dots i_m})_{i_1,\dots,i_m=1}^{d}$. In this case, we will denote $T_d^{(m)} := T^{d\times\cdots\times d}$ for simplicity. Also, since in our case the dimension d is fixed, we will denote it by $T^{(m)}$. Then $A\in T^{(m)}$ is a function from $\langle d\rangle^m$ to $\mathbb{R}$, denoted by $A = (A_I)_{I\in\langle d\rangle^m}$.

It will be useful to define some operations on $T_d^{(m)}$:

Definition 3.5.2. It is natural to define the operations of entrywise addition and scalar multiplication that make $T^{(m)}$ a vector space in the following way: for any $A, B \in T^{(m)}$ and $\gamma \in \mathbb{R}$

$$(A+B)_{i_1\dots i_m} := A_{i_1\dots i_m} + B_{i_1\dots i_m}, \quad (\gamma A)_{i_1\dots i_m} := \gamma A_{i_1\dots i_m}.$$

Moreover, given m matrices $B_1 = (b^{(1)}_{ij}), \dots, B_m = (b^{(m)}_{ij}) \in \mathbb{R}^{d\times d} = T^{(2)}$ and $A \in T^{(m)}$, we define the *multilinear matrix multiplication* by
$A' := (B_1, \dots, B_m) \odot A \in T^{(m)}$ where

$$A'_{i_1\dots i_m} := \sum_{j_1,\dots,j_m=1}^{d} b^{(1)}_{i_1 j_1} \cdots b^{(m)}_{i_m j_m} A_{j_1\dots j_m}. \tag{3.5.1}$$

For $A \in T^{(m)}$ and $k \le m$ matrices $B_1, \dots, B_k \in T^{(2)}$, we also define the product $A' := (B_1, \dots, B_k) \odot A \in T^{(m)}_d$ in the following way:

$$A'_{i_1\dots i_m} := \sum_{j_1,\dots,j_k=1}^{d} b^{(1)}_{i_1 j_1} \cdots b^{(k)}_{i_k j_k} A_{j_1\dots j_k i_{k+1}\dots i_m},$$

i.e., the multiplication acts on the first k-indices of A. For simplicity, when $B_1 = \dots = B_k := B$, we will denote $(B_1, \dots, B_k) \odot A$ by $B \odot^k A$. For example, if $d = 4$ and given $B = (b_{ij}) \in \mathbb{R}^{4\times 4}$, $A \in T^{(3)}$,

$$(B \odot A)_{i_1 i_2 i_3} = \sum_{j=1}^{4} b_{i_1 j} A_{j i_2 i_3},$$

and

$$B \odot^3 A = (B, B, B) \odot A.$$

Finally, we equip $T^{(m)}$ with an inner product:

Definition 3.5.3. Let $A = (A_{i_1\dots i_m}), B = (B_{i_1\dots i_m}) \in T^{(m)}$, we call $\langle A, B\rangle_{\mathscr{F}} \in \mathbb{R}$ the *Frobenius inner product* between the m-tensors A and B, defined by

$$\langle A, B\rangle_{\mathscr{F}} := \sum_{i_1,\dots,i_m=1}^{d} A_{i_1\dots i_m} B_{i_1\dots i_m}.$$

This induces a norm in $T^{(m)}$, called *Frobenius norm* in the natural way:

$$\|A\|_{\mathscr{F}} := \sqrt{\langle A, A\rangle_{\mathscr{F}}} = \left(\sum_{i_1,\dots,i_m=1}^{d} (A_{i_1\dots i_m})^2 \right)^{1/2} \ge 0.$$

Definition 3.5.4. The tensor $D = (D_I)_{I\in\langle d\rangle^m} \in T^{(m)}$ is called *symmetric*, if $\forall I \in \langle d\rangle^m$ it is true that $D_I = D_{\sigma(I)}$ for every permutation σ of m elements. Then $F^{(m)} \subset T^{(m)}$ (and occasionally $F^{(m)}_d$) denotes the set of symmetric m-tensors. Given $A \in T^{(m)}$, we define the *symmetric part of* A as the symmetric tensor defined by

$$\mathrm{Sym} A := \frac{1}{m!} \sum_{\sigma\in\mathscr{P}} \sigma(A) \in F^{(m)},$$

where $\mathscr{P}$ is the group of permutations of m elements and $\sigma(A)$ is the tensor with components $\sigma(A)_I := A_{\sigma(I)}$, $\forall I \in \langle d\rangle^m$.

Remark 3.5.5. For a symmetric tensor $D \in F^{(m)}$, clearly we do not need to define D_I for each $I = (i_1, ..., i_m) \in \langle d \rangle^m$ since the value of D_I depends only on the number of occurrences of each value in the index I. Therefore, we define the function $\varphi : \langle d \rangle^m \to S^{(m)}$ with

$$\varphi_k(I) := \sum_{j=1}^{m} \delta_{k,i_j}, \quad \forall k = 1, ..., d \quad \text{and for each} \quad I = (i_1, ..., i_m) \in \langle d \rangle^m \ ,$$

where $\delta_{k,i}$ denotes the Kronecker symbol. Hence, the component φ_k counts the occurrences of k in the multi-index I. Then, $\forall I \in \langle d \rangle^m$ we define the multi-index $\varphi(I) \in S^{(m)}$ as $\varphi(I) = (\varphi_1(I), ..., \varphi_d(I))$. We observe that $\varphi(I)$ is well defined, since $\sum_{k=1}^{d} \varphi_k(I) = m$, for any $I \in \langle d \rangle^m$.

For the computation of the Frobenius norm of a symmetric tensor it will be useful to introduce the following index classes:

Definition 3.5.6. For a fixed $I \in \langle d \rangle^m$ we define the *equivalence class of I under the action of φ* as

$$[I]_\varphi := \{J \in \langle d \rangle^m : \ \varphi(I) = \varphi(J)\} \ ,$$

and the set of classes

$$\langle d \rangle^m / \varphi := \{[I]_\varphi : \ I \in \langle d \rangle^m\} \ .$$

It is easy to show that there is a bijection between the quotient set $\langle d \rangle^m / \varphi$ and $S^{(m)}$ through the identification $[I]_\varphi \subset \langle d \rangle^m$ and $\alpha = \varphi(I)$, for each $\alpha \in S^{(m)}$. We observe that:

- If $\varphi(I) = \alpha = (\alpha_1, ..., \alpha_d)$, then $[I]_\varphi$ has exactly $\gamma_\alpha = \frac{m!}{\alpha_1! \cdots \alpha_d!}$ elements.
- If $D = (D_I)_{I \in \langle d \rangle^m}$ is symmetric, then $D_I = D_J$ if I and J are in the same class.

We will use these two properties in the proof of Proposition 3.5.18, for example to compute the Frobenius norm of a symmetric tensor.

Definition 3.5.7. Let $D = (D_I)$ be a symmetric m-tensor and $I \in \langle d \rangle^m$. Then, for any $\alpha = (\alpha_1, ..., \alpha_d) \in S^{(m)}$ we define

$$D_\alpha := D_I, \quad \text{if } \alpha = (\varphi_1(I), ..., \varphi_d(I)).$$

We observe that this notion is well-defined since D is symmetric and the property $\varphi(I) = \varphi(\sigma(I))$ holds.

The previous definition shows that φ induces a one-to-one correspondence between the indices of a symmetric m-tensor and the elements of $S^{(m)}$. This implies that the dimension of $F^{(m)}$ is equal to the cardinality of $S^{(m)}$, i.e. Γ_m (see (3.4.6)). Hence, for defining $D \in F^{(m)}$ we just need to define D_α for every $\alpha \in S^{(m)}$.

Next we define the order-m outer product and discuss the rank-1 decomposition of tensors, using a result from multilinear algebra ([13], Lemma 4.2).

Definition 3.5.8. Let $v_i := (v_1^{(i)}, ..., v_d^{(i)})$, $i = 1, ..., m$ be m vectors in $\mathbb{R}^d$. We define $v_1 \otimes \cdots \otimes v_m \in T^{(m)}$ as the m-tensor with components

$$(v_1 \otimes \cdots \otimes v_m)_I := v_{i_1}^{(1)} \cdots v_{i_m}^{(m)}, \quad \forall I = (i_1, ..., i_m) \in \langle d \rangle^m.$$

We call this operation between m vectors, *m-outer product.*

In the special case of all the vectors $v_i = v \in \mathbb{R}^d$, $i = 1, ..., m$ equal, we denote

$$v^{\otimes m} := v \otimes \cdots \otimes v,$$

and we observe that the tensor $v^{\otimes m}$ is symmetric by definition.

Proposition 3.5.9 ([13], Lemma 4.2)**.** *Let $D \in F_d^{(m)}$. Then, there exist an integer $s \in \langle \Gamma_m \rangle$, numbers $\lambda_1, ..., \lambda_s \in \mathbb{R}$, and vectors $v_1, ..., v_s \in \mathbb{R}^d$ such that*

$$D = \sum_{k=1}^{s} \lambda_k v_k^{\otimes m}. \tag{3.5.2}$$

The minimum s such that (3.5.2) *holds is called the symmetric rank of D.*

Remark 3.5.10. In [13] the result is stated for complex tensors. In that case it is possible to choose all the coefficients λ_i in (3.5.2) equal to one, due to the fact that $\mathbb{C}$ is a closed field. We remark that the same decomposition carries over to the real case, i.e. with real coefficients λ_i and real vectors v_i, by using the same proof [14].

It is easy to see that this rank-1 decomposition persists under a (constant) multilinear matrix multiplication:

Lemma 3.5.11. *Let $B \in \mathbb{R}^{d\times d}$. For any $D \in F_d^{(m)}$ decomposed as in formula* (3.5.2)*, the following decomposition holds:*

$$B \odot^m D = \sum_{k=1}^{s} \lambda_k (B v_k)^{\otimes^m}. \tag{3.5.3}$$

For rank-1 tensors, their inner product simplifies as follows:

Lemma 3.5.12. *Given $v_k = (v_i^{(k)}) \in \mathbb{R}^d$, $k = 1, ..., 2m$, then*

$$\langle v_1 \otimes \cdots \otimes v_m, v_{m+1} \otimes \cdots \otimes v_{2m} \rangle_{\mathscr{F}} = \prod_{i=1}^{m} \langle v_i, v_{i+m} \rangle, \tag{3.5.4}$$

where $\langle v_i, v_j \rangle$ is the inner product in $\mathbb{R}^d$.

A special case of this lemma is given by

Corollary 3.5.13. Given $v_1, v_2 \in \mathbb{R}^d$, then

$$\langle v_1^{\otimes^m}, v_2^{\otimes^m} \rangle_{\mathscr{F}} = \langle v_1, v_2 \rangle^m. \tag{3.5.5}$$

Next we shall derive some results on matrix-tensor products $B \odot^k A$:

Lemma 3.5.14. *Let $B = B^T \in \mathbb{R}^{d\times d}$ be such that $B \geq 0$. Then, for any $A \in T^{(m)}$*

$$\langle A, B \odot A \rangle_{\mathscr{F}} \geq 0. \tag{3.5.6}$$

For $B \in \mathbb{R}^{d\times d}$, $\|B\|$ we will denote in the sequel the *spectral norm of B.*

Lemma 3.5.15. *For any $A \in T_d^{(m)}$, $B \in \mathbb{R}^{d\times d}$ and $1 \leq k \leq m$,*

$$\|B \odot^k A\|_{\mathscr{F}} \leq \|B\|^k \|A\|_{\mathscr{F}}. \tag{3.5.7}$$

3.5.2 Time evolution of the tensors $D^{(m)}(t)$ in $V^{(m)}$

Proposition 3.4.9 gives the time evolution of each vector $d^{(m)}$. But for $m \geq 2$ it does not reveal its inherent structure. Therefore we shall now regroup the elements of $d^{(m)}$ as an order-m tensor and analyze its evolution.

Definition 3.5.16. Let $m \geq 1$, $t \geq 0$, and $d^{(m)}(t) = (d_\alpha(t))_{\alpha \in S^{(m)}} \in \mathbb{R}^{\Gamma_m}$ be the solution of the ODE $\frac{d}{dt} d^{(m)} = -C^{(m)} d^{(m)}$, with the matrix $C^{(m)}$ discussed in Remark 3.4.10. Then we define the symmetric m-tensor $D^{(m)}(t) = (D^{(m)}_\alpha(t))_{\alpha \in S^{(m)}}$ as

$$D^{(m)}_\alpha(t) := \frac{d_\alpha(t)}{\gamma_\alpha}, \tag{3.5.8}$$

where $\gamma_\alpha := \frac{m!}{\alpha!}$, for $\alpha = (\alpha_1, ..., \alpha_d)$.

For $m = 1$ we of course have $D^{(1)} = d^{(1)} = (d_\alpha)_{\alpha \in \langle d \rangle}$. We illustrate the above definition for the case $m = d = 2$ with $\Gamma_2 = 3$:

$$d^{(2)} = \begin{pmatrix} d_{(2,0)} \\ d_{(1,1)} \\ d_{(0,2)} \end{pmatrix}, \qquad D^{(2)} = \begin{pmatrix} d_{(2,0)} & \frac{d_{(1,1)}}{2} \\ \frac{d_{(1,1)}}{2} & d_{(0,2)} \end{pmatrix} \in F_2^{(2)} \subset T_2^{(2)} = \mathbb{R}^{2\times 2}.$$

Elementwise, the evolution of $D^{(m)}_\alpha$ easily carries over from Proposition 3.4.9:

Proposition 3.5.17. *For any $\alpha \in S^{(m)}$, the element $D^{(m)}_\alpha(t)$ evolves according to*

$$\frac{d}{dt} D^{(m)}_\alpha = -\sum_{j,l=1}^{d} \alpha_j C_{jl} D^{(m)}_{(\alpha^{(j-)})^{(l+)}}. \tag{3.5.9}$$

Proof. From (3.4.13) we obtain by substituting the definition (3.5.8) on both sides:

$$\frac{d}{dt} D^{(m)}_\alpha = -\frac{1}{\gamma_\alpha} \sum_{j,l=1}^{d} \mathbb{1}_{\alpha_j \geq 1} \gamma_{(\alpha^{(j-)})^{(l+)}} (\alpha^{(j-)})^{(l+)}_l C_{jl} D^{(m)}_{(\alpha^{(j-)})^{(l+)}}. \tag{3.5.10}$$

The claim (3.5.9) then follows from the relation

$$\gamma_\alpha \alpha_j = \gamma_{(\alpha^{(j-)})^{(l+)}} (\alpha^{(j-)})^{(l+)}_l \qquad \forall \alpha \in \mathbb{N}_0^d \text{ with } \alpha_j \geq 1, \tag{3.5.11}$$

which can be obtained as follows: It is trivial for $l = j$, and for $l \neq j$ it follows from the definition of γ_α and from the observation that $(\alpha^{(j-)})^{(l+)}_l = \alpha_l + 1$ and $(\alpha^{(j-)})^{(l+)}_j = \alpha_j - 1$. □

The advantage of this new structure consists in two facts:

- The Frobenius norm $\|D^{(m)}(t)\|_{\mathscr{F}}$ is proportional (uniformly in t) to the Euclidean norm $\left\|\tilde{d}^{(m)}(t)\right\|_2$ for which we want to prove a decay estimate like (3.4.15).
- The rank-1 decomposition of $D^{(m)}(t)$ is compatible with the Fokker-Planck flow in $V^{(m)}$, i.e., for each symmetric tensor $D^{(m)}(0)$ (considered as an initial condition in $V^{(m)}$), we can decompose $D^{(m)}(t)$ as a sum of order-m outer products of vectors that are solutions of the ODE $\frac{d}{dt} v(t) = -C v(t)$.

Concerning the first property we have

Proposition 3.5.18. *Given $m \geq 1$, then*

$$\left\| D^{(m)}(t) \right\|_{\mathscr{F}} = \frac{1}{\sqrt{m!}} \left\| \tilde{d}^{(m)}(t) \right\|_2 , \quad \forall t \geq 0. \tag{3.5.12}$$

Proof. We compute, using Remark 3.5.6,

$$\| D^{(m)}(t) \|_{\mathscr{F}}^2 = \sum_{I \in \langle d \rangle^m} D_I^{(m)}(t)^2 = \sum_{\alpha \in S^{(m)}} D_\alpha^{(m)}(t)^2 \gamma_\alpha,$$

where we used the identification $D_\alpha^{(m)}(t) := D_I^{(m)}(t)$ if $\alpha = \varphi(I)$ as well as $\left| [I]_\varphi \right| = \gamma_\alpha$.

Then, using the definition of $D^{(m)}(t)$, $\tilde{d}_\alpha(t) = \| g_\alpha \|_{\mathscr{H}} d_\alpha(t)$, and Lemma 3.4.2, we have

$$\begin{aligned} \left\| D^{(m)}(t) \right\|_{\mathscr{F}}^2 &= \sum_{\alpha \in S^{(m)}} \frac{d_\alpha(t)^2}{\gamma_\alpha} = \sum_{\alpha \in S^{(m)}} \frac{\tilde{d}_\alpha(t)^2}{\gamma_\alpha \| g_\alpha \|_{\mathscr{H}}^2} = \frac{1}{m!} \sum_{\alpha \in S^{(m)}} \tilde{d}_\alpha(t)^2 \\ &= \frac{1}{m!} \left\| \tilde{d}^{(m)}(t) \right\|_2^2 , \end{aligned}$$

concluding the proof. □

Concerning the second property we find that the rank-1 decomposition of $D^{(m)}(t)$ commutes with the time evolution by the Fokker-Planck equation:

Theorem 3.5.19. *Let $m \geq 1$ be fixed and let $D^{(m)} \in F^{(m)}$, having the rank-1 decomposition $D^{(m)} = \sum_{k=1}^s \lambda_k v_k^{\otimes m}$ with symmetric rank s, constants $\lambda_1, \ldots, \lambda_s \in \mathbb{R}$ and s vectors $v_k := (v_j^{(k)})_{j=1}^d \in \mathbb{R}^d$. Then, $D^{(m)}(t)$, $t > 0$, the solution to* (3.5.9) *with initial condition $D^{(m)}(0) = D^{(m)}$ has the decomposition*

$$D^{(m)}(t) = \sum_{k=1}^s \lambda_k [v_k(t)]^{\otimes m}, \tag{3.5.13}$$

where all vectors $v_k(t) \in \mathbb{R}^d$, $k = 1, \ldots, s$ satisfy the ODE $\frac{d}{dt} v_k(t) = -C v_k(t)$ with initial condition $v_k(0) = v_k$. Moreover, $D^{(m)}(t)$, $t > 0$ has the constant-in-t symmetric rank s.

Proof. We shall compute the evolution of the symmetric m-tensor $A(t) := \sum_{k=1}^s \lambda_k [v_k(t)]^{\otimes m}$, using that $\frac{d}{dt} v_k(t) = -C v_k(t)$. To this end we compute first the derivative $\frac{d}{dt} (w(t)^{\otimes m})_\alpha$ if the vector $w(t) = (w_1(t), \ldots, w_d(t))^T \in \mathbb{R}^d$ satisfies the ODE $\frac{d}{dt} w(t) = -C w(t)$.

Given $\alpha = (\alpha_1, ..., \alpha_d) \in S^{(m)}$, we have

$$\begin{aligned}
\frac{d}{dt}(w(t)^{\otimes m})_\alpha &= \frac{d}{dt}\prod_{j=1}^{d} w_j(t)^{\alpha_j} \\
&= \sum_{j=1}^{d} \alpha_j \left(w_1(t)^{\alpha_1}\cdots w_j(t)^{\alpha_j-1}\cdots w_d(t)^{\alpha_d}\right)\left(\frac{d}{dt}w_j(t)\right) \\
&= -\sum_{j=1}^{d} \alpha_j \left(w_1(t)^{\alpha_1}\cdots w_j(t)^{\alpha_j-1}\cdots w_d(t)^{\alpha_d}\right)\sum_{l=1}^{d} C_{jl} w_l(t) \\
&= -\sum_{j,l=1}^{d} \alpha_j C_{jl}\left(w_1(t)^{\alpha_1}\cdots w_j(t)^{\alpha_j-1}\cdots w_l(t)^{\alpha_l+1}\cdots w_d(t)^{\alpha_d}\right) \\
&= -\sum_{j,l=1}^{d} \alpha_j C_{jl}\left(w(t)^{\otimes m}\right)_{(\alpha^{(j-)})^{(l+)}},
\end{aligned}$$

and hence, by linearity

$$\frac{d}{dt}(A(t))_\alpha = -\sum_{j,l=1}^{d} \alpha_j C_{jl}\,(A(t))_{(\alpha^{(j-)})^{(l+)}}. \tag{3.5.14}$$

This ODE equals the evolution equation (3.5.9) for $D^{(m)}$, and hence $A(t) = D^{(m)}(t)$ follows. Next we consider the symmetric rank of $D^{(m)}(t)$, $t > 0$. If it would be smaller than s, a reversed evolution to $t = 0$ would lead to a contradiction to the symmetric rank of $D^{(m)}$. □

This theorem allows to reduce the evolution of the tensors $D^{(m)}(t)$ to the ODE for the vectors $v_k(t)$. This will be a key ingredient for proving sharp decay estimates of $D^{(m)}$ in the next section. Moreover it provides a compact formula for the evolution of $D^{(m)}(t)$.

Corollary 3.5.20. Let $m \geq 1$ be fixed. Then, $D^{(m)}(t)$, t>0, the solution to (3.5.9) follows the evolution

$$\frac{d}{dt}D^{(m)}(t) = -m\,\mathrm{Sym}(C \odot D^{(m)}(t)), \quad t > 0. \tag{3.5.15}$$

Proof. We shall use the decomposition (3.5.13) for $D^{(m)}(t)$. First, we compute the evolution of $[v(t)]^{\otimes m}$, if $\frac{d}{dt}v(t) = -Cv(t)$:

$$\begin{aligned}
\frac{d}{dt}([v(t)]^{\otimes m}) &= -\sum_{k=0}^{m-1}[v(t)]^{\otimes k}\otimes((Cv(t))\otimes[v(t)]^{\otimes(m-k-1)} \\
&= -m\,\mathrm{Sym}\big((Cv(t))\otimes[v(t)]^{\otimes(m-1)}\big).
\end{aligned}$$

In the last equality we have used, with $w := Cv(t)$, the general formula

$$\mathrm{Sym}(w\otimes v^{\otimes(m-1)}) = \frac{1}{m}\sum_{k=0}^{m-1}(v^{\otimes k}\otimes w\otimes v^{\otimes(m-k-1)}), \quad \forall v, w \in \mathbb{R}^d$$

that can be proven with a straightforward computation. By using the linearity of Sym in $T^{(m)}$, we obtain

$$\frac{d}{dt}D^{(m)}(t) = \frac{d}{dt}\sum_{k=1}^{s}\lambda_k[v_k(t)]^{\otimes m} = -m\left(\sum_{k=1}^{s}\lambda_k \mathrm{Sym}\big((Cv_k(t))\otimes[v_k(t)]^{\otimes(m-1)}\big)\right)$$
$$= -m\,\mathrm{Sym}\left(\sum_{k=1}^{s}\lambda_k(Cv_k(t))\otimes[v_k(t)]^{\otimes(m-1)}\right) = -m\,\mathrm{Sym}(C\odot D^{(m)}(t)).$$

□

3.6 Decay of the subspace evolution in $V^{(m)}$

First we shall rewrite our main decay result, Theorem 3.3.4 in terms of tensors for all subspaces $V^{(m)}$. We recall $h(t) := \|e^{-Ct}\|_{\mathscr{B}(\mathbb{R}^d)}$, which satisfies

$$h(t)\le 1\,,\qquad t\ge 0\,. \tag{3.6.1}$$

This follows from

$$\frac{d}{dt}\left\|e^{-Ct}x_0\right\|_2^2 = -2\langle C_S x(t), x(t)\rangle \le 0\,,\qquad x_0\in\mathbb{R}^d,$$

for $x(t) = e^{-Ct}x_0$. Using Theorem 3.3.14, the statement of (3.6.1) can be improved immediately to

$$h(t) < 1,\qquad t>0. \tag{3.6.2}$$

We have shown in (3.4.15) that the inequality (3.6.8), see below, holds with $m=1$, since $D^{(1)}(t) = d^{(1)}(t)$ satisfies the evolution $\dot{d}^{(1)} = -Cd^{(1)}$. Next we extend the estimate (3.6.8) to general $m\ge 1$. To this end we will show in the next theorem that the propagator norm in each $V^{(m)}$ is the m-th power of the propagator norm of the ODE $\dot{x} = -Cx$. This will be used to derive the decay estimates for $\|e^{-Lt}\|_{\mathscr{B}(V_0^\perp)}$.

Theorem 3.6.1. *For each $m\ge 1$, $D^{(m)}(0)\in F^{(m)}$, and $D^{(m)}(t)$ defined as in* (3.5.8)*, the following estimate holds:*

$$\left\|D^{(m)}(t)\right\|_{\mathscr{F}} \le h(t)^m\left\|D^{(m)}(0)\right\|_{\mathscr{F}}\,,\qquad t\ge 0\,. \tag{3.6.3}$$

Moreover,

$$\sup_{0\ne D^{(m)}(0)\in F^{(m)}}\frac{\|D^{(m)}(t)\|_{\mathscr{F}}}{\|D^{(m)}(0)\|_{\mathscr{F}}} = h(t)^m. \tag{3.6.4}$$

Proof. Given the initial condition $D^{(m)}(0)\in F^{(m)}$, Theorem 3.5.19 provides its rank-1 decomposition as

$$D^{(m)}(t) = \sum_{k=1}^{s}\lambda_k[v_k(t)]^{\otimes m} = \sum_{k=1}^{s}\lambda_k[e^{-Ct}v_k]^{\otimes m} = e^{-Ct}\odot^m D^{(m)}(0),\quad \forall t\ge 0, \tag{3.6.5}$$

with $v_k(t) = e^{-Ct}v_k$, for $k=1,\dots,s$, where we have used Lemma 3.5.11 in the last equality. Using (3.5.7) then yields:

$$\|D^{(m)}(t)\|_{\mathscr{F}} = \|e^{-Ct}\odot^m D^{(m)}(0)\|_{\mathscr{F}} \le \|e^{-Ct}\|^m\|D^{(m)}(0)\|_{\mathscr{F}}, \tag{3.6.6}$$

proving (3.6.3).

In order to prove the equality (3.6.4) we choose initial data of the form $D^{(m)}(0) := v^{\otimes m}$, $v \in \mathbb{R}^d$. In this case the Frobenius norm factorizes, i.e. $\|D^{(m)}(0)\|_{\mathscr{F}} = \|v\|_2^m$ and

$$\|D^{(m)}(t)\|_{\mathscr{F}} = \|(e^{-Ct}v)^{\otimes m}\|_{\mathscr{F}} = \|e^{-Ct}v\|_2^m$$

We conclude by observing that

$$\sup_{0\neq v\in\mathbb{R}^d} \frac{\|e^{-Ct}v\|_2^m}{\|v\|_2^m} = h(t)^m.$$

□

The key step in the above proof is to write the evolution of the tensor $D^{(m)}(t)$ as in (3.6.5), which allows for the simple estimate (3.6.6). In contrast, using the rank-1 decomposition in $\|D^{(m)}(t)\|_{\mathscr{F}}^2$ would not be helpful, since the vectors $v_k(t)$ are in general not orthogonal.

We conclude this chapter with the proof of our main result, Theorem 3.3.4, by using Theorem 3.6.1.

Proof of Theorem 3.3.4. The first step consists in proving the inequality

$$\left\|e^{-Lt}\right\|_{\mathscr{B}(V_0^\perp)} \leq h(t), \forall t \geq 0. \tag{3.6.7}$$

We can derive the estimate (3.6.7) from the same ones that hold for the tensors $D^{(m)}(t)$ at each level m. More precisely, (3.6.7) holds if

$$\|D^{(m)}(t)\|_{\mathscr{F}} \leq h(t)\|D^{(m)}(0)\|_{\mathscr{F}}, \qquad t \geq 0, \quad D^{(m)}(0) \in F^{(m)}, \quad m \geq 1, \tag{3.6.8}$$

where $D^{(m)}(t)$ is defined as in (3.5.8). Indeed,

$$\|f(t) - f_\infty\|_{\mathscr{H}}^2 = \sum_{m\geq 1} \|\Pi_m f(t)\|_{\mathscr{H}}^2 = \sum_{m\geq 1} \|\tilde{d}^{(m)}(t)\|_2^2 = \sum_{m\geq 1} m!\, \|D^{(m)}(t)\|_{\mathscr{F}}^2, \quad t \geq 0, \tag{3.6.9}$$

where we have used the orthonormal decomposition of $f(t)$, formulas (3.4.10), (3.5.12), and that the coefficient $d_0(t) \equiv 1$, (with the index $0 \in \mathbb{N}_0^d$), is constant in time since $Lg_0 = 0$ and the normalization $\int_{\mathbb{R}^d} f_0\, dx = 1$. Let us assume (3.6.8). Then,

$$\begin{aligned}\|f(t) - f_\infty\|_{\mathscr{H}}^2 &= \sum_{m\geq 1} m!\, \|D^{(m)}(t)\|_{\mathscr{F}}^2 \leq h(t)^2 \sum_{m\geq 1} m!\, \|D^{(m)}(0)\|_{\mathscr{F}}^2 \\ &= h(t)^2 \|f_0 - f_\infty\|_{\mathscr{H}}^2,\end{aligned}$$

proving (3.6.7).

Next, the proof of (3.6.8) is a direct consequence of Theorem 3.6.1 and $h(t) \leq 1$, yielding

$$\|D^{(m)}(t)\|_{\mathscr{F}} \leq (h(t))^m \|D^{(m)}(0)\|_{\mathscr{F}} \leq h(t)\|D^{(m)}(0)\|_{\mathscr{F}}.$$

Now that (3.6.7) has been proved, we need to show that it is actually an equality, in order to conclude the proof of (3.3.1). For this purpose, we observe that for $m = 1$, $D^{(1)} \in \mathbb{R}^d$ evolves according to the ODE $\dot{x} = -Cx$ (see (3.4.14)). Then, it is sufficient to choose an initial datum $f_0 \in V^{(1)}$ to achieve the equality, concluding the proof. □

Remark 3.6.2. Using (3.6.2), the decay estimates (3.6.3) show that the higher subspace components $D^{(m)}(t)$ decay, for each fixed $t > 0$, with a rate that increases exponentially in m. Due to the subspace decomposition (3.6.9), this enhanced decay of the higher subspace components translates into a parabolic-type regularization of the FP-semigroup for $t > 0$, cp. to Proposition 3.3.19.

3.7 Second quantization

In this last section we are going to write the FP-operator L in (3.2.5) in terms of the second quantization formalism. This "language" was introduced in quantum mechanics in order to simplify the description and the analysis of quantum many-body systems. The assumption of this construction is the indistinguishability of particles in quantum mechanics. Indeed, according to the statistics of particles, the exchange of two of them does not affect the status of the configuration, possibly up to a sign. Since we are dealing with symmetric tensors, we are going to consider the case in which the sign does not change, i.e. the wave function is identical after this exchange. This is the case of particles that are called *bosons*.

The functional spaces of second quantization are the so-called *Fock spaces*, that we are going to define in this section. When a single Hilbert space H describes a single particle, then it is convenient to build an infinite sum of symmetric tensorization of H in order to represent a system of (up to) infinitely many indistinguishable particles, i.e. the Fock space over H.

In the first part of this section the definitions of the Boson Fock space and second quantization operators are given. These constructions will be needed in order to write the FP-operator L as the second quantization of its corresponding drift matrix C. This will be the main result of the second part of this section as an application of well known results in the literature.

3.7.1 The Boson Fock space

In the next definition we will use the notion of m-fold tensor product over a Hilbert space H. This is a generalization of the space of order-m hypermatrices $T^{(m)}$ defined in §5, where the Hilbert space was the finite dimensional space $\mathbb{R}^d$. In the quantum mechanics literature, the role of the Hilbert space is often played by $L^2(\mathbb{R}^3;\mathbb{C})$, in order to describe the wave function of a quantum particle. For a more complete explanation of tensor products of Hilbert spaces and Fock spaces we refer to §II.4 in [30].

In the literature, Fock spaces are mostly considered for Hilbert spaces over the field $\mathbb{C}$. But since the FP-equations (3.1.1) and (3.2.5) are posed on $\mathbb{R}^d$ (and not over $\mathbb{C}^d$), we shall use here only real valued Fock spaces. Moreover, these FP-equations are considered here only for real valued initial data, and hence real valued solutions.

Definition 3.7.1. Let H be a Hilbert space and denote by $H^{(m)} := H \otimes H \otimes \cdots \otimes H$ (m times), for any $m \geq 1$. Set $H^{(0)} := \mathbb{C}$ (or $\mathbb{R}$) and define the *Fock space over* H as the completed direct sum

$$\mathscr{F}(H) = \bigoplus_{m=0}^{\infty} H^{(m)}. \tag{3.7.1}$$

Then, an element $\psi \in \mathscr{F}(H)$ can be represented as a sequence $\psi = \{\psi^{(m)}\}_{m=0}^{\infty}$, where $\psi^{(0)} \in \mathbb{C}$ (or $\mathbb{R}$), $\psi^{(m)} \in H^{(m)}, \forall m \geq 1$, so that

$$\|\psi\|_{\mathscr{F}(H)} := \sqrt{\sum_{m=0}^{\infty} \|\psi^{(m)}\|^2_{H^{(m)}}} < \infty. \tag{3.7.2}$$

Here $\|\cdot\|_{H^{(m)}}$ denotes the norm induced by the inner product in $H^{(m)}$ (see Proposition 1, §II.4 in [30]).

As we anticipated, we will rather work with a subspace of $\mathscr{F}(H)$, the so-called Boson Fock space that we are going to define. First we need to define the m-fold symmetric tensor product of H as follows:

Let $\mathscr{P}_m$ be the permutation group on m elements and let $\{\phi_k\}$; $k = 1, ..., \dim H$, be a basis for H. For each $\sigma \in \mathscr{P}_m$, we define its corresponding operator (we will still denote it with σ) acting on basis elements of $H^{(m)}$ by

$$\sigma(\phi_{k_1} \otimes \phi_{k_2} \otimes \cdots \otimes \phi_{k_m}) := \phi_{\sigma(k_1)} \otimes \phi_{\sigma(k_2)} \otimes \cdots \otimes \phi_{\sigma(k_m)}. \tag{3.7.3}$$

Then σ extends by linearity to a bounded operator on $H^{(m)}$. With the previous definition (3.7.3) we can define the operator $S_m := \frac{1}{m!}\sum_{\sigma \in \mathscr{P}_m} \sigma$ that acts on $H^{(m)}$. Its range $S_m H^{(m)}$ is called the *m-fold symmetric tensor product of* H. Let us see examples of $S_m H^{(m)}$.

Example 3.7.2. Let us consider first the case $H = L^2(\mathbb{R})$ and $H^{(m)} = L^2(\mathbb{R}) \otimes \cdots \otimes L^2(\mathbb{R})$. Since $H^{(m)}$ is isomorphic to $L^2(\mathbb{R}^m)$, it follows that an element $\psi^{(m)} \in S_m H^{(m)}$ is a function $\psi^{(m)}(x_1, ..., x_m)$ in $L^2(\mathbb{R}^m)$ left invariant under any permutation of the variables. It is used in quantum mechanics to describe the quantum states of m particles that are not distinguishable.

For our purposes, we will deal with $H = \mathbb{R}^d$. In this case it is easy to check that $S_m H^{(m)}$ corresponds to the space of symmetric m-tensors $F^{(m)}$ that we defined in §5, equipped with the Frobenius norm. □

Definition 3.7.3. The subspace of $\mathscr{F}(H)$,

$$\mathscr{F}_s(H) := \bigoplus_{m=0}^{\infty} S_m H^{(m)} \tag{3.7.4}$$

is called the *symmetric Fock space over* H or the *Boson Fock space over* H.

3.7.2 The second quantization operator

In order to write the FP-propagator in terms of the second quantization formalism, we need to define the second quantization operators (see §I.4 in [33] and §X.7 in [31]) acting on the Boson Fock space.

Let H be a Hilbert space and $\mathscr{F}_s(H)$ be the Boson Fock space over H. Let A be a contraction on H, i.e., a linear transform of norm smaller than or equal to 1. Then there is a unique contraction (Corollary I.15, [33]) $\Gamma(A)$ on $\mathscr{F}_s(H)$ so that

$$\Gamma(A) \restriction_{S_m H^{(m)}} = A \otimes \cdots \otimes A \qquad (m \text{ times}), \tag{3.7.5}$$

where the operator $A\otimes\cdots\otimes A$ is defined on each basis element $\psi^{(m)}=\psi_{i_1}\otimes\cdots\otimes\psi_{i_m}$ of $S_mH^{(m)}$ as

$$(A\otimes\cdots\otimes A)(\psi^{(m)}):=(A\psi_{i_1})\otimes\cdots\otimes(A\psi_{i_m}),$$

and equal to the identity when restricted to $H^{(0)}$. In order to prove the above existence of $\Gamma(A)$, the estimate $\|\Gamma(A)\restriction_{S_mH^{(m)}}\|\leq\|A\|^m$ is first showed in [33]. This allows to extend the operator $\Gamma(A)$ to the Boson Fock space by continuity, and by remaining a contraction. In the case $A=e^{-Ct}$ and $H=\mathbb{R}^d$, the operator $\Gamma(A)$ will be useful to show the link between the Fokker-Planck solution operator e^{-Lt} and the second quantization operators, defined in the following way:

Definition 3.7.4. Let H be a Hilbert space. Let A be an operator on H (with domain $G(A)$). The operator $d\Gamma(A)$ is defined as follows: Let $G_m(A)\subseteq S_mH^{(m)}$ be $G(A)\otimes\cdots\otimes G(A)$ and $G(d\Gamma(A)):=\dotplus_{m=0}^{\infty}G_m(A)$ (incomplete direct sum):

$$d\Gamma(A)\restriction_{S_mH^{(m)}}:=A\otimes\mathbb{1}\otimes\cdots\otimes\mathbb{1}+\cdots+\mathbb{1}\otimes\cdots\otimes\mathbb{1}\otimes A,\quad m\geq1,\tag{3.7.6}$$

and $d\Gamma(A)\restriction_{H^{(0)}}:=0$. The operator $d\Gamma(A)$ is called the *second quantization of A*.

In [33] the following property of the second quantization operator can be found (see I.41):

Let A generate a $\mathscr{C}_0$-contraction semigroup on H. Then the closure of $d\Gamma(A)$ generates a $\mathscr{C}_0$-contraction semigroup on $\mathscr{F}_s(H)$ and

$$e^{-d\Gamma(A)t}=\Gamma(e^{-At})\qquad\forall t\geq0.\tag{3.7.7}$$

3.7.3 Application to the operator e^{-Lt}

In the last part of this section we will show that the Fokker-Planck operator L is the second quantization of C. First, we shall identify the Hilbert space $L^2(\mathbb{R}^d,f_\infty^{-1})$ with a suitable Fock space.

The spectral decomposition and the tensor structure that we introduced in §5 suggest to consider the Boson Fock space over the finite dimensional Hilbert space $\mathbb{R}^d$, whose elements have components in the space of symmetric tensors $F^{(m)}$. Indeed, we can define an isomorphism Ψ between $L^2(\mathbb{R}^d,f_\infty^{-1})$ and $\mathscr{F}_s(\mathbb{R}^d)$ as follows:

Let $f\in L^2(\mathbb{R}^d,f_\infty^{-1})$. As we saw in §4, f admits the decomposition $f(x)=\sum_{m=0}^{\infty}\sum_{\alpha\in S^{(m)}}d_\alpha g_\alpha(x)$, for suitable coefficients $d_\alpha\in\mathbb{R}$. For each $m\geq1$, we define the symmetric tensor $\widetilde{D}^{(m)}\in F^{(m)}$ with components $\widetilde{D}^{(m)}_\alpha:=d_\alpha\frac{\sqrt{m!}}{\gamma_\alpha}\in\mathbb{R}$ (see (3.5.8)), $\forall\alpha\in S^{(m)}$. For $m=0$ we choose $\widetilde{D}^{(0)}:=\langle f,f_\infty\rangle_{L^2(f_\infty^{-1})}$. Hence, by observing that $F^{(m)}=S_mH^{(m)}$, $H:=\mathbb{R}^d$, we define the isometry

$$\Psi:f\in L^2(\mathbb{R}^d,f_\infty^{-1})\to\psi:=\{\widetilde{D}^{(m)}\}_{m=0}^{\infty}\in\mathscr{F}_s(\mathbb{R}^d).\tag{3.7.8}$$

It remains to check that $\|\psi\|_{\mathscr{F}_s(\mathbb{R}^d)}<\infty$. This follows from the Plancherel's equality together with (3.5.12). It leads to

$$\|f\|^2_{L^2(f_\infty^{-1})}=\sum_{m=0}^{\infty}\|\widetilde{D}^{(m)}\|^2_{\mathscr{F}}=\|\psi\|^2_{\mathscr{F}_s(\mathbb{R}^d)}.$$

Hence, up to an isomorphism, we can consider the FP-operator L also as acting on the Fock space $\mathscr{F}_s(\mathbb{R}^d)$. We conclude the section with the next proposition that allows to write L in the second quantization formalism.

Proposition 3.7.5. *Let L be the Fokker-Planck operator defined in* (3.2.5) *and let $C \in \mathbb{R}^{d\times d}$ be its corresponding drift matrix. Then, L, now considered as acting on $\mathscr{F}_s(\mathbb{R}^d)$, is the second quantization of C, considered as an operator from the Hilbert space $\mathbb{R}^d$ to itself, i.e., $L = d\Gamma(C)$.*

Proof. Due to the relation (3.7.7), it is sufficient to prove that the FP-propagator e^{-Lt} (considered on $\mathscr{F}_s(\mathbb{R}^d)$) satisfies the equality

$$e^{-Lt} = \Gamma(e^{-Ct}), \qquad \forall t \ge 0. \tag{3.7.9}$$

Equivalently, on each $S_m H^{(m)}$, $m \ge 1$, the formula

$$e^{-Lt}(\psi^{(m)}) = (e^{-Ct}\psi_{i_1}) \otimes \cdots \otimes (e^{-Ct}\psi_{i_m}), \tag{3.7.10}$$

holds for every basis element $\psi^{(m)} = \bigotimes_{k=1}^m \psi_{i_k}$ of $F^{(m)}$.

Given an initial condition $f_0 \in L^2(\mathbb{R}^d, f_\infty^{-1})$ and its corresponding solution $f(t) = e^{-Lt} f_0$ of (3.2.5), the isometry $\Psi\colon L^2(\mathbb{R}^d, f_\infty^{-1}) \to \mathscr{F}_s(H)$ maps then as follows:

$$\Psi f_0 = \psi_0 = \{\widetilde{D}^{(m)}(0)\}_{m=0}^\infty, \quad \text{and} \quad \Psi f(t) = \psi(t) = \{\widetilde{D}(t)^{(m)}\}_{m=0}^\infty,$$

respectively. Then, the factored evolution formula (3.6.5) for $D^{(m)}(t) = \sqrt{m!}\,\widetilde{D}^{(m)}(t)$ proves the equality (3.7.10), for each $m \ge 1$. Since the generator of a $\mathscr{C}_0$-semigroup is unique, we obtain $L = d\Gamma(C)$. □

While C is a bounded operator with domain $G(C) = \mathbb{R}^d$, its second quantization $d\Gamma(C)$ is unbounded with dense domain $G(d\Gamma(C)) \subsetneq \mathscr{F}_s(H)$, just like L is unbounded on $L^2(\mathbb{R}^d, f_\infty^{-1})$.

Finally, our main result, Theorem 3.3.4 reads in the language of second quantization

$$\left\| e^{-d\Gamma(C)t}\restriction_{\bigoplus_{m\ge1} S_m H^{(m)}} \right\|_{\mathscr{B}(\mathscr{F}_s(H))} = \|e^{-Ct}\|_{\mathbb{R}^{d\times d}}, \qquad t \ge 0. \tag{3.7.11}$$

Note that the restriction to $\bigoplus_{m\ge1} S_m H^{(m)}$ corresponds to the restriction to $V_0^\perp$ in (3.3.1), the orthogonal of the steady state f_∞.

We remark that Proposition 3.7.5 is a special case of Theorem 1 in [12], there formulated for an infinite dimensional Hilbert space setting. We still include a proof here to make this paper self-contained. Moreover, an explicit computation of the spectrum and second quantization formalism for FP-equations in the infinite dimensional setting were given in [35].

Remark 3.7.6. Many aspects of the above analysis seem to rely importantly on the explicit spectral decomposition of the FP-operator in §4.1, i.e. knowing the FP-eigenfunctions (as Hermite functions). We remark that this situation in fact carries over to FP-equations with linear coefficients plus a *nonlocal perturbation* of the form $\theta_f := \theta * f$ with the function $\theta(x)$ having zero mean, see Lemma 3.8 and Theorem 4.6 in [10]. For such nonlocally perturbed FP-equations, surprisingly, one still knows all the eigenfunctions as well as its (multi-dimensional) creation and annihilation operators.

3.8 Appendix: Deferred proofs

Proof of Lemma 3.5.11. We compute the components of the l.h.s. of (3.5.3). Using (3.5.2) with $v_k = (v_i^{(k)}) \in \mathbb{R}^d$, we have for any $(i_1, .., i_m) \in \langle d \rangle^m$:

$$(B \odot^m D)_{i_1 \dots i_m} = \sum_{j_1,\dots,j_m=1}^{d} B_{i_1 j_1} \cdots B_{i_m j_m} D_{j_1 \dots j_m} = \sum_{j_1,\dots,j_m=1}^{d} B_{i_1 j_1} \cdots B_{i_m j_m} \sum_{k=1}^{s} \lambda_k v_{j_1}^{(k)} \cdots v_{j_m}^{(k)}$$
$$= \sum_{k=1}^{s} \lambda_k (B v_k)_{i_1} \cdots (B v_k)_{i_m} = \left(\sum_{k=1}^{s} \lambda_k (B v_k)^{\otimes^m} \right)_{i_1 \cdots i_m},$$

concluding the proof. □

Proof of Lemma 3.5.12. By definition,

$$\langle v_1 \otimes \cdots \otimes v_m, v_{m+1} \otimes \cdots \otimes v_{2m} \rangle_{\mathscr{F}} = \sum_{i_1,\dots,i_m=1}^{d} (v_1 \otimes \cdots \otimes v_m)_{i_1 \dots i_m} (v_{m+1} \otimes \cdots \otimes v_{2m})_{i_1 \dots i_m}$$
$$= \sum_{i_1,\dots,i_m=1}^{d} v_{i_1}^{(1)} \cdots v_{i_m}^{(m)} v_{i_1}^{(m+1)} \cdots v_{i_m}^{(2m)}$$
$$= \left(\sum_{i_1=1}^{d} v_{i_1}^{(1)} v_{i_1}^{(m+1)} \right) \cdots \left(\sum_{i_m=1}^{d} v_{i_m}^{(m)} v_{i_m}^{(2m)} \right)$$
$$= \langle v_1, v_{m+1} \rangle \cdots \langle v_m, v_{2m} \rangle.$$

□

Proof of Lemma 3.5.14. We have

$$\langle A, B \odot A \rangle_{\mathscr{F}} = \sum_{i_1,\dots,i_m=1}^{d} A_{i_1 \dots i_m} (B \odot A)_{i_1 \dots i_m} = \sum_{j_1, i_1,\dots,i_m=1}^{d} A_{i_1 \dots i_m} B_{i_1 j_1} A_{j_1 i_2 \dots i_m}$$
$$= \sum_{i_2,\dots,i_m=1}^{d} \langle x^{(i_2 \dots i_m)}, B x^{(i_2 \dots i_m)} \rangle,$$

where, for $i_2, ..., i_m$ fixed, $x_{i_1}^{(i_2 \dots i_m)} := A_{i_1 i_2 \dots i_m}$ are vectors in $\mathbb{R}^d$. The claim then follows from $B \geq 0$. □

Proof of Lemma 3.5.15. First consider the *Case* $k = 1$. We have

$$\|B \odot A\|_{\mathscr{F}}^2 = \sum_{i_1,\dots,i_m=1}^{d} (\sum_{j_1=1}^{d} B_{i_1 j_1} A_{j_1 i_2 \dots i_m})^2 = \sum_{i_2,\dots,i_m=1}^{d} \|B x^{(i_2 \dots i_m)}\|^2 \tag{3.8.1}$$
$$\leq \sum_{i_2,\dots,i_m=1}^{d} \|B\|^2 \|x^{(i_2 \dots i_m)}\|^2 = \|B\|^2 \sum_{i_1,\dots,i_m=1}^{d} (x_{i_1}^{(i_2 \dots i_m)})^2 \tag{3.8.2}$$
$$= \|B\|^2 \|A\|_{\mathscr{F}}^2 \tag{3.8.3}$$

where, for $i_2, ..., i_m$ fixed, $x_{j_1}^{(i_2 \dots i_m)} := A_{j_1 i_2 \dots i_m}$ are vectors in $\mathbb{R}^d$. Note that the estimate (3.8.1) would hold as well if the matrix-tensor product does not operate on the first index (as in $B \odot A$), but on the $j-$th index, with some $1 \le j \le m$. Then (3.5.7) follows by iterated applications of (3.8.1). □

Proof of Proposition 3.3.20. (a) We recall that Theorem 3.3.4 and (3.6.1) imply

$$\tilde{h}(t) = \|e^{-Lt}\|_{\mathscr{B}(V_0^\perp)} = \|e^{-Ct}\|_2 = h(t) \le 1, \quad t \ge 0.$$

Then, Theorem 3.6.1 implies (3.6.3), $\forall m \ge 1$. From (3.4.10) we recall

$$\left\|\frac{f(t)}{f_\infty}\right\|^2_{L^2(f_\infty)} = \|f(t)\|^2_{\mathscr{H}} = \sum_{m\in\mathbb{N}_0} \|\tilde{d}^{(m)}(t)\|^2 = \sum_{\beta\in\mathbb{N}_0^d} |\tilde{d}_\beta(t)|^2, \tag{3.8.4}$$

and $\frac{f(t)}{f_\infty} = \sum_{\beta\in\mathbb{N}_0^d} \tilde{d}_\beta(t)\hat{g}_\beta$, where $\hat{g}_\beta := \frac{\tilde{g}_\beta}{f_\infty}$ is an orthonormal basis of $L^2(f_\infty)$.

Using (3.4.2) and the formula $H_n'(x) = nH_{n-1}(x)$ for Hermite polynomials we compute, for any $\beta \in \mathbb{N}_0^d$,

$$\partial_{x_j}\hat{g}_\beta = \frac{\beta_j H_{\beta_j-1}(x_j)}{\sqrt{\beta!}} \prod_{i\ne j} H_{\beta_i}(x_i), \text{ and } \|\partial_{x_j}\hat{g}_\beta\|_{L^2(f_\infty)} = \sqrt{\beta_j},$$

where we used $\|H_n\|_{L^2(f_\infty)} = \sqrt{n!}$. This yields, with (3.6.3) and (3.5.12),

$$\begin{aligned}\left\|\nabla\left(\frac{f(t)}{f_\infty}\right)\right\|^2_{L^2(f_\infty)} &= \sum_{\beta\in\mathbb{N}_0^d} |\tilde{d}_\beta(t)|^2|\beta| = \sum_{m\in\mathbb{N}_0} m\|\tilde{d}^{(m)}(t)\|^2 \\ &\le \sum_{m\in\mathbb{N}_0} m(\tilde{h}(t))^{2m}\|\tilde{d}^{(m)}(0)\|^2, \quad t > 0.\end{aligned} \tag{3.8.5}$$

From the hypothesis on $\tilde{h}$, we deduce $\tilde{h}(t) \le 1 - c_1 t^\alpha$ on $0 \le t \le \delta$ for some $0 < c_1 \le c$ and some $\delta > 0$. Then (3.8.5) can be estimated further by

$$\sum_{m\in\mathbb{N}_0} m(1-c_1t^\alpha)^{2m}\|\tilde{d}^{(m)}(0)\|^2 \le \frac{1}{ec_1}t^{-\alpha} \sum_{m\in\mathbb{N}_0} \|\tilde{d}^{(m)}(0)\|^2, \quad 0 \le c_1 t^\alpha \le 1.$$

where we used the elementary inequality $m(1-c_1t^\alpha)^{2m} \le \frac{1}{ec_1}t^{-\alpha}$, $m \in \mathbb{N}_0$. The main assertion of part (a) then follows from (3.8.4).

Finally we turn to the optimality of α: If (3.3.21) would hold for all $f_0 \in \mathscr{H}$ with some $\alpha_1 \in (0,\alpha)$, then part (b) of this proposition would imply $\tilde{h}(t) \le 1 - c_2 t^{\alpha_1}$. But this would contradict the assumption $\tilde{h}(t) = 1 - ct^\alpha + o(t^\alpha)$. Hence, $\alpha/2$ is indeed the minimal regularization exponent in (3.3.21).

(b) For $f_0 \in V^{(m)}$, $m \in \mathbb{N}$ we compute, by using (3.8.5) and (3.3.21),

$$\left\|\nabla\left(\frac{f(t)}{f_\infty}\right)\right\|^2_{L^2(f_\infty)} = m\,\|\tilde{d}^{(m)}(t)\|^2 \le \tilde{c}^2 t^{-\alpha}\|\tilde{d}^{(m)}(0)\|^2, \qquad 0 < t \le \delta. \tag{3.8.6}$$

Then, by taking in (3.8.6) the supremum w.r.t. the set $\{0 \ne \tilde{d}^{(m)}(0) \in \mathbb{R}^{\Gamma_m}\}$ and using (3.6.4), (3.5.12) we obtain the family of estimates

$$\tilde{h}(t)^{2m} = \sup_{0\ne D^{(m)}\in F^{(m)}} \frac{\|D^{(m)}(t)\|^2_{\mathscr{F}}}{\|D^{(m)}\|^2_{\mathscr{F}}} = \sup_{0\ne\tilde{d}^{(m)}(0)\in\mathbb{R}^{\Gamma_m}} \frac{\|\tilde{d}^{(m)}(t)\|^2}{\|\tilde{d}^{(m)}(0)\|^2} \le \frac{\tilde{c}^2}{m}t^{-\alpha}, \tag{3.8.7}$$

with $m \in \mathbb{N}$, $0 < t \le \delta$.

Next we will show that this family of estimates for $\tilde{h}(t)$ implies $\tilde{h}(t) \le 1 - c_2 t^\alpha$ for $0 \le t \le \delta_2$, with some $c_2 > 0$, $\delta_2 > 0$ (see Figure 3.2 for the case $\alpha = 1$). For each $m \in \mathbb{N}$ and $t \in I_\delta := (0, \delta]$, we rewrite (3.8.7) as

$$\tilde{h}(t) \le \left(\frac{\tilde{c}}{\sqrt{m}} t^{-\frac{\alpha}{2}} \right)^{\frac{1}{m}} = e^{-\frac{1}{2} \frac{\log(\bar{c} m t^\alpha)}{m}} =: g(m;t), \tag{3.8.8}$$

with $\bar{c} := \tilde{c}^{-2}$. For $t \in I_\delta$ fixed, we now consider the function $g(\mu;t)$ with continuous argument $\mu > 0$. $g(\cdot;t)$ has its unique minimum at $\mu_0(t) := \frac{e}{\bar{c}} t^{-\alpha}$ and it is strictly decreasing on $(0, \mu_0(t))$.

To estimate the minimum of g for the discrete argument $m \in \mathbb{N}$, we consider: For $0 \le t \le t_1 := \left(\frac{e-2}{\bar{c}}\right)^{1/\alpha}$ we have

$$\frac{2}{\bar{c}} t^{-\alpha} \le \left\lceil \frac{2}{\bar{c}} t^{-\alpha} \right\rceil < \frac{2}{\bar{c}} t^{-\alpha} + 1 \le \frac{e}{\bar{c}} t^{-\alpha} = \mu_0(t),$$

with $\lceil \cdot \rceil$ denoting the ceiling function. We choose the index $m(t) := \lceil \frac{2}{\bar{c}} t^{-\alpha} \rceil \in \mathbb{N}$ and use the monotonicity of $g(\cdot;t)$ on $(0, \mu_0(t)]$ to estimate:

$$\tilde{h}(t) \le \min_{m \in \mathbb{N}} g(m;t) \le g(m(t);t) \le g\left(\frac{2}{\bar{c}} t^{-\alpha}; t\right) = e^{-2c_2 t^\alpha},$$

with $c_2 := \frac{\log(2)\bar{c}}{8} > 0$.

With the elementary estimate $e^{-2c_2 y} \le 1 - c_2 y$ on some $[0, t_2]$, we obtain

$$\tilde{h}(t) \le e^{-2c_2 t^\alpha} \le 1 - c_2 t^\alpha, \qquad t \in [0, \delta_2],$$

with $\delta_2 := \min\{t_1, t_2^{1/\alpha}\}$.

Finally we turn to the minimality of α: If $\tilde{h}$ would even satisfy the decay estimate $\tilde{h}(t) \le 1 - \tilde{c}_2 t^{\alpha_1}$ with some $\alpha_1 \in (0, \alpha)$ and $\tilde{c}_2 > 0$, then (the proof of) part (a) of this proposition would imply the regularization estimate (3.3.21) with the exponent $\alpha_1/2$. But this would contradict the assumption on α being minimal in that estimate. □

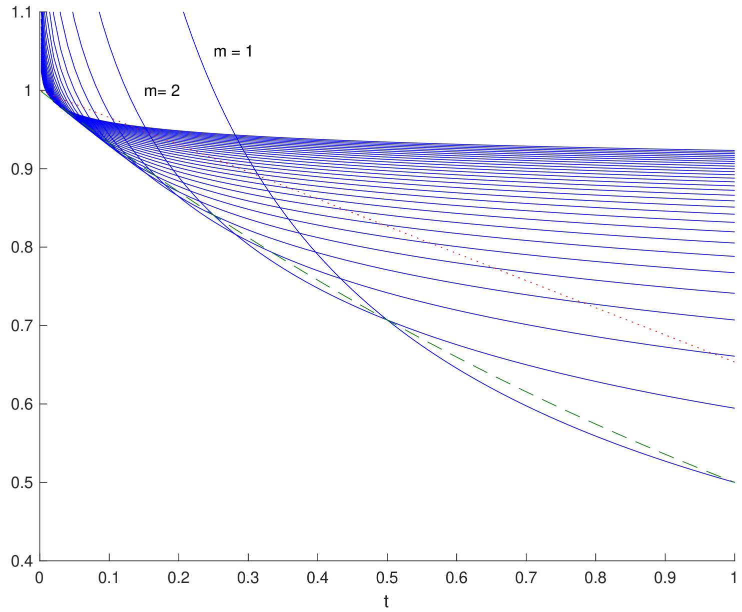

Figure 3.2: The family of decay estimates $h(t) \le g(m;t)$, $m \in \mathbb{N}$ with $\alpha = 1$, $\bar{c} = 4$ (solid, blue curves) implies $h(t) \le e^{-2c_2 t}$, (dashed, green curve), and hence $h(t) \le 1 - c_2 t$ (dotted, red line).

Bibliography

[1] F. Achleitner, A. Arnold, E. Carlen, *On multi-dimensional hypocoercive BGK models,* Kinetic and related models 11, no. 4 (2018), 953-1009.

[2] F. Achleitner, A. Arnold, E. Carlen, *The Hypocoercivity Index for the short and large time behaviour of ODEs*, https://arxiv.org/abs/2109.10784 (2021).

[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 282 (2019), 241-264, G. Giacomin et al. (eds.), Springer.

[4] P. Alphonse, J. Bernier, *Polar decomposition of semigroups generated by non-selfadjoint quadratic differential operators and regularizing effects,* Preprint. https://arxiv.org/abs/1909.03662v2 (2020)

[5] A. Arnold, A. Einav , T. Wöhrer, *On the rates of decay to equilibrium in degenerate and defective Fokker-Planck equations*, J. Differential Equations 264, no. 11 (2018), 6843-6872.

[6] 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.

[7] A. Arnold, P. A. Markowich, G. Toscani, A. Unterreiter, *On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations.* Comm. PDE 26, no. 1-2 (2001), 43-100.

[8] A. Arnold, S. Jin, T. Wöhrer, *Sharp Decay Estimates in Local Sensitivity Analysis for Evolution Equations with Uncertainties: from ODEs to Linear Kinetic Equations,* Journal of Differential Equations 268, no. 3 (2019), 1156-1204.

[9] A.Arnold, B.Signorello, *Optimal non-symmetric Fokker-Planck equation for the convergence to a given equilibrium.* Preprint.https://arxiv.org/abs/2106.15742 (2021).

[10] A. Arnold, D. Stürzer, *Spectral analysis and long-time behaviour of a Fokker- Planck equation with a non-local perturbation,* Rend. Lincei Mat. Appl. 25 (2014), 53-89. Erratum: Rend. Lincei Mat. Appl. 27 (2016), 147-149.

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

[12] A. Chojnowska-Michalik, B. Goldys, *Nonsymmetric Ornstein-Uhlenbeck semigroup as second quantized operator*, J. Math. Kyoto Univ. 36, no. 3 (1996), 481–498.

[13] P. Comon, G. Golub, L.-H. Lim, B. Mourrain, *Symmetric tensors and symmetric tensor rank,* SIAM J. Matrix Anal. Appl. 30, no. 3 (2008), 1254-1279.

[14] P. Comon, B. Mourrain, private communication, 26.9.2019.

[15] J. Dolbeault, X. Li, *Phi-entropies for Fokker-Planck and kinetic Fokker-Planck equations,* Math. Mod. and Meth. in Appl. Sci. 28 (2018), 2637-2666.

[16] S. Friedland, M. Stawiska, *Best Approximation on Semi-algebraic Sets and k-Border Rank Approximation of Symmetric Tensors*, preprint, arXiv:1311.1561 [math.AG], (2013).

[17] S. Gadat, L. Miclo, *Spectral decompositions and L2-operator norms of toy hypocoercive semi-groups*, Kinetic and Related Models 2 (2013), 317-372.

[18] M. Herda, L.M. Rodrigues, *Large-time behaviour of solutions to Vlasov-Poisson-Fokker-Planck equations: from evanescent collisions to diffusive limit*, J. Stat. Phys. 170 (2018), 895-931.

[19] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math. 119 (1967), 147-171.

[20] S. Kawashima, *L*arge-time behaviour of solutions to hyperbolic-parabolic systems of conservation laws and applications, Proc. Roy. Soc. Edinburgh Sect. A 106 (1987), 169-194.

[21] T. Lelièvre, F. Nier, G.A. Pavliotis, *Optimal Non-reversible Linear Drift for the Convergence to Equilibrium of a Diffusion*, Springer Science+Business Media New York (2013).

[22] L.-H. Lim, *Tensors and hypermatrices*, Chapter 15, 30 pp., in L. Hogben (Ed.), Handbook of Linear Algebra, 2nd Ed., CRC Press, Boca Raton, FL (2013).

[23] G. Metafune, D. Pallara, E. Priola, *Spectrum of Ornstein-Uhlenbeck operators in L^p spaces with respect to invariant measures*, J. Funct. Anal. 196, no. 1 (2002), 40-60.

[24] L. Miclo, P. Monmarché, *Étude spectrale minutieuse de processus moins indécis que les autres. (French) [Detailed spectral study of processes that are less indecisive than others]* Séminaire de Probabilités XLV, Lecture Notes in Math. 2078, Springer, Cham (2013), 459-481.

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

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

[27] L. Pareschi, G. Russo, G. Toscani, *Fast spectral methods for the Fokker-Planck-Landau collision operator*, J. Comput. Phys. 165, no.1 (2000), 216-236.

[28] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer Verlag (1983).

[29] L. Perko, *Differential Equations and Dynamical Systems*, Texts in Applied Mathematics 7, Springer Verlag (1991).

[30] M. Reed, B. Simon, *Methods of modern mathematical physics, Vol. 1,* Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London (1980).

[31] M. Reed, B. Simon, *Methods of modern mathematical physics, Vol. 2,* Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London (1975).

[32] C. Schmeiser, *Entropy methods,* https://homepage.univie.ac.at/christian.schmeiser/Entropy-course.pdf.

[33] B. Simon, *The* $P(\Phi)_2$ *Euclidean (Quantum) Field Theory,* Princeton University Press, Princeton (1974).

[34] B. Shizgal, *Spectral methods in chemistry and physics. Applications to kinetic theory and quantum mechanics. Scientific Computation*, Springer, Dordrecht (2015), xviii+415 pp.

[35] J.M. van Neerven, *Second quantization and the* L^p*-spectrum of nonsymmetric Ornstein-Uhlenbeck operators*, Infinite Dimensional Analysis, Quantum Probability and Related Topics 8, no. 3 (2005), 473–495.

[36] C. Villani, *Hypocoercivity*, Memoirs of the American Mathematical Society 202 (2009).

4 Optimal non-symmetric Fokker-Planck Equation

4.1 Introduction

The starting point of this paper is a linear, symmetric (or "reversible") Fokker-Planck (FP) equation on $\mathbb{R}^d$, $d \geq 2$ with a corresponding, typically anisotropic Gaussian steady state. It is known from the literature (see [11, 8], e.g.) that the convergence to equilibrium can be accelerated by adding to the FP-equation non-symmetric perturbations that do not alter the equilibrium. It is hence a natural goal to find the "optimal perturbation" (in a sense to be made precise) such that the corresponding solutions converge the fastest to the fixed steady state. For FP-equations with fixed or variable diffusion matrices, this problem was studied, respectively, in [11] and [8]. A closely related question for the (kinetic) 1D Goldstein-Taylor system was recently studied in [7]: For a fixed (anti-symmetric) transport operator, the authors found the best (symmetric) relaxation operator, yielding the fastest exponential decay to equilibrium. For the same model, but with constant-in-x relaxation, the propagator norm was previously computed in [9].

While we shall analyse this problem here in a pure PDE context, the origin of the question comes from a statistical and probabilistic setting: Let a given potential $V : \mathbb{R}^d \to \mathbb{R}$ satisfy $\int_{\mathbb{R}^d} e^{-V(x)} dx = 1$, and define the probability density function

$$f_{\infty,V}(x) := e^{-V(x)}. \tag{4.1.1}$$

To compute expectations with respect to the associated probability measure μ_V, e.g. via a Markov chain Monte Carlo algorithm (see [6]), one needs to construct an ergodic Markov process $(X_t)_{t\geq 0}$ with the unique invariant law μ_V, i.e.

$$\text{law}\,(X_t) \to \mu_V, \qquad \text{as } t \to \infty. \tag{4.1.2}$$

The efficiency of such algorithms can be measured by the speed of convergence in (4.1.2). This motivates to pursue the following objective: find the fastest among all possible processes that sample from the same equilibrium μ_V. A classical way to sample from μ_V is to consider a standard Brownian motion with drift $-\nabla V$.

The probability density function f_t of the process X_t at time t then solves the Fokker-Planck equation

$$\partial_t f_t = \text{div}_x(\nabla_x f_t + \nabla_x V f_t) =: -L f_t, \quad t > 0. \tag{4.1.3}$$

It is symmetric in the sense that its generator L is symmetric in the Hilbert space

$$\mathcal{H} := L^2(\mathbb{R}^d, f_{\infty,V}^{-1}) = \left\{ f : \mathbb{R}^d \to \mathbb{R}, \text{ measurable s.t.} \int_{\mathbb{R}^d} f(x)^2 f_{\infty,V}^{-1}(x)\, dx < \infty \right\}.$$

Under appropriate assumptions on the potential V (e.g. if $\frac{1}{2}|\nabla V(x)|^2 - \Delta V(x) \to +\infty$ for $|x| \to \infty$, see [13], A.19) it is possible to show that f_t converges to the unique equilibrium $f_{\infty,V}$. Moreover, L is coercive in $V_0^\perp$ with $V_0 := \text{span}_{\mathbb{R}}\{f_{\infty,V}\} \subset \mathcal{H}$, i.e. $\exists \lambda > 0$ such that

$$\langle Lf, f\rangle_{\mathcal{H}} \geq \lambda ||f||^2_{\mathcal{H}}, \qquad f \in V_0^\perp. \tag{4.1.4}$$

We shall assume in the sequel that this λ is chosen as large as possible, i.e. as the spectral gap of L. As a consequence, if f_t is a solution of (4.1.3), then

$$||f_t - f_{\infty,V}||_{\mathcal{H}} \leq e^{-\lambda t}||f_0 - f_{\infty,V}||_{\mathcal{H}}, \tag{4.1.5}$$

for any normalized initial condition $f_0 \in \mathcal{H}$ (see Proposition 9 in [13]). So we have a purely exponential convergence estimate.

We shall discuss in the next section that it is often possible to improve the rate of convergence towards $f_{\infty,V}$ by adding a non-reversible perturbation in (4.1.3) while preserving the steady state $f_{\infty,V}$ (as done in [8] and [11]). As a first step we consider the non-symmetric FP-equation

$$\partial_t f_t = \text{div}_x(\nabla_x f_t + (\nabla_x V + b) f_t) =: -L_b f_t, \tag{4.1.6}$$

with $b = b(x)$ such that $\text{div}_x(be^{-V}) = 0$, to keep the steady state condition $L_b f_{\infty,V} = 0$ still valid. In this paper we will only consider Fokker-Planck equations with linear drift, just as in [11, 8]. This corresponds to quadratic potentials

$$V(x) = \frac{x^T K^{-1} x}{2}, \; K \in \mathcal{S}^{>0} \tag{4.1.7}$$

and linear perturbations of the form $b = Ax$, $A \in \mathcal{M}$.

Notation: Here and in the sequel we denote with $\mathcal{M}$ the set of real $d \times d$ matrices, $\mathcal{S}^{>0}$ (resp. $\mathcal{S}^{\geq 0}$) the set of positive definite (resp. positive semi-definite) symmetric matrices, and with $\mathcal{A}$ the set of anti-symmetric matrices. The spectrum of $A \in \mathcal{M}$ is denoted by $\sigma(A)$. For a symmetric matrix A, $\lambda_{min}(A)$ and $\lambda_{max}(A)$ denote, respectively, its smallest and largest eigenvalue.

The following lemma (Lemma 1 in [11]) characterizes explicitly the admissible perturbations in (4.1.6).

Lemma 4.1.1. *Let $V(x)$ be given by* (4.1.7) *and let $b(x) = Ax$ where $A \in \mathcal{M}$. Then*

$$\text{div}_x(be^{-V}) = 0 \textit{ if and only if } A = JK^{-1} \textit{ with some } J \in \mathcal{A}. \tag{4.1.8}$$

Then the non-symmetric Fokker-Planck equation (4.1.6) becomes

$$\partial_t f_t = \text{div}_x(\nabla_x f_t + (I_d + J)K^{-1} x f_t), \tag{4.1.9}$$

where I_d denotes the identity matrix in $\mathcal{M}$, $J \in \mathcal{A}$ is arbitrary, and $f_{\infty,V}$ is still a steady state.

Note that (4.1.9) still satisfies (4.1.5) with the same rate λ (see §2.4 in [3]), but λ may be smaller than the spectral gap of L_b. However, the sharp decay rate can be recovered by hypocoercivity tools [13, 5]: Then one finds constants $\tilde{\lambda} > 0$ and $c \geq 1$ (depending on the fixed potential V and the matrix J) such that

$$||f_t - f_{\infty,V}||_{\mathcal{H}} \leq c e^{-\tilde{\lambda} t} ||f_0 - f_{\infty,V}||_{\mathcal{H}}, \quad \forall t \geq 0. \tag{4.1.10}$$

For the reversible FP-dynamics (with $b = 0$) the maximal decay rate $\tilde{\lambda}$ in the estimate (4.1.10) is λ, the biggest coercivity constant in the inequality (4.1.4). In this case, the multiplicative constant $c = 1$. The advantage of adding a non-reversible perturbation b is to possibly obtain a larger decay rate $\tilde{\lambda} > \lambda$, at the price of allowing for a multiplicative constant $c > 1$. In fact, the decay rate may be improved iff K is not a multiple of I_d, see §3.2, [11].

The question discussed in [11] is the following: Given the potential (4.1.7), which is the optimal non-reversible linear FP-equation of the form (4.1.9) (and with *t*ime-independent coefficients) such that its solutions converge to $f_{\infty,V}$ with largest decay rate? For the diffusion matrix fixed as I_d, as in (4.1.9), the authors give a complete answer in [11], Theorem 1. But if one generalizes the question, allowing to vary both b and the diffusion matrix, as done in [8], the best decay rate from [11] can be improved further. Finally, one can extend the question further and analyze if *t*ime-dependent coefficients can enhance the decay of linear FP-equations even more.

Let us put this paper more into context with the literature on entropy methods and hypocoercivity: The main goal of [1, 2, 3, 4, 5, 13] is to find explicit and sometimes even optimal decay rates for *a* given evolution equation. By contrast, the novelty in [11, 8] and here is to *f*ix an equilibrium density and then to *s*eek the evolution equation (within a certain class) that yields the fastest convergence towards the equilibrium.

This paper is organized as follows: In the next section we formulate this optimization problem and review the results from [8]. In §4.3 we present the main result: As the biggest improvement compared to [8], we shall be able to obtain multiplicative constants in (4.1.10) that are arbitrarily close to 1. In §4.4.1, §4.4.2 we will elucidate this result on 2D examples, giving sharp decay estimates and numerical illustrations. Moreover, we identify the non-symmetric perturbation of the FP-equation as a highly rotating drift term. Then, in §4.4.3 we discuss the issue of using time-dependent coefficient matrices to accelerate the decay behaviour, mostly focussing on a numerical case study in 2D. Finally, we conclude in §4.5.

4.2 Formulation of the optimization problem and existing results

Let $K \in \mathscr{S}^{>0}$ be given. We define the (typically) anisotropic Gaussian

$$f_{\infty,K}(x) := \frac{\det(K)^{-1/2}}{(2\pi)^{d/2}} \exp\left(-\frac{x^T K^{-1} x}{2}\right), \qquad x \in \mathbb{R}^d, \tag{4.2.1}$$

and the linear Fokker-Planck equation

$$\partial_t f_t = -L_{C,D} f_t := \operatorname{div}_x(D \nabla_x f_t + C x f_t), \qquad x \in \mathbb{R}^d,\ t \in (0,\infty), \tag{4.2.2}$$

for arbitrary x-independent matrices $D \in \mathscr{S}^{\geq 0}$ and $C \in \mathscr{M}$. Equation (4.2.2) is a generalization of the non-reversible (4.1.9), possibly with a degenerate (i.e. singular) diffusion matrix D. Moreover we define the set

$$\mathscr{I}(K) := \{(C, D) \in \mathscr{M} \times \mathscr{S}^{\geq 0},\ \operatorname{Tr}(D) \leq d : L_{C,D} f_{\infty,K} = 0\}. \tag{4.2.3}$$

The next lemma (Lemma 3.1 in [8]; for $D = I_d$ also Lemma 1 in [11]) gives a characterization of the pairs (C, D) in $\mathscr{I}(K)$.

Lemma 4.2.1. *For $K \in \mathscr{S}^{>0}$ fixed, the following two statements are equivalent:*

- $(C, D) \in \mathscr{I}(K)$;
- $D \in \mathscr{S}^{\geq 0}, \mathrm{Tr}(D) \leq d$, *and* $\exists J \in \mathscr{A}$ *such that* $C = (D + J)K^{-1}$.

In other words, for $K \in \mathscr{S}^{>0}$ given, we have

$$\mathscr{I}(K) = \{((D + J)K^{-1}, D) \in \mathscr{M} \times \mathscr{S}^{\geq 0} : \; J \in \mathscr{A}, \mathrm{Tr}(D) \leq d\}, \tag{4.2.4}$$

and $\mathscr{I}(K)$ is not empty.

Given a fixed covariance matrix $K \in \mathscr{S}^{>0}$ (and hence the fixed Gaussian $f_{\infty,K}$), the set $\mathscr{I}(K)$ represents the matrix pairs (C, D) such that their associated FP-equation admits $f_{\infty,K}$ as a normalized steady state. But reversely, for such a FP-equation, the (normalized) steady state $f_{\infty,K}$ does not have to be unique (e.g. $C = D = \mathrm{diag}(1, 0)$ admits (4.2.1) with any $K = \mathrm{diag}(1, \kappa)$, $\kappa > 0$). It is known from the literature (see for example Theorem 3.1, [5]) that the existence of a unique L^1-normalized steady state $f_{\infty,K}$ for (4.2.2) is equivalent to the following two conditions on $(C, D) \in \mathscr{M} \times \mathscr{S}^{\geq 0}$:

1. C is positive stable (i.e., C has a positive spectral gap $\rho(C) := \min\{\mathrm{Re}(\lambda) : \lambda \in \sigma(C)\}$);
2. hypoellipticity of (4.2.2) (i.e., there is no non-trivial C^T-invariant subspace of $\ker(D)$).

For our set-up, hypoellipticity can actually be inferred from $(C, D) \in \mathscr{I}(K)$; and more precisely:

Lemma 4.2.2. *For some fixed $K \in \mathscr{S}^{>0}$, let $(C, D) \in \mathscr{I}(K)$ and $\rho(C) > 0$. Then the corresponding FP-equation* (4.2.2) *is hypoelliptic.*

Proof. Normalized steady states of (4.2.2) are Gaussian with its covariance matrix Q satisfying the continuous Lyapunov equation

$$2D = CQ + QC^T \,. \tag{4.2.5}$$

Since $D \geq 0$ and $\rho(C) > 0$, (4.2.5) has a unique, symmetric and positive semi-definite solution Q (see, e.g., Theorem 2.2 in [12]), namely $Q = K$.

By the above mentioned equivalence to the uniqueness of the steady state, (4.2.2) is hypoelliptic. □

For each fixed steady state $f_{\infty,K}$ we now want to answer the following questions:

(Q1) Which FP-evolution(s) converge(s) the fastest, i.e. with largest decay rate λ_{opt} to the steady state in the operator norm of $e^{-L_{C,D}t}$ on $V_0^{\perp} \subset \mathscr{H} := L^2(\mathbb{R}^d, f_{\infty,K}^{-1})$?

(Q2) Second, when the best decay rate is fixed, what is the infimum of the multiplicative constant, c_{inf}, in the decay estimate (4.1.10)?

(Q3) Third, for a fixed $K \in \mathscr{S}^{>0}$ and the corresponding λ_{opt}, and for any $c > c_{inf}$, which pair(s) of matrices $(C_{opt}, D_{opt}) \in \mathscr{M} \times \in \mathscr{S}^{\geq 0}$ are such that $e^{-L_{C_{opt},D_{opt}}t}$ yields the convergence estimate (4.1.10) with the constants (λ_{opt}, c)?

(Q4) Forth, for such an optimal pair of matrices, what bound on C_{opt} can be found, and how does this bound grow w.r.t. to the space dimension d?

(Q5) Could something be gained by allowing for time-dependent matrices $C(t)$, $D(t)$?

Remark 4.2.3. We note that, without the additional constraint $\mathrm{Tr}(D) \le d$ in the definition of $\mathscr{I}(K)$, the problem of finding an optimal evolution in the above sense would be ill-posed: Indeed, if f_t converges to $f_{\infty,K}$ as $t \to \infty$, then $f_t^\alpha := f_{\alpha t}$, for any $\alpha > 0$ and pertaining to $(\alpha C, \alpha D)$, has the same equilibrium and converges α times faster to it. For this reason, we shall only consider diffusion matrices with a prescribed bound for the trace, as in [8]. In probabilistic language it corresponds to the requirement that the upper bound on the total amount of randomness simultaneously injected in the system is prescribed, and this bound is equal to the case $D = I_d$.

Next we shall optimize the decay rate within the family of FP-equations (4.2.2) satisfying $(C, D) \in \mathscr{I}(K)$. But our choice of matrix C will, in general, differ from the one constructed in [8]. We base this optimization on the fact that the sharp exponential decay rate of the FP-equation (4.2.2) equals $\rho(C)$ (at least for C diagonalizable, see [5], e.g.). Actually, (4.2.2) and its associated drift ODE, i.e. $\frac{d}{dt}x = -\widetilde{C}x$, with $\widetilde{C} := K^{-1/2}CK^{1/2}$ (and hence $\rho(C) = \rho(\widetilde{C})$), have an even closer connection, as proven in Theorem 2.3, [4]:

Theorem 4.2.4. *Let $K \in \mathscr{S}^{>0}$ be given. We consider a FP-equation* (4.2.2) *with* $(C, D) \in \mathscr{I}(K)$ *and C positive stable. Then, the propagator norms of* (4.2.2) *and of its corresponding drift ODE $\frac{d}{dt}x = -\widetilde{C}x$ are equal, i.e.*

$$\left\| e^{-L_{C,D}t} \right\|_{\mathscr{B}(V_0^\perp)} = \left\| e^{-\widetilde{C}t} \right\|_{\mathscr{B}(\mathbb{R}^d)}, \qquad \forall\, t \ge 0, \tag{4.2.6}$$

where $||\cdot||_{\mathscr{B}(V_0^\perp)}$ denotes the operator norm on $\mathscr{H}$ and orthogonality is considered w.r.t. $\mathscr{H}$. Moreover,

$$||A||_{\mathscr{B}(\mathbb{R}^d)} := \sup_{0 \ne x_0 \in \mathbb{R}^d} \frac{||Ax_0||_2}{||x_0||_2}$$

denotes the spectral matrix norm of any matrix $A \in \mathscr{M}$.

This result motivates to investigate the maximum spectral gap of C. Indeed, the next theorem (see Theorem 2.1 in [8]) identifies the maximum spectral gap of matrices of the form $C = (D + J)K^{-1}$, and its proof (in [8]) provides an explicit, algorithmic construction of a corresponding, optimal matrix pair (C, D).

Theorem 4.2.5. *For $K \in \mathscr{S}^{>0}$ given,*

$$\max\{\rho(C) : \ (C, D) \in \mathscr{I}(K)\} = \max(\sigma(K^{-1})) = \min(\sigma(K))^{-1}. \tag{4.2.7}$$

Concerning the above questions, the article [8] gives the following (partial) answers: The authors give a complete and positive answer to question Q1, obtaining the optimal decay rate $\lambda_{opt} = \max(\sigma(K^{-1}))$. Their optimal pair $(C_{opt}, D_{opt}) \in \mathscr{I}(K)$ is very degenerate, the rank of D_{opt} being one (and this will also be the case for our approach below). But concerning questions Q2 and Q3, they obtain an estimate for the multiplicative constant that grows dramatically with the dimension (in fact of order d^{40d^2}). This is obtained in [8] when considering a FP-equation with *t*ime-independent coefficients, i.e. the equation form introduced in (4.2.2). As a remedy, the authors then considered *t*ime-dependent coefficients, using a symmetric FP-equation with the matrices (K^{-1}, I_d) for small times and a non-symmetric FP-equation for large times. Discontinuous coefficients were used there for analytical reasons, to improve decay estimates. But since their estimates are not sharp, it is not clear if time-dependent coefficients are really able to enhance the decay property of the *e*xact FP-propagator norm, i.e. the true function of time, without estimates. We shall return to this question in §4.4.3 to elucidate question Q5.

While the main result of [8] is presented for the logarithmic relative entropy, the same argument works also for the L^2-norm, as already noted on page 5, [8]:

Theorem 4.2.6 (Theorem 2.2, [8])**.** *Let $K \in \mathscr{S}^{>0}$ be given.*

(a) For any $\tilde{c} > 1$ it is possible to construct a matrix pair $(C_{opt}, D_{opt}) \in \mathscr{I}(K)$ such that, for all normalized $f_0 \in \mathscr{H}$ and for all $t_0 > 0$,

$$||f_t - f_{\infty,K}||^2_{\mathscr{H}} \le \tilde{c}\frac{\max(\sigma(K^{-1}))}{2t_0} e^{-2\max(\sigma(K^{-1}))(t-t_0)}||f_0 - f_\infty||^2_{\mathscr{H}}, \qquad t \ge t_0, \tag{4.2.8}$$

where f_t solves the following system of FP-equations

$$\begin{cases} \partial_t f_t = \operatorname{div}_x(\nabla_x f_t + K^{-1} x f_t), & 0 \le t \le t_0, \\ \partial_t f_t = \operatorname{div}_x(D_{opt}\nabla_x f_t + C_{opt} x f_t), & t > t_0. \end{cases} \tag{4.2.9}$$

(b) For the choice $\tilde{c} = 2$ in part (a), the matrix C_{opt} can be estimated as

$$||C_{opt}||_{\mathscr{F}} \le 4d^2\sqrt{\kappa(K)}\,\lambda_{opt}, \tag{4.2.10}$$

where $||\cdot||_{\mathscr{F}}$ denotes the Frobenius norm $||A||_{\mathscr{F}} := \sqrt{\operatorname{Tr}\left(A^T A\right)}$, and $\kappa(K)$ is the condition number of K.

Optimizing the estimate (4.2.8) w.r.t. the switching time t_0, and using the trivial bound $\left\| e^{-L_{C,D}t}\right\|_{\mathscr{B}(V_0^\perp)} \le 1$ we obtain the following result:

Corollary 4.2.7. Under the assumptions of Theorem 4.2.6, and when choosing $t_0 := \min(\sigma(K))/2$, the following estimate holds for all normalized $f_0 \in \mathscr{H}$:

$$||f_t - f_{\infty,K}||^2_{\mathscr{H}} \le ||f_0 - f_\infty||^2_{\mathscr{H}} \times \begin{cases} 1, & 0 \le t \le t_0, \\ \min\left\{1, \tilde{c}\,\kappa(K)\,e^{1-2\max(\sigma(K^{-1}))t}\right\}, & t > t_0. \end{cases} \tag{4.2.11}$$

where f_t solves the FP-system (4.2.9).

Hence, Theorem 2.2 from [8] can only yield multiplicative constants $c = \sqrt{\tilde{c}\,\kappa(K)\,e} > \sqrt{\kappa(K)\,e}$, using the notation of (4.1.10).

In the next section we shall improve this result in three directions: Answering question Q2 we shall prove that c_{inf} is always 1, and concerning question Q3 we shall construct an optimal matrix pair $(C_{opt}(c), D_{opt}(c))$ for any given $c > 1$. Moreover, we shall not need to split the FP-evolution in time, in contrast to (4.2.9). Our key ingredient to obtain an improved result (compared to [11, 8]) is the equality of the propagator norms of the FP-equation and of its drift ODE, see Theorem 4.2.4. This reduces the quest for an optimal decay estimate to an analogous, and hence easier ODE problem, without having to invoke a hypocoercive entropy method as in the proof of Theorem 2.2, [8], or the block-diagonal decomposition of the FP-propagator as in the proof of Proposition 11, [11]. Finally, concerning question Q4 we shall show that our drift matrix $C_{opt}(c)$ grows like $\mathcal{O}(d^{3/2})$ (for any fixed $c>1$), compared to an $\mathcal{O}(d^2)$–growth in [8].

4.2.1 Time-dependent coefficients

In order to analyze also the decay behaviour of the split FP-equation (4.2.9), we shall next admit in the FP-equation (4.2.2) time-dependent coefficient matrices:

$$\partial_t f_t = -L(t) f_t := \operatorname{div}_x(D(t)\nabla_x f_t + C(t) x f_t), \qquad x\in\mathbb{R}^d,\ t\in(0,\infty). \tag{4.2.12}$$

Here we assume that each FP-operator $L(t)$, with $t\ge 0$ fixed, admits $f_{\infty,K}$ as a steady state, and that the covariance matrix $K\in\mathcal{S}^{>0}$ is given and independent of t. Hence, the coefficient matrices satisfy $(C(t),D(t))\in\mathcal{I}(K)$ $\forall\, t\ge 0$, and by Lemma 4.2.1:

$$C(t) = (D(t)+J(t))\,K^{-1}, \quad \text{with some } J(t)\in\mathcal{A},\ \forall\, t\ge 0.$$

We shall assume $\forall\, t\ge 0$ that $\rho(C(t))>0$. Hence, by Lemma 4.2.2 each FP-operator $L(t)$ is hypocoercive. For (4.2.12), Theorem 4.2.4 can be extended: In the following theorem $S(t_2,t_1)$ and $T(t_2,t_1)$, $0\le t_1\le t_2<\infty$ will denote, respectively, the propagator operators for the PDE (4.2.12) and the ODE (4.2.13) that map an initial condition at time t_1 to the solution at time t_2.

Theorem 4.2.8. *Let $K\in\mathcal{S}^{>0}$ be given. Let $(C(t),D(t))\in\mathcal{I}(K)$ be piecewise smooth functions of $t\ge 0$ (where points of discontinuity do not accumulate), such that the initial value problem for* (4.2.12) *admits a unique solution in $C([0,\infty);\mathcal{H})$. Then, the propagator norms of* (4.2.12) *and of its corresponding drift ODE,*

$$\frac{d}{dt}x = -\widetilde{C}(t)x, \quad t\in(0,\infty), \tag{4.2.13}$$

where $\widetilde{C}(t) := K^{-1/2}\,C(t)\,K^{1/2}$, are equal, i.e.:

$$\|S(t_2,t_1)\|_{\mathcal{B}(V_0^\perp)} = \|T(t_2,t_1)\|_{\mathcal{B}(\mathbb{R}^d)}, \qquad \forall 0\le t_1\le t_2<\infty. \tag{4.2.14}$$

Since this result is a straightforward extension of Theorem 2.3 in [4], we shall give only some hints on the notational differences in the §5.7.

4.3 Main result

The next theorem is the main result of this work. It states the existence of pairs $(C_{opt},D_{opt}) = (C_{opt}(c),D_{opt}(c))\in\mathcal{I}(K)$ that yield the maximum decay rate of the propagator norm of $e^{-L_{C_{opt},D_{opt}}t}$, and in parallel yielding a multiplicative constant c arbitrarily close to 1.

Theorem 4.3.1. *Let $K\in\mathcal{S}^{>0}$ be given.*

(a) Then, for any constant $c > 1$ there exists a pair $(C_{opt}, D_{opt}) = (C_{opt}(c), D_{opt}(c)) \in \mathscr{I}(K)$ such that

$$\left\| e^{-L_{C_{opt},D_{opt}} t} \right\|_{\mathscr{B}(V_0^\perp)} \leq c e^{-\max(\sigma(K^{-1}))t}, \qquad t \geq 0. \tag{4.3.1}$$

(b) The matrices from part (a) can be estimated as

$$\|C_{opt}\|_{\mathscr{F}} \leq \lambda_{opt} \left[d + \sqrt{\kappa(K)} \frac{2\pi c^2}{\sqrt{3}(c^2-1)} \sqrt{d}\,(d-1) \right], \quad \|D_{opt}\|_{\mathscr{F}} = d. \tag{4.3.2}$$

In the proof we shall build upon the strategy from §3 in [11], and only deviate from their strategy in Step 2 below. Nevertheless we outline the full proof, to make it readable independently.

Proof of Theorem 4.3.1(a). We recall that, given any matrix pair (C, D) in $\mathscr{I}(K)$, we can rewrite the drift matrix C (see Lemma 4.2.1) as

$$C = (D + J)K^{-1} = K^{1/2}(\widetilde{D} + \widetilde{J})K^{-1/2},$$

where $\widetilde{D} := K^{-1/2} D K^{-1/2}$ and $\widetilde{J} = K^{-1/2} J K^{-1/2}$. Moreover it is easy to check that the map $M \mapsto K^{-1/2} M K^{-1/2}$ is a bijection that leaves $\mathscr{S}^{\geq 0}$ and $\mathscr{A}$ invariant. We split the proof into three steps.

Step 1 We shall construct an optimal pair $(\widetilde{D}_{opt}, \widetilde{J}_{opt})$ and investigate the propagator norm of the ODE-evolution

$$\frac{d}{dt} x = -\widetilde{C}_{opt} x, \quad x_0 := x(0) \in \mathbb{R}^d, \ t \geq 0, \tag{4.3.3}$$

where $\widetilde{C}_{opt} := \widetilde{D}_{opt} + \widetilde{J}_{opt}$. More precisely, we shall provide a decay estimate for $\|e^{-\widetilde{C}_{opt} t}\|_{\mathscr{B}(\mathbb{R}^d)}$ by constructing an appropriate Lyapunov functional (following §2.1 of [1]).

Following the proof of Theorem 2.1 in [8] we recall that D can enable the maximum decay rate $\lambda_{opt} := \max(\sigma(K^{-1}))$, only if the range of D is a subset of Ω, i.e. the eigenspace of K^{-1} corresponding to λ_{opt}. Hence we let $v \in \mathbb{R}^d$ be a normalized eigenvector of K^{-1} associated to λ_{opt}. As in [8] we define the rank-1 matrix $D_{opt} := d(v \otimes v) \in \mathscr{S}^{\geq 0}$ with $\mathrm{Tr}(D_{opt}) = d$. It follows that

$$\widetilde{D}_{opt} = d K^{-1/2}(v \otimes v) K^{-1/2} = d\lambda_{opt}(v \otimes v) = \lambda_{opt} D_{opt}, \tag{4.3.4}$$

and hence

$$\frac{\mathrm{Tr}\left(\widetilde{D}_{opt}\right)}{d} = \lambda_{opt}.$$

For $\dim(\Omega) > 1$, we remark that the choice of $\widetilde{D}_{opt}$ made in (4.3.4) is just *one* simple option, which enables the decay rate λ_{opt}. For the construction of $\widetilde{J}_{opt} \in \mathscr{A}$ we use a particular basis of $\mathbb{R}^d$: Let $\{\psi_k\}_{k=1}^d$ be an orthonormal basis of $\mathbb{R}^d$ such that the following condition (Lemma 2, [11]) is satisfied: for all $k \in \{1, ..., d\}$,

$$\langle \psi_k, \widetilde{D}_{opt} \psi_k \rangle = \frac{\mathrm{Tr}\left(\widetilde{D}_{opt}\right)}{d} = \lambda_{opt}. \tag{4.3.5}$$

The existence of such basis is guaranteed by Proposition 3 in the same paper. The essence of the basis $\{\psi_k\}_{k=1}^d$ is to provide an equidistribution of $\mathrm{Tr}\left(\widetilde{D}_{opt}\right)$ into the directions $\{\psi_k\}_{k=1}^d$, while $\widetilde{D}_{opt}$ has only rank 1. This is the starting point to enable a uniform (in x_0 and t) decay estimate of all trajectories of (4.3.3), see (4.3.9) below. We observe that in [11] the hypotheses of Proposition 3 require $\widetilde{D}_{opt}$ to be invertible. However, this condition can be weakened to $\widetilde{D}_{opt} \in \mathscr{S}^{\geq 0}$, as already pointed out in [8]: $\widetilde{D}_{opt} + \epsilon I_d \in \mathscr{S}^{>0}$, and $\epsilon \to 0^+$ yields the above result.

Next, let $0 < \lambda_1 < \cdots < \lambda_d$ be arbitrary numbers in $\mathbb{R}$ that will be chosen later in a suitable way. We define the matrix $\widetilde{J}_{opt} := \Psi \widehat{J}_{opt} \Psi^{-1} \in \mathscr{A}$, with $\Psi := [\psi_1, ..., \psi_d]$ and $\widehat{J}_{opt}$ is the anti-symmetric matrix with elements (as in Lemma 2, [11]):

$$\left(\widehat{J}_{opt}\right)_{j,k} := \frac{\lambda_j + \lambda_k}{\lambda_j - \lambda_k} \langle \psi_j, \widetilde{D}_{opt} \psi_k \rangle, \qquad \forall j \neq k, \tag{4.3.6}$$

and 0 else.

Now, the strategy consists in finding a suitable symmetric matrix $P \in \mathscr{S}^{>0}$ that defines a modified norm $||\cdot||_P$ in $\mathbb{R}^d$ such that the trajectories of the ODE (4.3.3) decay with pure exponential decay rate λ_{opt} w.r.t. this norm.

Step 2 Let us proceed with the construction of its inverse matrix $Q := P^{-1} \in \mathscr{S}^{>0}$. We define $Q := \Psi \Lambda \Psi^{-1}$, with the matrix $\Lambda := \mathrm{diag}(\lambda_1, ..., \lambda_d)$. We observe that $Q \in \mathscr{S}^{>0}$ due to the orthonormality of Ψ and the positivity of λ_i. Moreover, by definition, Q has the eigenvectors ψ_i and eigenvalues λ_i. By using Lemma 2, [11] (or a straightforward computation using (4.3.5)) the following Lyapunov equation holds for Q:

$$\widetilde{J}_{opt} Q - Q \widetilde{J}_{opt} = -Q \widetilde{D}_{opt} - \widetilde{D}_{opt} Q + 2\lambda_{opt} Q. \tag{4.3.7}$$

Let us define the modified norm $||x||_P^2 := \langle x, Px \rangle$ on $\mathbb{R}^d$, where $P := Q^{-1} \in \mathscr{S}^{>0}$. Differentiating this norm along a trajectory of the ODE (4.3.3) we obtain with (4.3.7), multiplied on either side by $P = Q^{-1}$:

$$\frac{d}{dt}||x(t)||_P^2 = -\left\langle x(t), \left[P\left(\widetilde{D}_{opt} + \widetilde{J}_{opt}\right) + \left(\widetilde{D}_{opt} - \widetilde{J}_{opt}\right)P\right] x(t) \right\rangle = -2\lambda_{opt}||x(t)||_P^2. \tag{4.3.8}$$

Hence the modified norm decays with rate λ_{opt}, i.e.

$$||x(t)||_P^2 = e^{-2\lambda_{opt} t}||x(0)||_P^2, \qquad t \geq 0. \tag{4.3.9}$$

Transforming to the Euclidean vector norm, we obtain for the propagator

$$\left|\left|e^{-\widetilde{C}_{opt} t}\right|\right|_{\mathscr{B}(\mathbb{R}^d)} \leq \sqrt{\kappa(P)}\, e^{-\lambda_{opt} t}, \qquad t \geq 0, \tag{4.3.10}$$

where $\kappa(P)$ denotes the condition number of the matrix P.

Step 3 The multiplicative constant appearing in (4.3.10) can be adjusted by choosing the eigenvalues of P in the following way: Given any $c > 1$, and due to the fact that $\kappa(P) = \kappa(Q) = \frac{\lambda_d}{\lambda_1}$, it is sufficient to choose λ_d and λ_1 such that their quotient is (less or) equal to c^2. The remaining parameters $\lambda_2 < ... < \lambda_{d-1} \in (\lambda_1, \lambda_d)$ could be freely chosen at this point, but assigning them a precise value will be crucial in the proof of part (b).

To summarize, we have proved so far that, for any prescribed $c > 1$, there exists a pair of matrices $\widetilde{J}_{opt} \in \mathscr{A}$ and $\widetilde{D}_{opt} \in \mathscr{S}^{\geq 0}$ such that

$$\left|\left|e^{-\widetilde{C}_{opt} t}\right|\right|_{\mathscr{B}(\mathbb{R}^d)} \leq c e^{-\lambda_{opt} t}, \qquad t \geq 0. \tag{4.3.11}$$

We conclude the proof by combining Theorem 4.2.4 applied to the operator $e^{-L_{C_{opt}, D_{opt}} t}$, and the above inequality (4.3.11). □

For $K = \alpha I_d$, we remark that a trivial modification of the above proof admits the choice $(C_{opt}, D_{opt}) = (K^{-1}, I_d)$, $J = 0$, $P = I_d$. In this case the reversible dynamics is already optimal with $\lambda_{opt} = \alpha^{-1}$ and $c = 1$ in (4.3.1). Moreover, $||C||_{\mathcal{F}} = \lambda_{opt}\sqrt{d}$.

Proof of Theorem 4.3.1(b). First we compute the Frobenius norm of $D_{opt} := d(v \otimes v)$, with $v \in \mathbb{R}^d$ normalized eigenvector of K^{-1}:

$$||D_{opt}||^2_{\mathcal{F}} = d^2\,\mathrm{Tr}\left(D^2_{opt}\right) = d^2 ||v||_2^4 = d^2. \tag{4.3.12}$$

For estimating $||C_{opt}||_{\mathcal{F}}$ we recall $C_{opt} = D_{opt}K^{-1} + K^{1/2}\widetilde{J}_{opt}K^{-1/2}$, which implies using the inequality $||AB||_{\mathcal{F}} \le ||A||_{\mathcal{F}} ||B||_{\mathcal{B}(\mathbb{R}^d)}$ (see [10], p. 364):

$$||C_{opt}||_{\mathcal{F}} \le ||D_{opt}||_{\mathcal{F}} \max(\sigma(K^{-1})) + \sqrt{\kappa(K)}\,||\widetilde{J}_{opt}||_{\mathcal{F}}\,. \tag{4.3.13}$$

Since the Frobenius norm is unitarily invariant and $\widetilde{J}_{opt} = \Psi \widehat{J}_{opt}\Psi^{-1}$ we have $||\widetilde{J}_{opt}||_{\mathcal{F}} = ||\widehat{J}_{opt}||_{\mathcal{F}}$. For any $k = 1, ..., d$ we define $\alpha_k := \langle v, \psi_k\rangle$ and we observe that $\alpha_k^2 = \frac{1}{d}$: Indeed from (4.3.5) it follows that

$$\lambda_{opt} = \langle \psi_k, \widetilde{D}_{opt}\psi_k\rangle = d\lambda_{opt}\alpha_k^2.$$

Hence we can rewrite (4.3.6) as

$$\left(\widehat{J}_{opt}\right)_{j,k} = \frac{\lambda_j + \lambda_k}{\lambda_j - \lambda_k} d\lambda_{opt}\alpha_j\alpha_k, \qquad \forall j \ne k. \tag{4.3.14}$$

It follows that

$$||\widehat{J}_{opt}||^2_{\mathcal{F}} = d^2\lambda^2_{opt} \sum_{j\ne k=1}^{d} \left(\frac{\lambda_j + \lambda_k}{\lambda_j - \lambda_k}\right)^2 \alpha_j^2\alpha_k^2 = \lambda^2_{opt} \sum_{j\ne k=1}^{d} \left(\frac{\lambda_j + \lambda_k}{\lambda_j - \lambda_k}\right)^2. \tag{4.3.15}$$

Next we choose the parameters λ_k, $k = 1, ..., d$ as

$$\lambda_k := \frac{d-1}{c^2-1} + k - 1, \tag{4.3.16}$$

and they satisfy $0 < \lambda_1 < \cdots < \lambda_d$ and $\frac{\lambda_d}{\lambda_1} = c^2$. Moreover we have for $j \ne k$: $(\lambda_j + \lambda_k)^2 < (2\lambda_d)^2 = 4\left(\frac{c^2}{c^2-1}\right)^2 (d-1)^2$, which implies with (4.3.15):

$$||\widehat{J}_{opt}||^2_{\mathcal{F}} \le \lambda^2_{opt} 4\left(\frac{c^2}{c^2-1}\right)^2 (d-1)^2 \sum_{j\ne k=1}^{d} \frac{1}{(\lambda_j - \lambda_k)^2}. \tag{4.3.17}$$

With the following estimate of a hyperharmonic series

$$\sum_{j\ne k=1}^{d} \frac{1}{(\lambda_j - \lambda_k)^2} = \sum_{j=1}^{d}\sum_{\substack{k=1\\k\ne j}}^{d} \frac{1}{(j-k)^2} = \sum_{j=1}^{d}\left(\sum_{l=1}^{j-1}\frac{1}{l^2} + \sum_{l=1}^{d-j}\frac{1}{l^2}\right) \le \sum_{j=1}^{d}\frac{\pi^2}{3} = d\frac{\pi^2}{3},$$

and (4.3.17) we obtain

$$||\widehat{J}_{opt}||_{\mathcal{F}} \le \lambda_{opt}\frac{2\pi}{\sqrt{3}}\frac{c^2}{c^2-1}\sqrt{d}(d-1). \tag{4.3.18}$$

We conclude the proof by combining the (in)equalities (4.3.12), (4.3.13), and (4.3.18). □

Let us briefly compare the strategy of proofs for Theorem 4.3.1 here and for Theorem 2.2 in [8]: The main difference concerns how to connect the evolution of the drift ODE to the FP-equation (here via the equality of the propagator norms, and via a hypocoercive entropy method in [8]). Further, our choice of the parameters λ_k is (slightly) improved compared to the choice

$$\lambda_k = d + k, \tag{4.3.19}$$

in [11, Remark 7] and [8]. Finally, the proof of Theorem 4.3.1(b) provides a refined estimate of $||C_{opt}||_{\mathscr{F}}$.

Remark 4.3.2. We note that, for any $c > 1$, an optimal matrix pair $(C_{opt}(c), D_{opt}(c))$ is *n*ot unique: Using in the proof of Theorem 4.3.1 the matrices $\widetilde{J}^T_{opt}$, $\widetilde{C}^T_{opt}$ instead of, respectively, $\widetilde{J}_{opt}$, $\widetilde{C}_{opt}$ and the norm $||\cdot||_Q$ instead of $||\cdot||_P$ yields another non-symmetric FP-equation that satisfies the same estimates (4.3.1), (4.3.2).

4.4 Examples and numerical illustrations

In this section we shall illustrate the results of §4.3. For an explicit example in $\mathbb{R}^2$ we shall give a plot of the exact propagator norm for the FP-equation, which is accessible due to Theorem 4.2.4 for constant-in-time coefficients and due to Theorem 4.2.8 for the time-dependent case. First of all we shall illustrate Theorem 4.3.1(a), particularly focussing on the multiplicative constant in the exponential decay estimate (4.3.1).

4.4.1 Optimal decay estimates

As a first example we consider the covariance matrix $K = \mathrm{diag}(1,2) \in \mathbb{R}^{2\times 2}$. Then the maximum decay rate for FP-equations that converge to $f_{\infty,K}$ is $\lambda_{opt} = \min(\sigma(K))^{-1} = 1$. Next we shall construct one optimal pair of matrices (C_{opt}, D_{opt}) such that $e^{-L_{C_{opt},D_{opt}}t} f_0$ converges to $f_{\infty,K}$ with decay rate λ_{opt} and with a multiplicative constant arbitrary close to one. For any $c > 1$ we choose real numbers $0 < \lambda_1 < \lambda_2$ such that $\frac{\lambda_2}{\lambda_1} = c^2$. We abbreviate $\mu := \frac{\lambda_2+\lambda_1}{\lambda_2-\lambda_1} = \frac{c^2+1}{c^2-1} > 1$. Following the procedure described in the proof of Theorem 4.3.1(a) we first compute $D_{opt} = \widetilde{D}_{opt} = \mathrm{diag}(2,0) \in \mathbb{R}^{2\times 2}$. An orthonormal basis of $\mathbb{R}^2$ satisfying condition (4.3.5) is given by $\psi_1 := \frac{1}{\sqrt{2}}(1,1)^T$ and $\psi_2 := \frac{1}{\sqrt{2}}(-1,1)^T$. This defines the anti-symmetric matrix $\widetilde{J}_{opt} = \widehat{J}_{opt} = \begin{pmatrix} 0 & \mu \\ -\mu & 0 \end{pmatrix}$. Finally we compute

$$C_{opt} = \begin{pmatrix} 2 & \frac{\mu}{\sqrt{2}} \\ -\sqrt{2}\mu & 0 \end{pmatrix}, \quad \text{and} \quad \widetilde{C}_{opt} = \begin{pmatrix} 2 & \mu \\ -\mu & 0 \end{pmatrix}. \tag{4.4.1}$$

The spectral gaps of the drift matrix C_{opt} and the operator $L_{C_{opt},D_{opt}}$ coincide and are equal to 1. C_{opt} has the two distinct eigenvalues $1 \pm i\sqrt{\mu^2-1}$ (because $\mu > 1$), which are also eigenvalues of $L_{C_{opt},D_{opt}}$ (see Theorem 5.3 in [5] or Proposition 10 in [11]). Hence, $\lambda_{opt} = 1$ is indeed the largest possible, uniform decay rate of the FP-propagator $e^{-L_{C_{opt},D_{opt}}t}$ on $V_0^{\perp}$.

Thanks to Theorem 4.2.4 we can reduce the evaluation of the multiplicative constant c in the decay estimate (4.1.10) to the study of the propagator norm of the associated drift ODE $\frac{d}{dt}x = -\widetilde{C}_{opt}x$. In Theorem 3.7, [2] the authors provide the explicit form of the best multiplicative constant for an ODE $\frac{d}{dt}x = -Ax$ in $\mathbb{R}^2$ when the matrix $A \in \mathbb{R}^{2\times 2}$ is positive stable, diagonalizable and $\operatorname{Re}\tau_1 = \operatorname{Re}\tau_2$, with τ_j, $j = 1,2$ the eigenvalues of A: Then the best constant c_{min} in the exponential decay estimate for e^{-At} is given by

$$c_{min} = \sqrt{\frac{1+\alpha}{1-\alpha}}, \qquad \alpha := \left| \left\langle \frac{v_1}{||v_1||_2}, \frac{v_2}{||v_2||_2} \right\rangle \right|, \tag{4.4.2}$$

where $v_i \in \mathbf{C}^2$, $i = 1,2$ denote the eigenvectors of A. Since $\widetilde{C}_{opt}$ satisfies the hypotheses of Theorem 3.7 in [2], a straightforward computation gives $\alpha = \frac{1}{|\mu|}$, and the best multiplicative constant is $c_{min} = \sqrt{\frac{\lambda_2}{\lambda_1}} = c$, coinciding with the statement of Theorem 4.3.1(a).

We observe that $c \searrow 1$ implies $\mu \to \infty$. This corresponds to the $h$igh-rotational limit in the drift matrix of the FP-equation

$$\partial_t f_t = \operatorname{div}_x(D_{opt}\nabla_x f_t + C_{opt}(\mu) x f_t).$$

For increasing $|\mu|$, the latter tends to mix with increasing speed the dissipative and non-dissipative directions (i.e. x_1 and x_2, respectively) of the corresponding symmetric FP-equation (i.e. with $\mu = 0$).

As stated in Remark 4.3.2, replacing μ by $-\mu$ yields another FP-equation with the same optimal decay behaviour. Only the rotational direction is then reversed.

4.4.2 Numerical illustrations: time-independent FP-equations

To illustrate the construction of optimal coefficient matrices in Theorem 4.3.1(a) we revisit the 2D-example from [8], i.e. $K = \operatorname{diag}(1/\varepsilon, 1)$, $\varepsilon = 0.05$ which admits the optimal decay rate $\lambda_{opt} = 1$. For any given multiplicative constant $c > 1$, the optimal coefficient matrices constructed in Theorem 4.3.1(a) read:

$$D_{opt} = \widetilde{D}_{opt} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}, \quad C_{opt} = \begin{pmatrix} 0 & -\frac{\mu}{\sqrt{\varepsilon}} \\ \sqrt{\varepsilon}\mu & 2 \end{pmatrix}, \quad \widetilde{C}_{opt} = \begin{pmatrix} 0 & -\mu \\ \mu & 2 \end{pmatrix}, \quad \mu := \frac{c^2+1}{c^2-1}. \tag{4.4.3}$$

In Figure4.1 we present the exact propagator norms (as a function of time) of the FP-equation and of its drift ODE, i.e.

$$\left\| e^{-L_{C,D}t} \right\|_{\mathscr{B}(V_0^{\perp})} = \left\| e^{-\widetilde{C}t} \right\|_{\mathscr{B}(\mathbb{R}^d)}, \quad t \geq 0 \tag{4.4.4}$$

for several prescribed values of the multiplicative constant: $c = 1.5, 2, 3$. This figure includes also the r.h.s. of the corresponding exponential decay estimate (4.3.1), using a logarithmic scale for the ordinate axis. Being the exact upper envelops, this reveals that this estimate is indeed sharp, concerning both the exponential rate and the multiplicative constant. Also note that each curve of the propagator norm periodically touches (from above) the curve corresponding to the high-rotational limit, given by $e^{-\lambda_{opt}t}$.

Continuing with the same example, we shall next compare the results from Theorem 4.3.1(a) here and Theorem 2.2 in [8]. First we need to explain the criterion of comparison: For both results, and for a given constant $c > 1$ we seek a matrix pair (C, D) such that the inequality (4.3.1) holds. Since $c_{inf} = 1$, such a pair can always be found, but $\|C\|_{\mathscr{F}}$ becomes large as $c \searrow 1$ (see

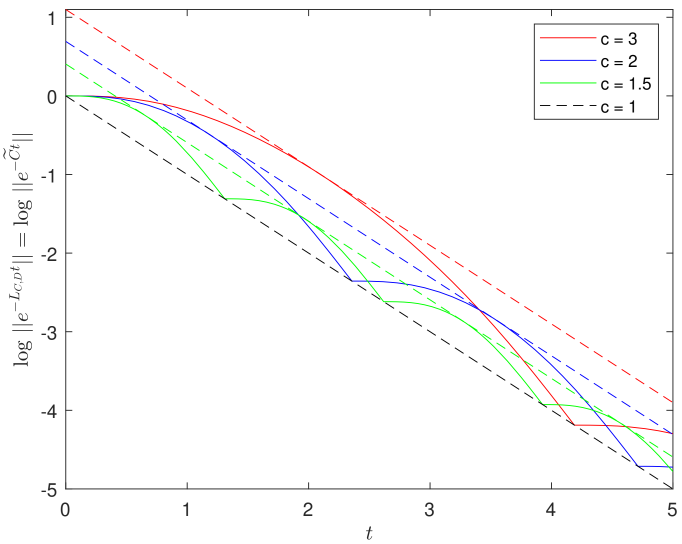

Figure 4.1: The solid curves show the FP- and ODE-propagator norms as functions of t for 3 values of the multiplicative parameter: $c = 3, 2, 1.5$ (top to bottom). The dashed curves give the corresponding (sharp) exponential bound of the form $ce^{-\lambda_{opt}t}$ for the 3 cases. The dashed black curve shows the exponential bound in the high-rotational limit, i.e. for $c \searrow 1$. Colors only online.

§4.4.1). So, asking (4.4.4) to be close to the high-rotational limit $e^{-\lambda_{opt}t}$ cannot be a useful criterion. Instead, for given $c > 1$ we want to find $(C, D) \in \mathscr{I}(K)$ such that (4.3.1) holds and $\|C\|_{\mathscr{F}}$ is minimal. This has also a practical implication for solving the FP-equation (4.2.2) numerically: $\|C\|_{\mathscr{F}}$ "small" allows to use "large" time steps.

For fixed $c = \sqrt{2}$, Theorem 4.3.1(a) here and Theorem 2.2 in [8] yield, respectively,

$$\widetilde{C}^{AS}_{opt} = \begin{pmatrix} 0 & -3 \\ 3 & 2 \end{pmatrix}, \quad \widetilde{C}^{GM}_{opt} = \begin{pmatrix} 0 & -7 \\ 7 & 2 \end{pmatrix},$$

with $\|C^{AS}_{opt}\|_{\mathscr{F}} = \sqrt{184.45}$ and $\|C^{GM}_{opt}\|_{\mathscr{F}} = \sqrt{986.45}$. The essential difference stems from the different choices of λ_1 and λ_2, (4.3.16) vs. (4.3.19). In Theorem 4.3.1(a), the estimate (4.3.1) is sharp, and hence the corresponding plot of the propagator norm has $\sqrt{2}\, e^{-t}$, i.e. the r.h.s. in (4.3.1), as its upper envelop (see Figure 4.2, left). Since the estimate from Theorem 2.2 in [8] is not sharp, the anti-symmetric part of $\widetilde{C}^{GM}_{opt}$ is larger than "necessary", and hence the corresponding plot of the propagator norm stays well below the estimate $\sqrt{2}\, e^{-t}$. With a view towards numerical applications the latter is rather disadvantageous.

Figure 4.2 also shows the decay of the propagator norm of the symmetric FP-equations in (4.2.9) and (4.4.8). Their respective decay rates are $\frac{1}{\lambda_{max}(K)} = \varepsilon$ and $\frac{d}{\mathrm{Tr}(K)} = \frac{2\varepsilon}{1+\varepsilon}$, both well below $\lambda_{opt} = 1$, the rate of the optimal hypocoercive FP-equations.

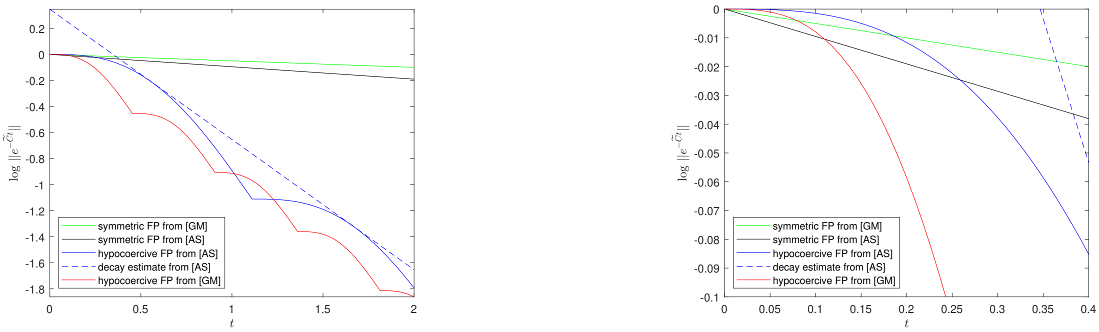

Figure 4.2: Left: For $c = \sqrt{2}$, the solid blue and red curves show the FP- and ODE-propagator norms as functions of t for the hypocoercive FP-equations constructed, respectively, in Theorem 4.3.1(a) here and Theorem 2.2 in [8]. The dashed blue curve gives the corresponding exponential bound $\sqrt{2}\,e^{-t}$; it is sharp for Theorem 4.3.1(a). The solid green and black curves show the FP- and ODE-propagator norms for the symmetric FP-equations in (4.2.9) and (4.4.8), respectively.
Right: a zoom of the plot, close to $t = 0$. Colors only online.

4.4.3 Numerical illustrations: time-dependent FP-equations

In [8] the authors used a FP-equation of the split form (4.2.9) with piecewise constant coefficient matrices in order to approach the given equilibrium quickly. Following this approach, we shall next discuss if time-dependent coefficient matrices $C(t)$, $D(t)$ can accelerate the convergence in FP-equations, compared to the case of constant matrices C, D that was analysed in §4.3.

As a first step we shall show that the initial decay of hypocoercive FP-evolutions, as constructed in the proof of Theorem 4.3.1(a) (recall that $\operatorname{rank}(\widetilde{D}_{opt}) = 1$), can always be improved, e.g. in the spirit of the split FP-equation (4.2.9) proposed in [8]. The following lemma gives, at $t = 0$, the largest possible decay rate of the FP-equation (4.2.2) as well as of its drift ODE $\dot{x} = -\widetilde{C}x$ (both when considering their propagator norms).

Lemma 4.4.1. *Let $K \in \mathcal{S}^{>0}$ be given. For any $(C, D) \in \mathcal{I}(K)$, the maximum decay rate of $\left\| e^{-L_{C,D}t} \right\|_{\mathcal{B}(V_0^\perp)}$ at $t = 0$ equals $\frac{d}{\operatorname{Tr}(K)}$. It is obtained by the symmetric FP-equation with $C = \frac{d}{\operatorname{Tr}(K)} I_d$ and $D = \frac{d}{\operatorname{Tr}(K)} K$.*

Proof. Due to Theorem 4.2.4 we want to maximize the decay of the corresponding ODE-propagator norm,

$$\left\| e^{-\widetilde{C}t} \right\|^2_{\mathcal{B}(\mathbb{R}^d)} = \lambda_{max}\left(e^{-\widetilde{C}^T t} e^{-\widetilde{C}t} \right)$$

at $t = 0$. A Taylor expansion yields

$$\left\| e^{-\widetilde{C}t} \right\|_{\mathcal{B}(\mathbb{R}^d)} = 1 - \lambda_{min}(\widetilde{C}_s)\, t + \mathcal{O}(t^2) \quad \text{as } t \to 0, \tag{4.4.5}$$

where $\widetilde{C}_s := \frac{1}{2}(\widetilde{C} + \widetilde{C}^T)$ is the symmetric part of $\widetilde{C}$. We recall from the proof of Theorem 4.3.1(a) that $\widetilde{C} := K^{-1/2} C K^{1/2}$ and $\widetilde{D} := K^{-1/2} D K^{-1/2} = \widetilde{C}_s \geq 0$.

Thus we are led to the following optimization problem: Find $\widetilde{C}_s \in \mathcal{S}^{\geq 0}$ with

$$\operatorname{Tr}(D) = \operatorname{Tr}(K^{1/2} \widetilde{C}_s K^{1/2}) =: \tau \leq d, \tag{4.4.6}$$

such that $\lambda_{min}(\widetilde{C}_s)$ is maximal. Since $\widetilde{J}$, the anti-symmetric part of $\widetilde{C}$, does not appear within this problem, we set it to 0, for simplicity.

For such an optimal $\widetilde{C}_s$, (4.4.6) actually has to be an equality: Otherwise we would have

$$K^{1/2}\widetilde{C}_s K^{1/2} \le \tau\, I_d < d\, I_d$$

and the matrix $\widetilde{C}_s$ could be "enlarged", e.g. by the matrix

$$A := \frac{d-\tau}{d^2-\tau}(d\,K^{-1} - \widetilde{C}_s) \in \mathscr{S}^{>0}.$$

Then, $\widetilde{C}_s + A$ still satisfies the constraint (4.4.6):

$$\operatorname{Tr}\left(K^{1/2}[\widetilde{C}_s + A]K^{1/2}\right) = d,$$

but $\lambda_{min}(\widetilde{C}_s + A) > \lambda_{min}(\widetilde{C}_s)$, contradicting the optimality of $\widetilde{C}_s$.

Next we shall prove that the optimal matrix satisfies

$$\widetilde{C}_s = \widetilde{C} = \frac{d}{\operatorname{Tr}(K)}\, I_d = C. \tag{4.4.7}$$

If the optimal $0 \ne \widetilde{C}_s \in \mathscr{S}^{\ge 0}$ was not proportional to I_d, we could "reduce" $\widetilde{C}_s$ by the matrix

$$B := \widetilde{C}_s - \lambda\, I_d \ge 0 \quad \text{with } \lambda := \lambda_{min}(\widetilde{C}_s),$$

without changing the smallest eigenvalue. Moreover $\widetilde{C}_s - B = \lambda\, I_d$ satisfies

$$\operatorname{Tr}\left(K^{1/2}[\lambda\, I_d]K^{1/2}\right) = \operatorname{Tr}(K^{1/2}\widetilde{C}_s K^{1/2}) - \operatorname{Tr}(K^{1/2}BK^{1/2}) \le d,$$

and hence $\lambda\, I_d$ is another optimal matrix of the above optimization problem. From the equality requirement in (4.4.6) and $\operatorname{Tr}(K^{1/2}\widetilde{C}_s K^{1/2}) = d$ we then conclude $B = 0$. Hence $\widetilde{C}_s$ is proportional to I_d, and equality in (4.4.6) yields $\lambda_{min}(\widetilde{C}_s) = \frac{d}{\operatorname{Tr}(K)}$, finishing the proof. □

With this lemma we can identify the symmetric FP-equation with steady state $f_{\infty,K}$ that exhibits maximum initial decay as

$$\partial_t f_t = \frac{d}{\operatorname{Tr}(K)} \operatorname{div}_x(K\nabla_x f_t + x f_t), \qquad x \in \mathbb{R}^d,\ t \in (0,\infty). \tag{4.4.8}$$

Its initial decay rate, $\frac{d}{\operatorname{Tr}(K)}$ is larger then that of (4.2.9), namely $\frac{1}{\lambda_{max}(K)}$. We recall that the optimal FP-equations constructed in the proof of Theorem 4.3.1(a) are all hypocoercive, satisfying $\operatorname{rank}(\widetilde{D}_{opt}) = 1$, where $\widetilde{D}_{opt} = (\widetilde{C}_{opt})_s$. Hence $\lambda_{min}\big((\widetilde{C}_{opt})_s\big) = 0$, and the corresponding propagator norm behaves like $1 + \mathscr{O}(t^2)$, see (4.4.5). Therefore it is obvious that, for small time, the symmetric FP-equations (4.2.9) and (4.4.8) both decrease the FP-propagator norm faster than the hypocoercive FP-evolutions from Theorem 4.3.1(a). This is illustrated on a 2D example in Figure 4.2, right.

For the rest of this section we shall base our discussion of using time-dependent coefficients on the concrete example from §4.4.2, again with $\varepsilon = 0.05$, since a general theory of it seems

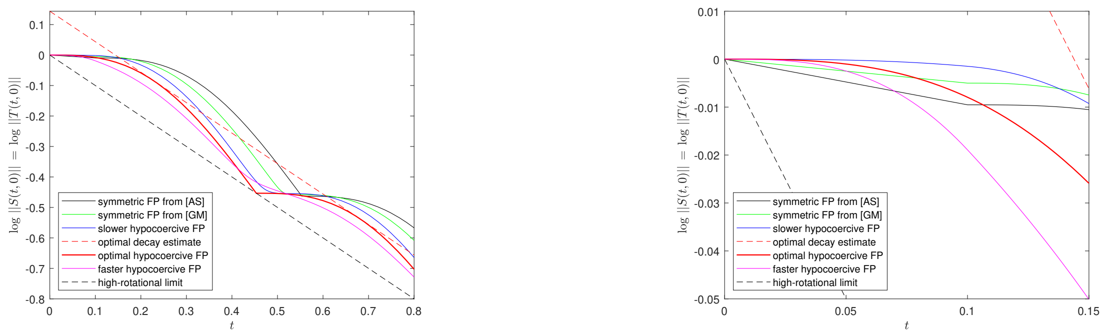

Figure 4.3: Left: For $c = \sqrt{4/3}$, the FP- and ODE-propagator norms are given for hypocoercive FP-equations with piecewise constant coefficients, using 5 different values on $0 \le t \le 0.1$: The solid red curve corresponds to the optimal, constant matrices from Theorem 4.3.1(a) as reference case, and the dashed red curve is the corresponding decay estimate (4.3.1). The initially symmetric FP-equations from (4.4.8) and (4.2.9) are given by the black and green solid curves, respectively. Hypocoercive FP-equations with slower and faster rotational drift are represented, respectively, by the blue and magenta solid curves.
Right: a zoom of the plot, close to $t = 0$. Colors only online.

unreachable to us for the moment. In a numerical case study we shall analyze the FP-propagator norm $\|S(t,0)\|_{\mathscr{B}(V_0^\perp)}$, as a function of time. In the past it would have been quite a challenge to compute (not just to estimate) this norm. But due to Theorem 4.2.8 this has become easy for FP-equations with linear drift.

In Figure 4.3 we shall compare the decay of the FP- and corresponding ODE-propagator norms for 5 cases of FP-equations with piecewise constant coefficient matrices, as in (4.2.9):

$$S(t,0) = \begin{cases} e^{-L_{C_i,D_i} t} & 0 \le t \le t_0 \\ e^{-L_{C_1,D_1}(t-t_0)} e^{-L_{C_i,D_i} t_0} & t > t_0 \end{cases}, \quad T(t,0) = \begin{cases} e^{-\widetilde{C}_i t} & 0 \le t \le t_0 \\ e^{-\widetilde{C}_1 (t-t_0)} e^{-\widetilde{C}_i t_0} & t > t_0 \end{cases}.$$

Choosing $t_0 = 0.1$, we use on the interval (t_0, ∞) always the same matrices, namely those from (4.4.3) with $\mu = 7$, which is the optimal hypocoercive FP-evolution from Theorem 4.3.1 for the multiplicative constant $c = \sqrt{4/3}$. For the interval $[0, t_0]$ we shall compare the following cases:

(FP1) This reference case uses the same coefficients as for $t > t_0$, i.e.:

$$D_1 = \operatorname{diag}(0,2), \quad C_1 = \begin{pmatrix} 0 & -\frac{7}{\sqrt{\varepsilon}} \\ 7\sqrt{\varepsilon} & 2 \end{pmatrix}, \quad \widetilde{C}_1 = \begin{pmatrix} 0 & -7 \\ 7 & 2 \end{pmatrix}. \tag{4.4.9}$$

Figure 4.3 also includes the sharp upper and lower envelops of the resulting propagator norm (as function of t).

(FP2) The symmetric FP-equation from [8], and shown in (4.2.9) is determined by the matrices

$$D_2 = I_2, \quad C_2 = \widetilde{C}_2 = \operatorname{diag}(\varepsilon, 1).$$

(FP3) The symmetric FP-equation (4.4.8) with maximum initial decay is determined by the matrices

$$D_3 = \frac{2\varepsilon}{1+\varepsilon}\,\mathrm{diag}(\frac{1}{\varepsilon},1), \quad C_3 = \widetilde{C}_3 = \frac{2\varepsilon}{1+\varepsilon} I_2.$$

(FP4) A hypocoercive FP-equation with slower rotational part than in (4.4.9) is determined by the matrices

$$D_4 = \mathrm{diag}(0,2), \quad C_4 = \begin{pmatrix} 0 & -\frac{3}{\sqrt{\varepsilon}} \\ 3\sqrt{\varepsilon} & 2 \end{pmatrix}, \quad \widetilde{C}_4 = \begin{pmatrix} 0 & -3 \\ 3 & 2 \end{pmatrix}.$$

(FP5) A hypocoercive FP-equation with faster rotational part than in (4.4.9) is determined by the matrices

$$D_5 = \mathrm{diag}(0,2), \quad C_5 = \begin{pmatrix} 0 & -\frac{11}{\sqrt{\varepsilon}} \\ 11\sqrt{\varepsilon} & 2 \end{pmatrix}, \quad \widetilde{C}_5 = \begin{pmatrix} 0 & -11 \\ 11 & 2 \end{pmatrix}.$$

Note that (FP4) and (FP5) are both of the form (4.4.3).

First we need to fix the criterion for comparing these 5 FP-equations with split coefficients. As one sees from Figure 4.3, adapting the FP-equation only on the initial time interval $[0, t_0]$ has a highly nonlocal-in-t effect. Hence, it does not make sense to compare the norm-curves pointwise in time. Following the paradigm of §4.3, it is appropriate to compare again the corresponding best exponential decay estimates (4.3.1). Since all compared FP-equations coincide for large time, or more precisely on (t_0,∞), their exponential decay rate is the same, and it suffices to compare the multiplicative constant of the (sharp) decay estimates.

Now we shall replace in the reference FP-equation (FP1) the initial phase by a symmetric evolution: With both options (FP2) and (FP3) the propagator norm decays initially faster than for the reference FP-equation (see Figure 4.3, right), but this backfires at later times: In both cases the upper envelop for the whole norm-function on $[0,\infty)$ and hence the multiplicative constant c is larger than for the reference case (FP1) (see Figure 4.3, left).

Finally we shall replace in the reference FP-equation (FP1) the initial phase by a hypocoercive evolution having an anti-symmetric part of $\widetilde{C}$ that differs from case (FP1). With the slower rotational part in case (FP4) the multiplicative constant c is increased (see Figure 4.3, left), but when using initially the faster rotational part from case (FP5), the multiplicative constant c is decreased!

While we present in Figure 4.3 the plots only for $t_0 = 0.1$, the results for other values of $t_0 > 0$ are qualitatively the same. Choosing $t_0 \approx 0.1434$ (i.e. the first point of tangency between $\|e^{-\widetilde{C}_5 t}\|_{\mathscr{B}(\mathbb{R}^d)}$ and its sharp exponential decay estimate $\sqrt{6/5}e^{-t}$, see Figure 4.4) in the split case (FP5) reduces the multiplicative constant to $c = \sqrt{6/5}$. Note that this is also the sharp constant for the *n*on-split FP-equation involving the matrices (C_5, D_5). This means that the same decay quality (in the above defined sense) can be obtained with the constant coefficient matrices (C_5, D_5) for all time or just a short initial layer with (C_5, D_5) and then evolving with (C_1, D_1) for $t > t_0$. The multiplicative constant can be reduced even further, e.g. with the following choice of matrices on the interval $[0, 0.11413]$ (see Figure 4.4):

(FP6)

$$D_6 = \mathrm{diag}(0,2), \quad C_6 = \begin{pmatrix} 0 & -\frac{13.8}{\sqrt{\varepsilon}} \\ 13.8\sqrt{\varepsilon} & 2 \end{pmatrix}, \quad \widetilde{C}_6 = \begin{pmatrix} 0 & -13.8 \\ 13.8 & 2 \end{pmatrix}$$

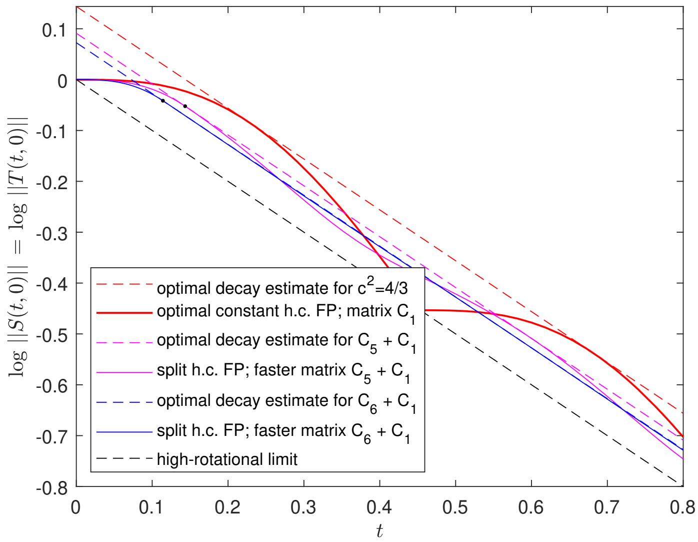

Figure 4.4: For $c = \sqrt{4/3}$, the FP- and ODE-propagator norms are given for hypocoercive (h.c.) FP-equations with piecewise constant coefficients, using 3 different values on $0 \le t \le t_0$: The solid red curve corresponds to the optimal, constant matrices from Theorem 4.3.1(a) as reference case. Hypocoercive FP-equations with the faster rotational drift matrices (FP5), (FP6) are represented by the magenta and blue solid curves, respectively. The dashed curves are the corresponding decay estimate (4.3.1). The discontinuity points t_0 of the coefficient matrices are marked with black dots.

This example of time-dependent FP-coefficients is also algorithmically relevant, since $\|C_1\|_{\mathscr{F}} < \|C_6\|_{\mathscr{F}}$. Hence, longer time steps could be used in a discretization of the split FP-equation for $t > t_0$.

4.5 Conclusion

For any given anisotropic Gaussian steady state (4.2.1) with covariance matrix K, we analysed the construction of non-symmetric FP-equations (4.2.2) that show fastest decay towards the unique normalized steady state $f_{\infty,K}$. Building upon preceding results (in particular [11, 8]) we proved that optimal exponential decay with small multiplicative constants (as in (4.1.10), and uniformly in f_0) can be achieved with a single FP-equation, without having to split off an initial evolution phase. Thereby, the maximum decay rate $\lambda_{opt} = \max(\sigma(K^{-1}))$, and the infimum of the multiplicative constants $c_{inf} = 1$. By contrast, the best multiplicative constant obtainable in [8] was bounded below by $\sqrt{\kappa(K)\,e}$. Hence, the gain provided here for the multiplicative constant is particularly important when $\kappa(K)$ is large, i.e. when the original, symmetric FP-dynamics includes very different time scales due to very different eigenvalues in K.

More precisely, for any given multiplicative constant $c > 1$ we were able to construct explicitly a non-symmetric FP-equation of form (4.2.2) with constant drift matrix $C_{opt}(c)$ and diffusion matrix $D_{opt}(c)$ such that the exponential decay estimate (4.1.10) holds with the parameters (λ_{opt}, c). For given c and variable space dimension d, we were able to reduce the growth estimate on these drift matrices to $\mathscr{O}(d^{3/2})$, down from $\mathscr{O}(d^2)$ given in [8].

In explicit 2D examples we illustrated, both analytically and numerically, that the infimum of the multiplicative constant, $c_{inf} = 1$ corresponds to the limit of adding a highly rotational, non-symmetric drift to the original FP-equation.

To round off our analysis we presented a numerical case study on a FP-equation in 2D with piecewise constant coefficient matrices. This showed two unexpected phenomena: First, no symmetric FP-evolution on an initial time layer was able to improve the overall decay behaviour; in fact it always got worse than in the time-independent case. Second, replacing on an initial time layer the non-symmetric drift by a higher rotational one (and then returning to the original drift for all time) can reduce the multiplicative constant for the whole evolution to a level that pertains to a "larger" drift matrix C.

4.6 Proof of Theorem 4.2.8

Proof-idea. First, the coordinate transformation $\tilde{x} := K^{-1/2}x$ and $\tilde{f}(\tilde{x}) := (\det K)^{1/2} f(K^{1/2}\tilde{x})$ transforms (4.2.12) into the *n*ormalized FP-equation

$$\partial_t \tilde{f}_t = -\widetilde{L}(t)\tilde{f}_t := \operatorname{div}_{\tilde{x}}(\widetilde{D}(t)\nabla_{\tilde{x}}\tilde{f}_t + \widetilde{C}(t)\tilde{x}\tilde{f}_t), \qquad \tilde{x} \in \mathbb{R}^d,\ t \in (0,\infty), \tag{4.6.1}$$

where $\widetilde{D}(t) := K^{-1/2}D(t)K^{-1/2}$. This FP-equation is naturally considered in $\widetilde{\mathscr{H}} := L^2(\mathbb{R}^d, \tilde{f}_\infty^{-1})$, and the (transformed) steady state is

$$\tilde{f}_\infty(\tilde{x}) = (2\pi)^{-d/2}\, e^{-|\tilde{x}|^2/2}.$$

This transformation preserves the norm of the solution: $\|f_t\|_{\mathscr{H}} = \|\tilde{f}_t\|_{\widetilde{\mathscr{H}}}$, $t \geq 0$. Hence the propagator norms of (4.2.12) and (4.6.1) coincide: $\|S(t_2,t_1)\|_{\mathscr{B}(V_0^\perp)} = \left\|\widetilde{S}(t_2,t_1)\right\|_{\mathscr{B}(\widetilde{V}_0^\perp)}$.

Next, one decomposes $\widetilde{\mathscr{H}}$ into mutually orthogonal subspaces $\widetilde{V}^{(m)}$, $m \in \mathbb{N}_0$, which are each invariant under the operators $\widetilde{L}(t)$ $\forall\, t \geq 0$:

$$\widetilde{\mathscr{H}} = \bigoplus_{m \in \mathbb{N}_0}^{\perp} \widetilde{V}^{(m)},$$

with

$$\widetilde{V}^{(m)} := \operatorname{span}\{g_\alpha(\tilde{x}) := (-1)^{|\alpha|} \nabla^\alpha \tilde{f}_\infty(\tilde{x}) : \alpha \in \mathbb{N}_0^d, |\alpha| = m\}.$$

Decomposing the solution of (4.6.1) into these subspaces as

$$\tilde{f}_t(\tilde{x}) = \sum_{\alpha \in \mathbb{N}_0^d} \tilde{d}_\alpha(t) \frac{g_\alpha(\tilde{x})}{\|g_\alpha\|_{\widetilde{\mathscr{H}}}},$$

yields the estimates

$$\sum_{|\alpha|=m} |\tilde{d}_\alpha(t_2)|^2 \leq h(t_2, t_1)^{2m} \left(\sum_{|\alpha|=m} |\tilde{d}_\alpha(t_1)|^2 \right), \quad 0 \leq t_1 \leq t_2 < \infty,\ m \in \mathbb{N},$$

with

$$h(t_2, t_1) := \|T(t_2, t_1)\|_{\mathscr{B}(\mathbb{R}^d)} \leq 1, \quad 0 \leq t_1 \leq t_2 < \infty.$$

On the one hand this shows that

$$\|\tilde{f}_{t_2} - \tilde{f}_\infty\|_{\widetilde{\mathscr{H}}} = \|\widetilde{S}(t_2, t_1)(\tilde{f}_{t_1} - \tilde{f}_\infty)\|_{\widetilde{\mathscr{H}}} \leq \|T(t_2, t_1)\|_{\mathscr{B}(\mathbb{R}^d)} \|\tilde{f}_{t_1} - \tilde{f}_\infty\|_{\widetilde{\mathscr{H}}}, \quad 0 \leq t_1 \leq t_2 < \infty.$$

On the other hand we can use initial conditions $\tilde{f}_{t_1} \in \widetilde{V}^{(1)}$, noting as in [4, §4.2] that the coefficient vector $\tilde{d}^{(1)}(t) := \left(\tilde{d}_\alpha(t)\right)_{|\alpha|=1} \in \mathbb{R}^d$ evolves according to

$$\frac{d}{dt} \tilde{d}^{(1)} = -\widetilde{C}(t)\, \tilde{d}^{(1)},$$

i.e. the drift ODE of the FP-equation. This implies the reverse inequality

$$\left\|\widetilde{S}(t_2, t_1)\right\|_{\mathscr{B}(\widetilde{V}_0^{\perp})} \geq \|T(t_2, t_1)\|_{\mathscr{B}(\mathbb{R}^d)}, \qquad \forall 0 \leq t_1 \leq t_2 < \infty,$$

and hence the equality (4.2.14) follows. □

Acknowledgement

The authors were partially supported by the FWF (Austrian Science Fund) funded SFB #F65 and the FWF-doctoral school W 1245. We acknowledge fruitful discussions with José A. Carrillo, who originally proposed this topic to us. We are also grateful to the anonymous referees, whose suggestions helped to improve this work.

Bibliography

[1] F. Achleitner, A. Arnold, E. A. Carlen, *On multi-dimensional hypocoercive BGK models.* Kinetic and Related Models 11, No. 4 (2018), 953-1009.

[2] 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 282 (2019), 241-264, G. Giacomin et al. (eds.), Springer.

[3] A. Arnold, P. A. Markowich, G. Toscani, A. Unterreiter, *On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations.* Comm. PDE 26, No. 1-2 (2001), 43-100.

[4] A. Arnold, C. Schmeiser, B. Signorello, *Propagator norm and sharp decay estimates for Fokker-Planck equations with linear drift.* Preprint. https://arxiv.org/abs/2003.01405.

[5] 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.

[6] P. Diaconis, *The Markov chain Monte Carlo revolution.* Bulletin of the American Mathematical Society 46, no.2 (2009), 179-205.

[7] H. Dietert, J. Evans, *Finding the jump rate for fastest decay in the Goldstein-Taylor model.* Preprint. https://arxiv.org/abs/2103.10064.

[8] A. Guillin, P. Monmarché, *Optimal linear drift for the speed of convergence of an hypoelliptic diffusion.* Electron. Commun. Probab. 21, no. 74 (2016), 1-14; Erratum: Electron. Commun. Probab. 22, no. 15 (2017), 1-2.

[9] L. Miclo, P. Monmarché, *Étude spectrale minutieuse de processus moins indécis que les autres,* In: C. Donati-Martin, A. Lejay, A. Rouault (eds) Séminaire de Probabilités XLV, Lecture Notes in Mathematics, vol 2078, pp. 459–481. Springer, Heidelberg (2013). English summary available at https://www.ljll.math.upmc.fr/~monmarche

[10] R.A. Horn, C.R. Johnson, *Matrix Analysis*, 2nd edition, Cambridge University Press (2013).

[11] T. Lelièvre, F. Nier, G.A. Pavliotis, *Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion.* Journal of Statistical Physics, 152, no. 2 (2013), 237-274.

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

[13] C. Villani, *Hypocoercivity*, Memoirs of the American Mathematical Society 202 (2009).

5 On the Goldstein-Taylor Equation with Space-Dependent Relaxation

5.1 Introduction

The object of this work is the large time analysis of the Goldstein-Taylor equations on the one-dimensional torus $\mathbb{T}$, i.e. on $[0, 2\pi]$ with periodic boundary conditions, and for $t \in (0, \infty)$:

$$\begin{aligned} \partial_t f_+(x,t) + \partial_x f_+(x,t) &= \frac{\sigma(x)}{2}(f_-(x,t) - f_+(x,t)), \\ \partial_t f_-(x,t) - \partial_x f_-(x,t) &= -\frac{\sigma(x)}{2}(f_-(x,t) - f_+(x,t)), \\ f_\pm(x,0) &= f_{\pm,0}(x), \end{aligned} \tag{5.1.1}$$

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 $\sigma \in L^\infty_+(\mathbb{T}) := \{f \in L^\infty(\mathbb{T}) \mid \operatorname{essmin} f > 0\}$ is the relaxation coefficient, and $f_{\pm,0}$ are the initial conditions. Since (5.1.1) 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

$$h_{\text{avg}} := \frac{1}{2\pi}\int_0^{2\pi} h(x)\,dx. \tag{5.1.2}$$

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 [23, 15], respectively. (5.1.1) can also be seen as a special 1D case of a BGK-model (named after the three physicists Bhatnagar, Gross, and Krook [9]) with a discrete set of velocities. Such equations commonly 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 [10, 18] 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 was devoted to the following three topics: scaling limits, asymptotic preserving (AP) numerical schemes, and large time behaviour. In a diffusive scaling, the Goldstein-Taylor model can be viewed as a hyperbolic approximation to the heat equation [21]. Various AP-schemes for this model in the stiff relaxation regime (i.e. for $\sigma \to \infty$) were constructed and analysed in [17, 16, 4]. Since the large time convergence of solutions to (5.1.1) towards its unique steady state is also the topic of this work, we shall review the related literature in more detail:

Analytically, the main difficulty of (5.1.1) is with its hypocoercivity, as defined in [25]: More specifically, 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, generated by the kernel of the relaxation operator, 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 (5.1.1). The Goldstein-Taylor model was also considered in the analysis of [5], if one chooses the velocity matrix to be $V = \text{diag}(1,-1)$ and the relaxation matrix $\boldsymbol{A}(x)$ to be

$$\boldsymbol{A}(x) = \frac{\sigma(x)}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \geq 0.$$

Exponential convergence to the steady state is then proved in the aforementioned work for the system (5.1.1) *with inflow boundary conditions*. Such boundary conditions make the problem significantly easier than in the periodic set-up envisioned here, in particular it allows for $\sigma(x)$ to be zero on a subset of $\mathbb{T}$, an issue that proves to be far more difficult in our setting.
In [12] the authors proved polynomial decay towards the equilibrium, allowing $\sigma(x)$ to vanish at finitely many points.
In [24] the author proved exponential decay for solutions to (5.1.1) for a more general $\sigma(x) \geq 0$. That work is based on a (non-local in time) *weak coercive estimate* on the damping.
All of the papers mentioned so far did not focus on the optimality of the (exponential) decay rate. Using the equivalence between (5.1.1) and the telegrapher's equation, the authors of [8] have shown that this optimal decay rate, $\mu(\sigma)$, is the minimum of σ_{avg} and the spectral gap of the telegrapher's equation (excluding the case when some of those eigenvalues with real part equal to $\mu(\sigma)$ are defective). The precise value of this spectral gap, however, is hardly accessible - even for simple non-constant relaxation functions $\sigma(x)$ (see e.g. §5.7). Moreover, it is based on the restrictive requirement $f_{\pm,0} \in H^1(\mathbb{T})$, and cannot be extended to other discrete velocity models in 1D. The reason for the latter is that [8] heavily relies on the equivalence of (5.1.1) to the telegrapher's equation.

The issues above motivated our subsequent analysis: We introduce a method for L^2–initial data that can be extended to other discrete velocity BGK-models (as illustrated below for a 3−velocities system), and that yields sharp rates for constant σ. Moreover, and most importantly, it is applicable in the general non-homogeneous $\sigma \in L^\infty_+(\mathbb{T})$ case and yields in these cases an explicit, quantitative lower bound for the decay rate. In this case, however, it will not achieve an optimal rate of convergence[1] 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 [25, 13, 1, 2].
This paper is structured as follows: In §5.2 we give the analytical setting of the problem and present our main convergence result (Theorem 5.2.2). In §5.3 we recall some analytical results which will be needed in the analysis that will follow, and explore some properties of the entropy functional E_θ and the anti-derivative of functions on $\mathbb{T}$, defined in (5.2.2) and (5.2.3), respectively. §5.4 is devoted to the case where $\sigma(x) = \sigma$ is constant, which will motivate our more general approach: Based on a modal decomposition of the Goldstein-Taylor system and its spectral analysis we derive the entropy functional E_θ, first on a modal level and then as a pseudo-differential operator in physical space. We conclude by proving part (a) of our main

[1] at least compared to the H^1-result in [8]

theorem. Continuing to §5.5, we will prove, using a perturbative approach to the problem, part (b) of our main theorem. The robustness of our method will be shown in §5.6 where we use it to obtain an explicit rate of convergence for a 3−velocities Goldstein-Taylor model. Finally, in §5.7 we discuss a potential way to improve the technique from §5.5, and explicitly show the lack of optimality of it for a particular case of $\sigma(x)$.

5.2 The setting of the problem and main results

To better understand the Goldstein-Taylor system, (5.1.1), one starts by recasting it in the macroscopic variables

$$u := f_+ + f_-, \qquad v := f_+ - f_-,$$

representing the spatial (mass) density and the flux density, respectively. The macroscopic variables yield the following system of equations on $\mathbb{T} \times (0,\infty)$:

$$\begin{aligned} \partial_t u(x,t) + \partial_x v(x,t) &= 0, \\ \partial_t v(x,t) + \partial_x u(x,t) &= -\sigma(x) v(x,t), \\ u(.,0) = u_0 := f_{+,0} + f_{-,0}, \quad & v(.,0) = v_0 := f_{+,0} - f_{-,0}\,, \end{aligned} \tag{5.2.1}$$

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, [20]). Moreover, when one tries to understand the qualitative behaviour of (5.2.1), 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 large 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 [8].

We now set the framework that will assist us in the investigation of the large time behaviour of (5.2.1), in a relatively general case. The natural Hilbert space to consider this problem is $L^2(\mathbb{T})^{\otimes 2}$, 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 (5.1.1) and (5.2.1) are (only) hypocoercive, the symmetric part of their generators (i.e. the operators on their r.h.s.) are not coercive on $L^2(\mathbb{T})^{\otimes 2}$. Hence, the standard L^2–norm cannot serve as a usable Lyapunov functional. As is typical for hypocoercive equations (see [25, 13, 1]), a possible remedy to this problem is to consider a "twisted" norm (often also referred to as *entropy functional*), constructed in a way that this functional strictly decays along each trajectory $(u(t), v(t))$.

The following functional, which will be our entropy functional, is not an ansatz, and its origin will be derived in §5.4. Moreover, we will show that it will yield the sharp exponential decay for constant σ, when one chooses $\theta = \theta(\sigma)$ appropriately.

Definition 5.2.1. Let $f, g \in L^2(\mathbb{T})$ and let $\theta > 0$ be given. Then we define the *entropy* $E_\theta(f, g)$ as

$$E_\theta(f,g) := \|f\|^2 + \|g\|^2 - \frac{\theta}{2\pi}\int_0^{2\pi} \mathrm{Re}\left(\partial_x^{-1} f(x)\overline{g(x)}\right) dx. \tag{5.2.2}$$

Here, the *anti-derivative* of f is defined as

$$\partial_x^{-1} f(x) := \int_0^x f(y)dy - \left(\int_0^x f(y)dy\right)_{\mathrm{avg}}, \tag{5.2.3}$$

with the average defined in (5.1.2). The normalization constant in (5.2.3) is chosen such that $(\partial_x^{-1} f)_{\mathrm{avg}} = 0$.

Several recent studies (like [13, 1]) considered the Goldstein-Taylor system with constant σ. This case can be investigated fairly easy as one is able to utilise Fourier analysis in this setting, and construct a Lyapunov functional as a sum of quadratic functionals of the Fourier modes. However, 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 nigh 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, yielding quantitative estimates for the convergence. As the nature of this approach is perturbative, our decay rates are not optimal. The methodology itself, however, 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 paper, with the use of the vector notation

$$f(t) := \begin{pmatrix} f_+(t) \\ f_-(t) \end{pmatrix}, \quad f_0 := \begin{pmatrix} f_{+,0} \\ f_{-,0} \end{pmatrix}, \tag{5.2.4}$$

is the following:

Theorem 5.2.2. *Let $u, v \in C([0,\infty); L^2(\mathbb{T}))$ be mild*[2] *real valued solutions to* (5.2.1) *with initial datum u_0, $v_0 \in L^2(\mathbb{T})$. Denoting by $u_{\mathrm{avg}} = (u_0)_{\mathrm{avg}}$ we have:*

a) *If $\sigma(x) = \sigma$ is constant we have that:*
If $\sigma \neq 2$ then

$$E_{\theta(\sigma)}\left(u(t) - u_{\mathrm{avg}}, v(t)\right) \le E_{\theta(\sigma)}\left(u_0 - u_{\mathrm{avg}}, v_0\right) e^{-2\mu(\sigma)t}$$

where

$$\theta(\sigma) := \begin{cases} \sigma, & 0 < \sigma < 2 \\ \frac{4}{\sigma}, & \sigma > 2 \end{cases}, \quad \mu(\sigma) := \begin{cases} \frac{\sigma}{2}, & 0 < \sigma < 2 \\ \frac{\sigma}{2} - \sqrt{\frac{\sigma^2}{4} - 1}, & \sigma > 2 \end{cases},$$

and if $\sigma = 2$ then for any $0 < \epsilon < 1$

$$E_{\frac{2(2-\epsilon^2)}{2+\epsilon^2}}\left(u(t) - u_{\mathrm{avg}}, v(t)\right) \le E_{\frac{2(2-\epsilon^2)}{2+\epsilon^2}}\left(u_0 - u_{\mathrm{avg}}, v_0\right) e^{-2(1-\epsilon)t}.$$

[2] We use *mild solution* in the terminology of semigroup theory [20].

Consequently if $\sigma \neq 2$

$$\left\| f(t) - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| \leq C_\sigma \left\| f_0 - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| e^{-\mu(\sigma) t}, \tag{5.2.5}$$

where

$$C_\sigma := \begin{cases} \sqrt{\frac{2+\sigma}{2-\sigma}}, & 0 < \sigma < 2 \\ \sqrt{\frac{\sigma+2}{\sigma-2}}, & \sigma > 2 \end{cases}, \quad f_\infty = \frac{u_{\text{avg}}}{2}, \tag{5.2.6}$$

and the decay rate $\mu(\sigma)$ *is sharp.*
For $\sigma = 2$ *we have that*

$$\left\| f(t) - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| \leq \frac{\sqrt{2}}{\epsilon} \left\| f_0 - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| e^{-(1-\epsilon) t}. \tag{5.2.7}$$

b) *If $\sigma(x)$ is non-constant such that*

$$0 < \sigma_{\min} := \inf_{x \in \mathbb{T}} \sigma(x) < \sup_{x \in \mathbb{T}} \sigma(x) =: \sigma_{\max} < \infty,$$

then by defining

$$\theta^* := \min\left(\sigma_{\min}, \frac{4}{\sigma_{\max}}\right) \tag{5.2.8}$$

and

$$\alpha^* := \alpha^*(\sigma_{\min}, \sigma_{\max}) := \begin{cases} \frac{\sigma_{\min}\left(4 + 2\sqrt{4 - \sigma_{\min}^2} - \sigma_{\min}\sigma_{\max}\right)}{4 + 2\sqrt{4 - \sigma_{\min}^2} - \sigma_{\min}^2}, & \sigma_{\min} < \frac{4}{\sigma_{\max}} \\ \sigma_{\max} - \sqrt{\sigma_{\max}^2 - 4}, & \sigma_{\min} \geq \frac{4}{\sigma_{\max}} \end{cases} \tag{5.2.9}$$

we have that

$$E_{\theta^*}\left(u(t) - u_{\text{avg}}, v(t)\right) \leq E_{\theta^*}\left(u_0 - u_{\text{avg}}, v_0\right) e^{-\alpha^* t},$$

and as result

$$\left\| f(t) - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| \leq \sqrt{\frac{2 + \theta_*}{2 - \theta_*}} \left\| f_0 - \begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix} \right\| e^{-\frac{\alpha^*}{2} t}, \tag{5.2.10}$$

with f_∞ *defined in* (5.2.6).

Part (a) of this theorem will be proved in §5.4.4, and Part (b) in §5.5. In many of the proofs which will eventually lead to the proof of this theorem we will assume that (u, v) is a classical solution, pertaining to u_0, v_0 in the periodic Sobolev space $H^1(\mathbb{T})$. The general result will follow by a simple density argument.

Remark 5.2.3. It is simple to see that if $\sigma(x)$ satisfies the conditions of (b), then, as $\sigma_{\min}$ and $\sigma_{\max}$ approach a positive constant $\sigma \neq 2$, we find that

$$\theta^* \to \min\left(\sigma, \frac{4}{\sigma}\right), \quad \text{and} \quad \alpha^* \to \begin{cases} \sigma - \sqrt{\sigma^2 - 4}, & \sigma > 2 \\ \sigma, & \sigma < 2 \end{cases},$$

recovering the results of part (a) of the above theorem.
In addition, one should note that when $\sigma_{\text{min}} > \frac{4}{\sigma_{\text{max}}}$ we have that

$$\alpha^*(\sigma_{\text{min}}, \sigma_{\text{max}}) = 2\mu(\sigma_{\text{max}}),$$

where $\mu(\sigma)$ was defined in part (a) of the Theorem. This validates the intuition that, if σ_{max} is "large enough", the convergence rate of the solution can be estimated using the "worst convergence rate", corresponding to $\mu(\sigma_{\text{max}})$ of the $\sigma(x) = \sigma$ case.
Lastly, one notices that when $\sigma_{\text{min}} = \frac{4}{\sigma_{\text{max}}}$

$$\frac{\sigma_{\text{min}}\left(4+2\sqrt{4-\sigma_{\text{min}}^2}-\sigma_{\text{min}}\sigma_{\text{max}}\right)}{4+2\sqrt{4-\sigma_{\text{min}}^2}-\sigma_{\text{min}}^2} = \sigma_{\text{max}} - \sqrt{\sigma_{\text{max}}^2-4},$$

which shows the continuity of α^* on the curve that stitches the two formulas in (5.2.9).

5.3 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 of this section will be deferred to §5.7.

We begin with the well known Poincaré inequality:

Lemma 5.3.1 (Poincaré Inequality). *Let $f \in H^1_{per}(\mathbb{T})$ with $f_{\text{avg}} = 0$. Then*

$$\|f\| \le \|f'\|. \tag{5.3.1}$$

Next we focus our attention on some simple, yet crucial, properties of the anti-derivative function which was defined in (5.2.3).

Lemma 5.3.2. *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

$$\widehat{\partial_x^{-1} f}(k) = \begin{cases} \frac{\hat{f}(k)}{ik}, & k \neq 0 \\ 0, & k = 0 \end{cases}. \tag{5.3.2}$$

Remark 5.3.3. (ii), (iv), and the fact that f is a function on the torus, imply that if $f_{\text{avg}} = 0$ we are allowed to use integration by parts with $\partial_x^{-1} f(x)$ on this boundaryless manifold without qualms.

The last simple lemma in this revolves around our newly defined functional, E_θ.

Lemma 5.3.4. *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* (5.2.2), *satisfies*

$$E_\theta(f,g) \le \left(1 + \frac{|\theta|}{2}\right)\left(\|f\|^2 + \|g\|^2\right). \tag{5.3.3}$$

If in addition $|\theta| < 2$ we have that

$$E_\theta(f,g) \ge \left(1 - \frac{|\theta|}{2}\right)\left(\|f\|^2 + \|g\|^2\right). \tag{5.3.4}$$

In particular, if $0 \le \theta < 2$ we have that

$$\left(1 - \frac{\theta}{2}\right)\left(\|f\|^2 + \|g\|^2\right) \le E_\theta(f,g) \le \left(1 + \frac{\theta}{2}\right)\left(\|f\|^2 + \|g\|^2\right). \tag{5.3.5}$$

Lastly, we shall prove the following theorem, which (finally) brings the system (5.2.1) into play, and on which we will rely on frequently in our future estimation.

Proposition 5.3.5. *Let $u, v \in C([0,\infty); L^2(\mathbb{T}))$ be (real valued) mild solutions to* (5.2.1) *with initial datum $u_0,\ v_0 \in L^2(\mathbb{T})$. Then for any $\theta \in \mathbb{R}$*

$$\begin{aligned}\frac{d}{dt}E_\theta\left(u(t) - u_{\text{avg}}, v(t)\right) = &-\theta\left\|u(t) - u_{\text{avg}}\right\|^2 + \frac{1}{2\pi}\int_0^{2\pi}(\theta - 2\sigma(x))\, v(x,t)^2 dx \\ &+ \frac{\theta}{2\pi}\int_0^{2\pi}\sigma(x)\partial_x^{-1}\left(u(x,t) - u_{\text{avg}}\right)v(x,t)dx - \theta\left(v(t)_{\text{avg}}\right)^2,\end{aligned} \tag{5.3.6}$$

where

$$u_{\text{avg}} = \frac{1}{2\pi}\int_0^{2\pi} u_0(x)dx = \frac{1}{2\pi}\int_0^{2\pi} u(x,t)dx, \quad \forall t > 0. \tag{5.3.7}$$

Proof. We begin by noticing that the validity of (5.3.7) follows immediately from the fact that u is a mild solution and the conservation of mass property of the system (5.2.1). 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 $t \ge 0$. With this observation in mind, we can assume without loss of generality that $u_{\text{avg}} = 0$.

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

$$\frac{d}{dt}\left(\|u(t)\|^2 + \|v(t)\|^2\right) = -\frac{1}{\pi}\int_0^{2\pi}\sigma(x)v(x,t)^2 dx. \tag{5.3.8}$$

We now turn our attention to the mixed term of $E_\theta(u,v)$:

$$\begin{aligned}
&\frac{d}{dt}\frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1}u(x,t)v(x,t)dx\\
&=\frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1}\left(\partial_t u\right)(x,t)v(x,t)dx+\frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1}u(x,t)\partial_t v(x,t)dx\\
&=-\frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1}\left(\partial_x v\right)(x,t)v(x,t)dx-\frac{\theta}{2\pi}\int_0^{2\pi}\partial_x^{-1}u(x,t)[\partial_x u(x,t)+\sigma(x)v(x,t)]dx.
\end{aligned}$$

Using points (ii) and (iii) of Lemma 5.3.2, together with Remark 5.3.3, we find that the above equals

$$\begin{aligned}
-\frac{\theta}{2\pi}\int_0^{2\pi}\left(v(x,t)-v(t)_{\text{avg}}\right)v(x,t)dx+\frac{\theta}{2\pi}\int_0^{2\pi}u(x,t)^2dx\\
-\frac{\theta}{2\pi}\int_0^{2\pi}\sigma(x)\partial_x^{-1}u(x,t)v(x,t)dx.
\end{aligned}$$

Subtracting this from (5.3.8) (as there is a minus in definition (5.2.2)) yields (5.3.6). □

5.4 Constant relaxation function

In recent years, the investigation of the Goldstein-Taylor model on $\mathbb{T}$ with constant relaxation function σ was frequently tackled with a modal decomposition in the Fourier space w.r.t. x. This approach allows for an extension to other discrete velocity models and even some continuous velocities models [1], but is not suitable for the non-homogeneous case.
Before beginning with our investigation we review a few recent results:
In [13, §1.4] exponential convergence to equilibrium was shown, but without the sharp rate. In [1, §4.1] a hypocoercive decay estimate of the form

$$\left\| f(t)-\begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix}\right\|_{L^2}\le c\,e^{-\mu t}\left\| f_0-\begin{pmatrix} f_\infty \\ f_\infty \end{pmatrix}\right\|_{L^2},$$

with the vector notation from (5.2.4) and the sharp rate

$$\mu(\sigma)=\begin{cases}\frac{\sigma}{2}, & 0<\sigma<2\\ \frac{\sigma}{2}-\sqrt{\frac{\sigma^2}{4}-1}, & \sigma>2\end{cases}$$

was obtained (see also Fig. 5.2 below). A further study on the minimal constant c in the above was provided in [3, Th. 1.1].

With these results in mind, we turn our attention to the following (recast) Goldstein-Taylor equation with a constant relaxation rate:

$$\begin{aligned}
\partial_t u(x,t)&=-\partial_x v(x,t),\\
\partial_t v(x,t)&=-\partial_x u(x,t)-\sigma v(x,t)\,.
\end{aligned}\tag{5.4.1}$$

In order to be able to discover our entropy functional, we shall consider the straightforward modal analysis in detail. This will allow us to obtain not only explicit decay rates for each Fourier

mode, but also an "optimal Lyapunov functional" for such given mode, with which we will then be able to construct a non-modal entropy functional in terms of a pseudo-differential operator as defined in (5.2.2).
As was mentioned in §5.2, this will give us intuition to the large time behaviour of the equation in several cases even when $\sigma(x)$ is not constant.

5.4.1 Fourier analysis and the spectral gap

One natural way to understand the large time behaviour of (5.4.1) relies on a simple Fourier analysis *together with* 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 variable), we see that (5.4.1) is equivalent to infinity many decoupled ODE systems:

$$\frac{d}{dt}\begin{pmatrix}\widehat{u}(k)\\ \widehat{v}(k)\end{pmatrix} = -\begin{pmatrix}0 & ik\\ ik & \sigma\end{pmatrix}\begin{pmatrix}\widehat{u}(k)\\ \widehat{v}(k)\end{pmatrix} := -\mathbf{C}_k\begin{pmatrix}\widehat{u}(k)\\ \widehat{v}(k)\end{pmatrix}, \qquad k\in\mathbb{Z}. \tag{5.4.2}$$

The eigenvalues of the matrices $\mathbf{C}_k \in \mathbb{C}^{2\times 2}$ 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

$$\mathbf{C}_0 = \begin{pmatrix}0 & 0\\ 0 & \sigma\end{pmatrix} \tag{5.4.3}$$

  we can conclude immediately that $\widehat{u}(0,t) = \widehat{u_0}(0)$ and $\widehat{v}(0,t) = \widehat{v_0}(0)e^{-\sigma t}$, 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 $0<|k|<\frac{\sigma}{2}$ one finds two real eigenvalues, whose minimum 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 large 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 $\frac{\sigma}{2}$. Moreover, that eigenvalue is defective (i.e. corresponds to a Jordan block of size 2) and the large time behaviour of $\widehat{u}(k)$ and $\widehat{v}(k)$ is controlled by $(1+t)\,e^{-\frac{\sigma}{2}t}$.

- *Case III:* For $|k|>\frac{\sigma}{2}$, one finds two complex conjugate eigenvalues, whose real part equals $\frac{\sigma}{2}$. Thus the large time behaviour of $\widehat{u}(k)$ and $\widehat{v}(k)$ is controlled by $e^{-\frac{\sigma}{2}t}$.

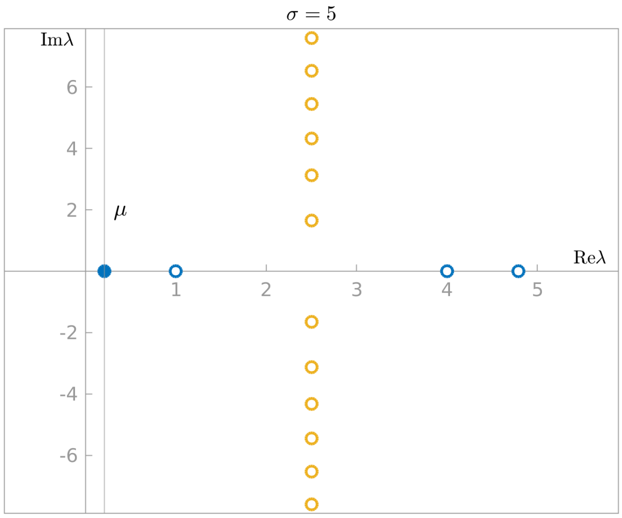

Figure 5.1: The eigenvalues $\lambda_{\pm,k}$ of $\mathbf{C}_k$, $|k| \in \mathbb{N}$ for $\sigma = 5$. The spectral gap is $\mu = (5-\sqrt{21})/2$.

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. This spectral gap of (5.4.1) will be denoted by $\mu(\sigma)$.
Case I, i.e. $0 < |k| < \frac{\sigma}{2}$, 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 $0 < |k| < \frac{\sigma}{2}$, 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 possibilities of Case III, implying that all modes are controlled by $e^{-\frac{\sigma}{2}t}$.
- $\sigma = 2$: We have possibilities of 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 rate one would find is $e^{-(1-\epsilon)t}$ for any given fixed $\epsilon > 0$.
- $\sigma > 2$: We have possibilities from 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. $\frac{\sigma}{2} \in \mathbb{N}\backslash\{1\}$, 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 $e^{-\left(\frac{\sigma}{2}-\sqrt{\frac{\sigma^2}{4}-1}\right)t}$, a decay rate that is realised on the $k = \pm 1$ modes, and the coefficient in the exponent is the spectral gap of the Goldstein-Taylor system (5.4.1).

An illustration of the eigenvalues of the matrices $\mathbf{C}_k$ for $|k| \in \mathbb{N}$ and $\sigma = 5$ can be viewed in Fig.

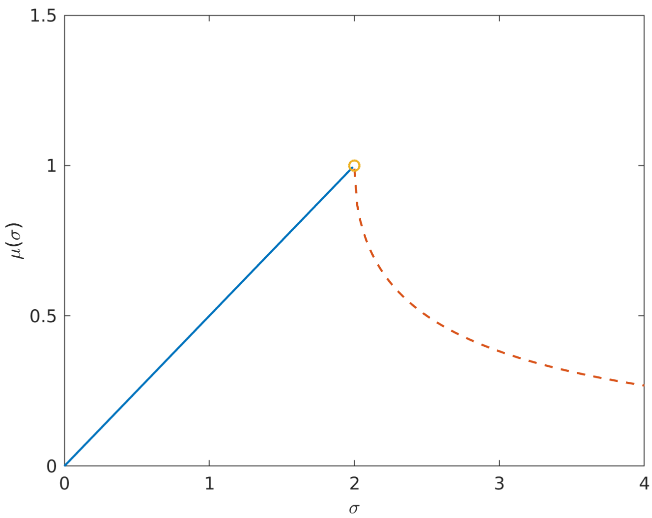

Figure 5.2: The exponential decay rate, $\mu(\sigma)$, of the solution pair $(u(t)-u_{\text{avg}}, v(t))$ grows linearly until $\sigma = 2$ where the defectiveness appears (hence the circle). From that point onwards the decay rate decreases, and is of order $O\left(\frac{1}{\sigma}\right)$.

5.2. Before we turn our attention to properly consider these cases and “uncover” our spatial entropy, we remind the reader of the hypocoercivity technique which will allow us to transform the spectral information of $\mathbf{C}_k$ into a an appropriate, twisted norm with which we will show the desired decay of the k-th mode.

5.4.2 Hypocoercivity and modal Lyapunov functionals

In the previous subsection we have concluded that, barring the zero mode, all the Fourier modes of (5.4.2) decay exponentially (excluding potentially those with $|k| = \frac{\sigma}{2}$ where a polynomial correction is required). The lack of positive definiteness of the governing matrix, $\mathbf{C}_k$, stops us from seeing this behaviour in the Euclidean norm on $\mathbb{C}^2$. However, by modifying the norm with the help of another, closely related, positive definite matrix $\mathbf{P}_k$, one can construct a new Lyapunov functional, which is equivalent to the Euclidean norm, that decays with the expected exponential rate (at least for a non-defective $\mathbf{C}_k$).
This is exactly the idea that motivated Arnold and Erb, and which is expressed in the following theorem (see [6], [1, Lemma 2]):

Theorem 5.4.1. *Let the matrix $\boldsymbol{C} \in \mathbb{C}^{n\times n}$ be positive stable (i.e. have only eigenvalues with positive real parts). Let*

$$\mu = \min\left\{\operatorname{Re}\lambda \mid \lambda \textit{ is an eigenvalue of } \boldsymbol{C}\right\}.$$

Then:

i) If all eigenvalues with real part equal to μ are non-defective, there exists a Hermitian, positive definite matrix $\mathbf{P}$ such that

$$\boldsymbol{C}^*\mathbf{P} + \mathbf{P}\boldsymbol{C} \geq 2\mu\mathbf{P}. \tag{5.4.4}$$

ii) If at least one eigenvalue with real part equal to μ is defective, then for any $\epsilon > 0$, one can find a Hermitian, positive definite matrix $\mathbf{P}_\epsilon$ such that

$$\boldsymbol{C}^* \mathbf{P}_\epsilon + \mathbf{P}_\epsilon \boldsymbol{C} \geq 2\left(\mu - \epsilon\right) \mathbf{P}_\epsilon\,, \tag{5.4.5}$$

where $\boldsymbol{C}^*$ denotes the Hermitian transpose of $\boldsymbol{C}$.

We remark that the matrices $\mathbf{P}$ and $\mathbf{P}_\epsilon$ are never unique.
One can utilise the theorem in the following way: Assuming the eigenvalues associated to $\mathbf{C}$'s spectral gap, μ, are non-defective, then by defining the norm

$$\left\| y \right\|_{\mathbf{P}}^2 := \left\langle y, \mathbf{P} y \right\rangle = y^* \mathbf{P} y,$$

one sees that, if $y(t)$ solves the ODE $\dot{y} = -\mathbf{C} y$, then

$$\frac{d}{dt} \left\| y \right\|_{\mathbf{P}}^2 = -\left\langle y, \left(\mathbf{C}^* \mathbf{P} + \mathbf{P}\mathbf{C}\right) y \right\rangle \leq -2\mu \left\| y \right\|_{\mathbf{P}}^2, \tag{5.4.6}$$

resulting in the correct decay rate. The same approach works in the second case of Theorem 5.4.1.
Besides the general idea of this methodology, Arnold and Erb have given a recipe (one that was later extended in [7] to defective cases, using a time dependent matrix $\mathbf{P}$) to finding the matrix $\mathbf{P}, \mathbf{P}_\epsilon$:
Assuming that $\mathbf{C}$ is diagonalisable, and letting $\{\omega_i\}_{i=1,\dots,n}$ be the eigenvectors of $\mathbf{C}^*$, the matrix $\mathbf{P} > 0$ can be chosen to be

$$\mathbf{P} = \sum_{i=1}^{n} b_i \omega_i \otimes \omega_i^*, \tag{5.4.7}$$

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 $\mathbf{C}$ 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 $\mathbf{C}$ non-defective, all matrices $\mathbf{P}$ satisfying (5.4.4) are indeed of the form (5.4.7), see [3, Lemma 3.1].

We now turn our attention back to the Fourier transformed Goldstein-Taylor system (5.4.2) and determine the modal Lyapunov functionals using the above recipe. A short computation, where the weights b_1, b_2 are chosen such that both diagonal elements of $\mathbf{P}$ are 1, finds the following matrices (For Case III we also require $b_1 = b_2$, as this minimises the number of the resulting admissible matrices $\mathbf{P}_k$ satisfying (5.4.4).):

- Case I: $0 < |k| < \frac{\sigma}{2}$. In this case we have:

$$\mathbf{P}_k^{(I)} := \begin{pmatrix} 1 & -\frac{2ki}{\sigma} \\ \frac{2ki}{\sigma} & 1 \end{pmatrix}, \tag{5.4.8}$$

- Case II: $|k| = \frac{\sigma}{2} \in \mathbb{N}$. As this case fosters defective eigenvalues, we will only consider the case $\sigma = 2$ (as was mentioned beforehand), and state the matrix corresponding to $k = \pm 1$ and a given fixed $\epsilon > 0$:

$$\mathbf{P}_{\epsilon,\pm 1}^{(II)} := \begin{pmatrix} 1 & \mp \frac{i(2-\epsilon^2)}{2+\epsilon^2} \\ \pm \frac{i(2-\epsilon^2)}{2+\epsilon^2} & 1 \end{pmatrix} \tag{5.4.9}$$

- Case III: $|k| > \frac{\sigma}{2}$. In this case we have:

$$\mathbf{P}_k^{(III)} := \begin{pmatrix} 1 & -\frac{i\sigma}{2k} \\ \frac{i\sigma}{2k} & 1 \end{pmatrix} \tag{5.4.10}$$

For each mode $k \neq 0$, its *modal Lyapunov functional* will be given by $\|\binom{\hat{u}(k,t)}{\hat{v}(k,t)}\|^2_{\mathbf{P}_k}$, where the matrix $\mathbf{P}_k$ is chosen according to the above three cases. In Case II, the parameter $\epsilon > 0$ can be chosen arbitrarily small.

5.4.3 Derivation of the spatial entropy $E_\theta(u, v)$

The goal of this subsection is twofold: Finding a modal entropy to our system, and translating it to a spatial entropy that is modal-independent.
To begin with we shall define a *modal entropy* to quantify the exponential decay of solutions to (5.4.2) towards its steady state:

$$\widehat{u_\infty}(k) = \begin{cases} \widehat{u_0}(k=0) = (u_0)_{\text{avg}}, & k = 0 \\ 0, & k \neq 0 \end{cases} \quad ; \qquad \widehat{v_\infty}(k) = 0\ ,\ k \in \mathbb{Z}\ . \tag{5.4.11}$$

Since the matrix $\mathbf{C}_0$ from (5.4.3) has no spectral gap, the mode $k = 0$ plays a special role, and hence will be treated separately.
Once found, we will want to relate that modal-based entropy to the *spatial entropy* E_θ from Definition 5.2.1, which is not based on a modal decomposition. To this end we already remark that the off-diagonal factors ik in (5.4.8) and $1/ik$ in (5.4.10) correspond in physical space, roughly speaking, to a first derivative and an anti-derivative, respectively.

As in §5.4.1 we shall distinguish three cases of σ:
$\mathbf{0 < \sigma < 2}$: All modes $k \neq 0$ satisfy $|k| > \frac{\sigma}{2}$, and hence are of Case III. We recall from §5.4.1 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 $\|\binom{\hat{u}(k,t)}{\hat{v}(k,t)}\|^2_{\mathbf{P}_k}$, where $\mathbf{P}_k$ satisfies the inequality (5.4.4) with $\mu = \frac{\sigma}{2}$. $\mathbf{P}_k = \mathbf{P}_k^{(III)}$ is the most convenient choice.
We define the modal entropy for any $\{\hat{u}(k), \hat{v}(k)\}_{k\in\mathbb{Z}}$ such that $\hat{u}(0) = 0$ as

$$\mathscr{E}(\hat{u}, \hat{v}) := \sum_{k\in\mathbb{Z}\setminus\{0\}} \left\| \begin{pmatrix} \hat{u}(k) \\ \hat{v}(k) \end{pmatrix} \right\|^2_{\mathbf{P}_k^{(III)}} + \left\| \begin{pmatrix} \hat{u}(0) \\ \hat{v}(0) \end{pmatrix} \right\|^2 \tag{5.4.12}$$

$$= \sum_{k\in\mathbb{Z}} \left(|\hat{u}(k)|^2 - \sigma \operatorname{Re}\left(\frac{\hat{u}(k)}{ik} \overline{\hat{v}(k)} \right) + |\hat{v}(k)|^2 \right), \tag{5.4.13}$$

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 5.3.2, we find that

$$\mathscr{E}(\hat{u}, \hat{v}) = E_\sigma(u, v), \tag{5.4.14}$$

which shows why we consider the spatial entropy functional from Definition 5.2.1 in this case.
We note that, since $u_{\text{avg}}(t)$ is conserved, part (iv) of Lemma 5.3.2, explains why we have chosen

to use the anti-derivative of u, and not of v.

$\sigma > 2$: This situation is more complicated than the previous one, as we have a mixture of at least two of the aforementioned three cases: finitely many k−s in $\mathbb{Z}$ for which $0 < |k| < \frac{\sigma}{2}$ (i.e. Case I), Case II for two k−s if $\frac{\sigma}{2} \in \mathbb{N}$, while the rest satisfy $|k| > \frac{\sigma}{2}$ (i.e. Case III). Following the above methodology to construct the modal entropy, we would need to use a combination of $\mathbf{P}_k^{(I)}$ and $\mathbf{P}_k^{(III)}$, given by (5.4.8) and (5.4.10), 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 §5.4.1 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 $\mathbf{P}_k$, $|k| > 1$. Moreover, using these "optimal" functionals will result in worsening of (i.e. enlargement of) the multiplicative constant in the L^2 hypocoercive estimation (5.2.5). Due to these reasons we will use the matrix

$$\mathbf{P}_k^{\text{suff}} := \mathbf{P}_k^{(III)}\left(\sigma \to \frac{4}{\sigma}\right) = \begin{pmatrix} 1 & -\frac{2i}{k\sigma} \\ \frac{2i}{k\sigma} & 1 \end{pmatrix} > 0 \tag{5.4.15}$$

when $k \neq 0$, which satisfies $\mathbf{P}_{\pm 1}^{\text{suff}} = \mathbf{P}_{\pm 1}^{(I)}$ for the crucial lowest modes. It also satisfies the following result, which implies exponential decay of all modal Lyapunov functionals $\|\binom{\hat{u}(k,t)}{\hat{v}(k,t)}\|^2_{\mathbf{P}_k^{\text{suff}}}$, $k \neq 0$ with rate $2\mu = \sigma - \sqrt{\sigma^2 - 4}$.

Lemma 5.4.2. *Let $\sigma > 2$. Then*

$$\boldsymbol{C}_k^* \mathbf{P}_k^{\text{suff}} + \mathbf{P}_k^{\text{suff}} \boldsymbol{C}_k - 2\mu \mathbf{P}_k^{\text{suff}} \geq 0 \qquad \forall\, k \neq 0\,.$$

The proof of this lemma is straightforward[3]. Proceeding like in (5.4.12) we define the modal entropy for any $\{\widehat{u}(k), \widehat{v}(k)\}_{k\in\mathbb{Z}}$ such that $\widehat{u}(0) = 0$:

$$\mathscr{E}(\widehat{u}, \widehat{v}) := \sum_{k\in\mathbb{Z}\setminus\{0\}} \left\| \begin{pmatrix} \widehat{u}(k) \\ \widehat{v}(k) \end{pmatrix} \right\|^2_{\mathbf{P}_k^{\text{suff}}} + \left\| \begin{pmatrix} \widehat{u}(0) \\ \widehat{v}(0) \end{pmatrix} \right\|^2 .$$

Due to (5.4.14) and (5.4.15) it is related to the spatial entropy functional from Definition 5.2.1 as

$$\mathscr{E}(\widehat{u}, \widehat{v}) = E_{\frac{4}{\sigma}}(u, v)\,.$$

$\sigma = 2$: Just like in the previous case, the lowest frequency modes $k = \pm 1$ control the large time behaviour. However, the matrices $\mathbf{C}_{\pm 1}$ 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 $\epsilon > 0$:

$$\mathbf{P}_{\epsilon,k}^{\text{suff}} = \mathbf{P}_k^{(III)}\left(\sigma \to \frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2}\right) = \begin{pmatrix} 1 & -\frac{i(2-\epsilon^2)}{k(2+\epsilon^2)} \\ \frac{i(2-\epsilon^2)}{k(2+\epsilon^2)} & 1 \end{pmatrix} > 0\,, \tag{5.4.16}$$

[3] In a sense, the same computation that shows this inequality is embedded in the proof of the exponential decay of E_θ in the next subsection.

which satisfies $\mathbf{P}^{\text{suff}}_{\epsilon,\pm 1} = \mathbf{P}^{(II)}_{\epsilon,\pm 1}$ for the crucial lowest model. It also satisfies the following result, which implies exponential decay of all modal Lyapunov functionals $\|\binom{\hat{u}(k,t)}{\hat{v}(k,t)}\|^2_{\mathbf{P}^{\text{suff}}_{\epsilon,k}}$, $k \neq 0$ with rate of at least $2\mu = 2(1-\epsilon)$.

Lemma 5.4.3. *Let $\sigma = 2$. Then*

$$\boldsymbol{C}_k^* \mathbf{P}^{\text{suff}}_{\epsilon,k} + \mathbf{P}^{\text{suff}}_{\epsilon,k} \boldsymbol{C}_k - 2\mu \mathbf{P}^{\text{suff}}_{\epsilon,k} > 0 \qquad \forall\, k \neq 0\,.$$

Proceeding like in (5.4.12) we define the modal entropy for any $\{\widehat{u}(k), \widehat{v}(k)\}_{k\in\mathbb{Z}}$ such that $\widehat{u}(0) = 0$:

$$\mathscr{E}(\widehat{u},\widehat{v}) := \sum_{k\in\mathbb{Z}\setminus\{0\}} \left\| \begin{pmatrix} \widehat{u}(k) \\ \widehat{v}(k) \end{pmatrix} \right\|^2_{\mathbf{P}^{\text{suff}}_{\epsilon,k}} + \left\| \begin{pmatrix} \widehat{u}(0) \\ \widehat{v}(0) \end{pmatrix} \right\|^2 .$$

Due to (5.4.14) and (5.4.16) it is related to the spatial entropy functional from Definition 5.2.1 as

$$\mathscr{E}(\widehat{u},\widehat{v}) = E_{\frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2}}(u,v)\,.$$

5.4.4 The evolution of the spatial entropy E_θ

In the previous subsection we have shown how, depending on the value of σ, the entropies E_σ, $E_{\frac{4}{\sigma}}$ and $E_{\frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2}}$ are the correct candidates to show the exponential convergence to equilibrium. A closer look at (5.4.6) shows that each modal Lyapunov functional $\|\binom{\hat{u}(k,t)}{\hat{v}(k,t)}\|^2_{\mathbf{P}_k}$ decays exponentially, and hence also the spatial entropy E_θ. Recalling the decay rates presented in §5.4.3 for the three regimes of σ, confirms that we have actually already proved most of part (a) of Theorem 5.2.2. 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 §5.5.

Theorem 5.4.4. *Under the same conditions of Theorem 5.2.2 with $\sigma(x) = \sigma$, one has that*

i) If $0 < \sigma < 2$ then

$$E_\sigma\left(u(t) - u_{\text{avg}}, v(t)\right) \leq E_\sigma\left(u_0 - u_{\text{avg}}, v_0\right) e^{-\sigma t}.$$

ii) If $\sigma > 2$ then

$$E_{\frac{4}{\sigma}}\left(u(t) - u_{\text{avg}}, v(t)\right) \leq E_{\frac{4}{\sigma}}\left(u_0 - u_{\text{avg}}, v_0\right) e^{-\left(\sigma - \sqrt{\sigma^2-4}\right)t}.$$

iii) If $\sigma = 2$ then for any $0 < \epsilon < 1$

$$E_{\frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2}}\left(u(t) - u_{\text{avg}}, v(t)\right) \leq E_{\frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2}}\left(u_0 - u_{\text{avg}}, v_0\right) e^{-2(1-\epsilon)t}.$$

Proof. In order to prove this theorem we shall obtain differential inequalities for E_θ, from which we will conclude the desired result by a simple application of Gronwall's inequality. Using Proposition 5.3.5 we find that:
<u>If $0<\sigma<2$:</u>

$$\begin{aligned}\frac{d}{dt}E_\sigma\left(u(t)-u_{\text{avg}},v(t)\right)&=-\sigma\left\|u(t)-u_{\text{avg}}\right\|^2-\sigma\left\|v(t)\right\|^2\\&+\frac{\sigma^2}{2\pi}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx-\sigma\left(v(t)_{\text{avg}}\right)^2\\=-\sigma E_\sigma\left(u(t)-u_{\text{avg}},v(t)\right)&-\sigma\left(v(t)_{\text{avg}}\right)^2\le-\sigma E_\sigma\left(u(t)-u_{\text{avg}},v(t)\right).\end{aligned}$$

Note that, since $v_{\text{avg}}(t)=(v_0)_{\text{avg}}e^{-\sigma t}$, we can compute $E_\theta\left(u(t)-u_{\text{avg}},v(t)\right)$ explicitly.
<u>If $\sigma>2$:</u>

$$\begin{aligned}\frac{d}{dt}E_{\frac{4}{\sigma}}\left(u(t)-u_{\text{avg}},v(t)\right)&=-\frac{4}{\sigma}\left\|u(t)-u_{\text{avg}}\right\|^2-\left(2\sigma-\frac{4}{\sigma}\right)\left\|v(t)\right\|^2\\&+\frac{4}{2\pi}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx-\frac{4}{\sigma}\left(v(t)_{\text{avg}}\right)^2\\\le-\left(\sigma-\sqrt{\sigma^2-4}\right)E_{\frac{4}{\sigma}}&\left(u(t)-u_{\text{avg}},v(t)\right)+\left(\sigma-\sqrt{\sigma^2-4}-\frac{4}{\sigma}\right)\left\|u(t)-u_{\text{avg}}\right\|^2\\+\left(\frac{4}{\sigma}-\sigma-\sqrt{\sigma^2-4}\right)\left\|v(t)\right\|^2&+\frac{4}{2\pi}\left(1-\frac{\sigma-\sqrt{\sigma^2-4}}{\sigma}\right)\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx.\end{aligned}$$

The desired inequality, $\frac{d}{dt}E_{\frac{4}{\sigma}}\le-\left(\sigma-\sqrt{\sigma^2-4}\right)E_{\frac{4}{\sigma}}$, is valid if and only if

$$\begin{aligned}&\frac{4}{2\pi}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx\\&\le\left(\sigma-\sqrt{\sigma^2-4}\right)\left\|u(t)-u_{\text{avg}}\right\|^2+\left(\sigma+\sqrt{\sigma^2-4}\right)\left\|v(t)\right\|^2.\end{aligned}\tag{5.4.17}$$

Cauchy-Schwarz inequality, together with Poincaré inequality (Lemma 5.3.1) and Lemma 5.3.2, imply that

$$\begin{aligned}&\frac{4}{2\pi}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx\le4\left\|u(t)-u_{\text{avg}}\right\|\left\|v(t)\right\|\\&=2\left(\sqrt{\sigma-\sqrt{\sigma^2-4}}\left\|u(t)-u_{\text{avg}}\right\|\right)\left(\sqrt{\sigma+\sqrt{\sigma^2-4}}\left\|v(t)\right\|\right).\end{aligned}$$

Together with the fact that $2|ab|\le a^2+b^2$ this shows (5.4.17), concluding the proof in this case.
<u>If $\sigma=2$:</u>

$$\begin{aligned}\frac{d}{dt}E_{\frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2}}\left(u(t)-u_{\text{avg}},v(t)\right)&=-\frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2}\left\|u(t)-u_{\text{avg}}\right\|^2-\frac{2\left(2+3\epsilon^2\right)}{2+\epsilon^2}\left\|v(t)\right\|^2\\+\frac{1}{2\pi}\cdot\frac{4\left(2-\epsilon^2\right)}{2+\epsilon^2}&\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx-\frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2}\left(v(t)_{\text{avg}}\right)^2\\\le-2(1-\epsilon)E_{\frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2}}&\left(u(t)-u_{\text{avg}},v(t)\right)-2\epsilon\left(1-\frac{2\epsilon}{2+\epsilon^2}\right)\left\|u(t)-u_{\text{avg}}\right\|^2\\-2\epsilon\left(1+\frac{2\epsilon}{2+\epsilon^2}\right)\left\|v(t)\right\|^2&+\frac{1}{2\pi}\cdot\frac{4\epsilon\left(2-\epsilon^2\right)}{2+\epsilon^2}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)dx.\end{aligned}$$

Like before, the desired inequality will follow if

$$\frac{1}{2\pi}\cdot\frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)\,dx$$
$$\le\left(1-\frac{2\epsilon}{2+\epsilon^2}\right)\left\|u(t)-u_{\text{avg}}\right\|^2+\left(1+\frac{2\epsilon}{2+\epsilon^2}\right)\|v(t)\|^2.$$

This is valid since

$$\frac{1}{2\pi}\cdot\frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2}\int_0^{2\pi}\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)v(x,t)\,dx$$
$$\le\frac{2\sqrt{4+\epsilon^4}}{2+\epsilon^2}\left\|u(t)-u_{\text{avg}}\right\|\|v(t)\|=2\left(\sqrt{1-\frac{2\epsilon}{2+\epsilon^2}}\left\|u(t)-u_{\text{avg}}\right\|\right)\left(\sqrt{1+\frac{2\epsilon}{2+\epsilon^2}}\|v(t)\|\right)$$
$$\le\left(1-\frac{2\epsilon}{2+\epsilon^2}\right)\left\|u(t)-u_{\text{avg}}\right\|^2+\left(1+\frac{2\epsilon}{2+\epsilon^2}\right)\|v(t)\|^2,$$

where we used Cauchy-Schwarz inequality, Poincaré inequality, and Lemma 5.3.2 again. The theorem is now complete. □

As the last part of this section, we finally prove part (a) of Theorem 5.2.2:

Proof of part (a) *of Theorem 5.2.2.* The decay estimates of $E_{\theta(\sigma)}$ are already shown in Theorem 5.4.4. To show (5.2.5) and (5.2.7) we recall that

$$f_+=\frac{u+v}{2},\quad f_-=\frac{u-v}{2},$$

and

$$\|f\|^2+\|g\|^2\le\frac{2}{2-\theta}E_\theta\left(f,g\right),\quad E_\theta\left(f,g\right)\le\frac{2+\theta}{2}\left(\|f\|^2+\|g\|^2\right)$$

for $0<\theta<2$ and $f_{\text{avg}}=0$, according to Lemma 5.3.4. Thus, using the definition of f_∞ from (5.2.6) we see that

$$\left\|f_+(t)-f_\infty\right\|^2+\left\|f_-(t)-f_\infty\right\|^2$$
$$=\frac{1}{2}\left\|u(t)-u_{\text{avg}}\right\|^2+\frac{1}{2}\|v(t)\|^2\le\frac{1}{2-\theta}E_\theta\left(u(t)-u_{\text{avg}},v(t)\right)$$
$$\le\frac{1}{2-\theta}E_\theta\left(u_0-u_{\text{avg}},v_0\right)e^{-2\mu(\sigma)t}\le\frac{1}{2}\cdot\frac{2+\theta}{2-\theta}\left(\left\|u_0-u_{\text{avg}}\right\|^2+\|v_0\|^2\right)e^{-2\mu(\sigma)t}$$
$$=\frac{2+\theta}{2-\theta}\left(\left\|f_{+,0}-f_\infty\right\|^2+\left\|f_{-,0}-f_\infty\right\|^2\right)e^{-2\mu(\sigma)t},$$

which shows the result for the appropriate choices of $\theta(\sigma)$ and $\mu(\sigma)$. For $\sigma=2$ we choose

$$\theta(2)=\frac{2\left(2-\epsilon^2\right)}{2+\epsilon^2},\quad\mu(2)=1-\epsilon\,.$$

The sharpness of the decay rate for $\sigma\neq2$ 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.

5.5 x–dependent relaxation function

The large time behaviour of solutions to the Goldstein-Taylor equation (5.1.1), or equivalently its recast form (5.2.1), becomes increasingly harder to understand, if the relaxation function, $\sigma(x)$, is not a constant. However, as shown in §5.4, 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 to deal with more complicated systems, as will be shown in the next section.

A fundamental theorem to establish our main result, Theorem 5.2.2 (b), is the following:

Theorem 5.5.1. *Let $u, v \in C([0,\infty); L^2(\mathbb{T}))$ be mild solutions to* (5.2.1) *with initial datum u_0, $v_0 \in L^2(\mathbb{T})$. Denoting by $u_{\text{avg}} = (u_0)_{\text{avg}}$ we have that for any given $0 < \alpha, \theta < 2$ the conditions*

$$\alpha < \theta, \quad \theta + \alpha < 2\sigma_{\min} \tag{5.5.1}$$

and

$$\sup_{x\in\mathbb{T}} \left(\theta^2 (\sigma(x) - \alpha)^2 - 4(\theta - \alpha)(2\sigma(x) - \theta - \alpha)\right) \leq 0, \tag{5.5.2}$$

imply that

$$E_\theta\left(u(t) - u_{\text{avg}}, v(t)\right) \leq E_\theta\left(u_0 - u_{\text{avg}}, v_0\right) e^{-\alpha t}, \quad t \geq 0. \tag{5.5.3}$$

Proof. Using (5.3.6) from Proposition 5.3.5, and the fact that $\theta\left(v(t)_{\text{avg}}\right)^2 \geq 0$, we find that

$$\begin{aligned} &\frac{d}{dt}E_\theta\left(u(t) - u_{\text{avg}}, v(t)\right) \leq -\alpha E_\theta\left(u(t) - u_{\text{avg}}, v(t)\right) - (\theta - \alpha)\left\|u(t) - u_{\text{avg}}\right\|^2 \\ &-\frac{1}{2\pi}\int_0^{2\pi} (2\sigma(x) - \theta - \alpha) v(x,t)^2 dx + \frac{\theta}{2\pi}\int_0^{2\pi} (\sigma(x) - \alpha)\,\partial_x^{-1}\left(u(x,t) - u_{\text{avg}}\right) v(x,t) dx. \end{aligned} \tag{5.5.4}$$

The proof of the theorem will follow from the above inequality if we can show that

$$\begin{aligned} &\frac{\theta}{2\pi}\int_0^{2\pi} (\sigma(x) - \alpha)\,\partial_x^{-1}\left(u(x,t) - u_{\text{avg}}\right) v(x,t) dx \\ &\leq (\theta - \alpha)\left\|u(t) - u_{\text{avg}}\right\|^2 + \frac{1}{2\pi}\int_0^{2\pi} (2\sigma(x) - \theta - \alpha) v(x,t)^2 dx. \end{aligned} \tag{5.5.5}$$

Due to condition (5.5.1) we have that

$$\inf_{x\in\mathbb{T}} (2\sigma(x) - \theta - \alpha) = 2\sigma_{\min} - \theta - \alpha > 0.$$

Hence, we obtain with Cauchy-Schwarz, Young's inequality $|ab| \leq \frac{a^2}{\theta} + \frac{\theta b^2}{4}$, and the Poincaré inequality, (5.3.1), that

$$\begin{aligned} &\left|\frac{\theta}{2\pi}\int_0^{2\pi} (\sigma(x) - \alpha)\,\partial_x^{-1}\left(u(x,t) - u_{\text{avg}}\right) v(x,t) dx\right| \\ &\leq \frac{\theta}{2\pi}\int_0^{2\pi} \sqrt{2\sigma(x) - \theta - \alpha}\,|v(x,t)|\,\frac{|\sigma(x) - \alpha|}{\sqrt{2\sigma(x) - \theta - \alpha}}\left|\partial_x^{-1}\left(u(x,t) - u_{\text{avg}}\right)\right| dx \end{aligned} \tag{5.5.6}$$

$$\leq \frac{\theta}{2\pi}\left(\int_0^{2\pi}(2\sigma(x)-\theta-\alpha)\,v(x,t)^2dx\right)^{\frac{1}{2}}\left(\int_0^{2\pi}\frac{(\sigma(x)-\alpha)^2}{2\sigma(x)-\theta-\alpha}\left(\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)\right)^2dx\right)^{\frac{1}{2}}$$
$$\leq \frac{1}{2\pi}\int_0^{2\pi}(2\sigma(x)-\theta-\alpha)\,v(x,t)^2dx+\frac{1}{2\pi}\int_0^{2\pi}\frac{\theta^2\,(\sigma(x)-\alpha)^2}{4\,(2\sigma(x)-\theta-\alpha)}\left(\partial_x^{-1}\left(u(x,t)-u_{\text{avg}}\right)\right)^2dx$$
$$\leq \frac{1}{2\pi}\int_0^{2\pi}(2\sigma(x)-\theta-\alpha)\,v(x,t)^2dx+\sup_{x\in\mathbb{T}}\left(\frac{\theta^2\,(\sigma(x)-\alpha)^2}{4\,(2\sigma(x)-\theta-\alpha)}\right)\left\|u(t)-u_{\text{avg}}\right\|^2.$$

The above implies that (5.5.5) will be valid when

$$\sup_{x\in\mathbb{T}}\frac{\theta^2\,(\sigma(x)-\alpha)^2}{4\,(2\sigma(x)-\theta-\alpha)}\leq\theta-\alpha,$$

which is equivalent, due to the positivity of the denominator, to (5.5.2). The proof is thus complete. □

Remark 5.5.2. It is worth to note that the conditions expressed in (5.5.1) are crucial in our estimation. Indeed, they tell us that

$$(\theta-\alpha)\left\|u(t)-u_{\text{avg}}\right\|^2\quad\text{and}\quad\int_0^{2\pi}(2\sigma(x)-\theta-\alpha)\,v(x,t)^2dx$$

are non-negative. If one part of the condition would not be true, we would be able to "cook" initial data such that the mixed u–v–term in (5.5.5) is zero, and the above terms add up to something strictly negative - breaking the functional inequality we are aiming to attain.

The next step towards proving part (b) in Theorem 5.2.2 is to look for θ and α such that conditions (5.5.1) and (5.5.2) are satisfied.
We recall the definition of θ^* from Theorem 5.2.2:

$$\theta^*:=\min\left(\sigma_{\min},\frac{4}{\sigma_{\max}}\right),$$

which in a sense captures the "worst possible" behaviour when comparing $\sigma(x)$ to the constant case (with $\sigma\neq 2$). We show the following:

Lemma 5.5.3. *Assume that* $0<\sigma_{\min}<\sigma_{\max}<\infty$, *where* $\sigma_{\min}$ *and* $\sigma_{\max}$ *were defined in Theorem 5.2.2. Then*

$$\alpha^*:=\alpha^*\left(\sigma_{\min},\sigma_{\max}\right):=\begin{cases}\frac{\sigma_{\min}\left(4+2\sqrt{4-\sigma_{\min}^2}-\sigma_{\min}\sigma_{\max}\right)}{4+2\sqrt{4-\sigma_{\min}^2}-\sigma_{\min}^2}, & \sigma_{\min}<\frac{4}{\sigma_{\max}}\\ \sigma_{\max}-\sqrt{\sigma_{\max}^2-4}, & \sigma_{\min}\geq\frac{4}{\sigma_{\max}}\end{cases}$$

is such that θ^* *and* α^* *satisfy conditions* (5.5.1) *and* (5.5.2).

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 always have that $0 < \theta^* < 2$.
We continue by considering condition (5.5.2), and finding appropriate parameters which will give condition (5.5.1) automatically. Denoting by

$$f(\alpha,\theta,y) := \theta^2 (y-\alpha)^2 - 4(\theta-\alpha)(2y-\theta-\alpha)$$

for (α,θ) that satisfy condition (5.5.1) and $y \in [\sigma_{\min},\sigma_{\max}]$, we find that for fixed α and θ, f is an upward parabola in y whose non-positive part lies between its roots

$$y_\pm(\alpha,\theta) := \alpha + \frac{2(\theta-\alpha)}{\theta^2}\left(2 \pm \sqrt{4-\theta^2}\right).$$

Thus, condition (5.5.2) is satisfied if and only if

$$y_-(\alpha,\theta) \le \sigma_{\min}, \quad \text{and} \quad \sigma_{\max} \le y_+(\alpha,\theta).$$

A simple calculation shows that for $0 < \theta < 2$

$$y_-(\alpha,\theta) \le \sigma_{\min} \iff \alpha \le \frac{\theta\left(2\sqrt{4-\theta^2} - (4-\sigma_{\min}\theta)\right)}{2\sqrt{4-\theta^2} - \left(4-\theta^2\right)} =: \gamma_{\min}(\theta),$$

$$\sigma_{\max} \le y_+(\alpha,\theta) \iff \alpha \le \frac{\theta\left(2\sqrt{4-\theta^2} + (4-\sigma_{\max}\theta)\right)}{2\sqrt{4-\theta^2} + \left(4-\theta^2\right)} =: \gamma_{\max}(\theta).$$

This means that, if we choose $\alpha(\theta)$ for a fixed θ so that condition (5.5.2) is valid, we must have that

$$\alpha(\theta) \le \min\left(\gamma_{\min}(\theta), \gamma_{\max}(\theta)\right).$$

One can continue and show that (see §5.7):

(i) For $\theta \le \sigma_{\min}$ and $0 < \theta < 2$ we have that $\gamma_{\max}(\theta) \le \gamma_{\min}(\theta)$.

(ii) For $\theta \le \frac{4}{\sigma_{\max}}$ and $0 < \theta < \sigma_{\max}$ we have that $0 < \gamma_{\max}(\theta) < \theta$.

With these observations we deduce that for any

$$\theta \in (0,\theta^*] = \left(0, \min\left(\sigma_{\min}, \frac{4}{\sigma_{\max}}\right)\right] \cap (0,2)$$

we have $\theta < \sigma_{\max}$ and hence

$$\gamma_{\max}(\theta) = \min\left(\gamma_{\min}(\theta), \gamma_{\max}(\theta)\right) \quad \text{and} \quad \gamma_{\max}(\theta) < \theta.$$

Hence, the pair $\left(\theta, \alpha = \gamma_{\max}(\theta)\right)$ satisfies not only condition (5.5.2) but also

$$\gamma_{\max}(\theta) + \theta < 2\theta \le 2\theta^* \le 2\sigma_{\min} \quad \text{and} \quad \gamma_{\max}(\theta) < \theta,$$

i.e. condition (5.5.1). We conclude that θ and $\alpha = \gamma_{\max}(\theta)$ satisfy both desired conditions, for any $\theta \in (0, \theta^*]$.
Noticing that

$$\gamma_{\max}\left(\theta^*\right) = \left\{\begin{array}{ll} \dfrac{\sigma_{\min}\left(2\sqrt{4-\sigma_{\min}^2}+(4-\sigma_{\max}\sigma_{\min})\right)}{2\sqrt{4-\sigma_{\min}^2}+\left(4-\sigma_{\min}^2\right)}, & \sigma_{\min} < \dfrac{4}{\sigma_{\max}} \\ \dfrac{\frac{8}{\sigma_{\max}}\sqrt{4-\frac{16}{\sigma_{\max}^2}}}{2\sqrt{4-\frac{16}{\sigma_{\max}^2}}+4-\frac{16}{\sigma_{\max}^2}}, & \sigma_{\min} > \dfrac{4}{\sigma_{\max}} \end{array}\right\} = \alpha^*\left(\sigma_{\min}, \sigma_{\max}\right),$$

we conclude the proof. □

Remark 5.5.4. The choice of $\alpha^*(\sigma_{\min}, \sigma_{\max}) = \gamma_{\max}(\theta^*)$ is not accidental. Indeed, one can easily show that

$$\frac{d}{d\theta}\gamma_{\max}(\theta) = \frac{8-2\sigma_{\max}\theta}{\left(4-\theta^2\right)^{\frac{3}{2}}},$$

and as such

$$\max_{\theta\in(0,\theta^*]}\gamma_{\max}(\theta) = \gamma_{\max}\left(\theta^*\right).$$

As the parameter $\alpha^* = \gamma_{\max}(\theta^*)$ corresponds to the decay rate of our entropy according to Theorem 5.5.1, our choice of $\alpha^*(\sigma_{\min}, \sigma_{\max})$ was motivated by maximising the decay rate that is possible with our methodology.

We now posses all the tools which are required to prove part (b) of Theorem 5.2.2.

Proof of part (b) *of Theorem 5.2.2.* The convergence estimation for $E_{\theta^*}\left(u(t)-u_{\text{avg}}, v(t)\right)$ follows immediately from Theorem 5.5.1 and Lemma 5.5.3. To obtain (5.2.10) we use Lemma 5.3.4 in a similar fashion to the way we proved part (a). □

5.6 Convergence to equilibrium in a 3−velocity Goldstein-Taylor model

The Goldstein-Taylor model can be thought of as a simplification of the BGK equation [9, 1]

$$\partial_t f(x,v,t) + v\cdot\nabla_x f(x,v,t) - \nabla_x V(x)\cdot\nabla_v f(x,v,t) = M(v)\int f(x,v,t)dv - f(x,v,t),$$

where the variable v is now in the discrete velocity space $\{v_1, \ldots, v_n\}$, the variable x is in the torus $\mathbb{T}$, and the potential $V(x)$ is zero. The r.h.s. of the above BGK equation corresponds to a projection onto the Maxwellian $M(v)$; in the discrete velocity case this Maxwellian is replaced by a constant matrix that determines the large time behaviour of the new model. Under the natural physical assumption of symmetry in the velocities (i.e. $\sum_{i=1}^n v_i = 0$) and the expectation that the solutions will converge towards a state that is *equally distributed in* v and constant in x [4], we

[4] If one wants to approximate the BGK equation with a Maxwellian relaxation function, then the column vector $(\frac{1}{n}, \ldots, \frac{1}{n})^T$ inside the relaxation matrix would have to be replaced by a *discrete Maxwellian*, as was done [2, §4.2].

find one potential multi-velocity extension of the Goldstein-Taylor model on $\mathbb{T}\times(0,\infty)$:

$$\partial_t \begin{pmatrix} f_1(x,t) \\ \vdots \\ f_n(x,t) \end{pmatrix} + \mathcal{V} \begin{pmatrix} f_1(x,t) \\ \vdots \\ f_n(x,t) \end{pmatrix} = \sigma(x) \left(\begin{pmatrix} \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{pmatrix} \otimes (1,\dots,1) - \mathbf{I} \right) \begin{pmatrix} f_1(x,t) \\ \vdots \\ f_n(x,t) \end{pmatrix}, \tag{5.6.1}$$

with the the diagonal matrix $\mathcal{V} := \mathrm{diag}[v_1,\dots,v_n]$, and the discrete velocities

$$\{v_1,\dots,v_n\} = \begin{cases} \left\{-k+\frac{1}{2},\dots,-\frac{1}{2},\frac{1}{2},\dots,k-\frac{1}{2}\right\}, & n=2k \\ \{-k,\dots,-1,0,1,\dots,k\}, & n=2k-1 \end{cases}, \quad n\in\mathbb{N},\ n\geq 2.$$

The matrix on the r.h.s. of (5.6.1) takes the form

$$\boldsymbol{Q} = \frac{1}{n} \begin{pmatrix} 1-n & 1 & \dots & 1 \\ 1 & 1-n & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \dots & 1-n \end{pmatrix}$$

which has $(1,1,\dots,1)^T$ 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 differences such as $\left\{f_i - f_j\right\}_{i,j=1,\dots,n}$ converge to zero. For more information we refer the interested reader to [1].

In this section we will consider a simple 3−velocity Goldstein-Taylor model, which is governed by the following system of equations on $\mathbb{T}\times(0,\infty)$

$$\begin{aligned} \partial_t f_1(x,t) + \partial_x f_1(x,t) &= \frac{\sigma(x)}{3}\left(f_2(x,t)+f_3(x,t)-2f_1(x,t)\right), \\ \partial_t f_2(x,t) &= \frac{\sigma(x)}{3}\left(f_1(x,t)+f_3(x,t)-2f_2(x,t)\right), \\ \partial_t f_3(x,t) - \partial_x f_3(x,t) &= \frac{\sigma(x)}{3}\left(f_1(x,t)+f_2(x,t)-2f_3(x,t)\right). \end{aligned} \tag{5.6.2}$$

Much like our Goldstein-Taylor equation, (5.1.1), we can recast the above with the variables

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix}, \tag{5.6.3}$$

which yields the following set of equations:

$$\begin{aligned} \partial_t u_1(x,t) + \sqrt{\frac{2}{3}}\partial_x u_2(x,t) &= 0, \\ \partial_t u_2(x,t) + \sqrt{\frac{2}{3}}\partial_x u_1(x,t) + \frac{1}{\sqrt{3}}\partial_x u_3(x,t) &= -\sigma(x) u_2(x,t), \\ \partial_t u_3(x,t) + \frac{1}{\sqrt{3}}\partial_x u_2(x,t) &= -\sigma(x) u_3(x). \end{aligned} \tag{5.6.4}$$

The orthogonal transformation (5.6.3) has a strong geometrical reasoning behind it, as it diagonalises the appropriate "interaction matrix", $\boldsymbol{Q}$. It is also worth to mention that much like (5.2.1), this transformations brings us to the macroscopic variables. Indeed, up to some scaling u_1 is the mass, u_2 is the flux, and u_3 is a linear combination of the kinetic energy and the mass.

Following our intuition we expect that by denoting

$$u_\infty := \frac{1}{2\sqrt{3}\pi}\int_{\mathbb{T}} \left(f_{1,0}(x)+f_{2,0}(x)+f_{3,0}(x)\right)dx,$$

we will find that

$$u_1(t,x) \overset{t\to\infty}{\longrightarrow} u_\infty, \quad u_2(t,x) \overset{t\to\infty}{\longrightarrow} 0, \quad u_3(t,x) \overset{t\to\infty}{\longrightarrow} 0.$$

To prove this result we shall introduce an appropriate Lyapunov functional. To find this functional, we have two options, even for the simple case of constant σ (which is our base case): Proceeding as in §5.4.2, we could use a modal decomposition of (5.6.4) and the (optimal) positive definite matrices $\mathbf{P}_k$ to construct an entropy functional with sharp decay, and then rewrite it in physical space, using pseudo-differential operators. This construction, which is analogous to the construction of $E_\theta(f,g)$ from (5.2.2), can become extremely cumbersome in dimension 3 and higher.

As a simpler alternative we shall hence rather follow the strategy from [1, §4.3] and [2, §2.3]: In Fourier space, the system matrix of (5.6.4) reads as

$$\mathbf{C}_k = \begin{pmatrix} 0 & \sqrt{\frac{2}{3}}ik & 0 \\ \sqrt{\frac{2}{3}}ik & \sigma & \frac{1}{\sqrt{3}}ik \\ 0 & \frac{1}{\sqrt{3}}ik & 0 \end{pmatrix}.$$

We note that, for $k \neq 0$, the *hypocoercivity index*[5] of $\mathbf{C}_k$, as well as of (5.6.4) is one, since this index is always bounded from above by the kernel dimension of the symmetric part of the generator, cf. [2]. For such index-1 problems, Theorem 2.6 from [2] shows that the choice

$$\mathbf{P}_k = \begin{pmatrix} 1 & \frac{\lambda}{ik} & 0 \\ -\frac{\lambda}{ik} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad k \neq 0,$$

with an appropriate $\lambda \in \mathbb{R}$, always yields a (simple) Lyapunov functional for (5.6.4), typically with a sub-optimal decay rate. Much like in §5.5, this guides us to the definition of our functional, expressed in the following theorem:

Theorem 5.6.1. *Let $u_1, u_2, u_3 \in C([0,\infty); L^2(\mathbb{T}))$ be mild real valued solutions to* (5.6.4) *with initial datum $u_{1,0}, u_{2,0}, u_{3,0} \in L^2(\mathbb{T})$. Denoting by*

$$\mathfrak{E}_\theta\left(f,g,h\right) := \|f\|^2 + \|g\|^2 + \|h\|^2 - \frac{\theta}{2\pi}\int_0^{2\pi} \left(\partial_x^{-1} f(x)\, g(x)\right)dx,$$

we have that

$$\mathfrak{E}_\theta\left(u_1(t)-u_\infty, u_2(t), u_3(t)\right) \le \mathfrak{E}_\theta\left(u_{1,0}-u_\infty, u_{2,0}, u_{3,0}\right)e^{-\alpha t}, \quad t \ge 0, \tag{5.6.5}$$

[5] This index characterises the degree of degeneracy of ODE or PDE-evolution equations, cf. [2].

for any $\theta > 0$ *and* $\alpha > 0$ *such that*

$$\sqrt{\frac{2}{3}}\theta + \alpha < 2\sigma_{\min}, \quad \alpha \le \sqrt{\frac{2}{3}}\theta, \tag{5.6.6}$$

and

$$\left(\sup_{x\in\mathbb{T}} \frac{\theta^2 (\sigma(x)-\alpha)^2}{8\sigma(x) - 4\sqrt{\frac{2}{3}}\theta - 4\alpha}\right) + \left(\sup_{x\in\mathbb{T}} \frac{\theta^2}{12(2\sigma(x)-\alpha)}\right) \le \sqrt{\frac{2}{3}}\theta - \alpha. \tag{5.6.7}$$

Remark 5.6.2. For $0 < \theta < 2$, $\mathfrak{E}_\theta(f,g,h)$ is equivalent to $\|f\|^2 + \|g\|^2 + \|h\|^2$. Indeed, following Lemma 5.3.4 we see that

$$\left(1 - \frac{|\theta|}{2}\right)\left(\|f\|^2 + \|g\|^2\right) + \|h\|^2 \le \mathfrak{E}_\theta(f,g,h) \le \left(1 + \frac{|\theta|}{2}\right)\left(\|f\|^2 + \|g\|^2\right) + \|h\|^2.$$

Proof of Theorem 5.6.1. 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 (5.6.4) 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

$$(u_1(t))_{\text{avg}} = \left(u_{1,0}\right)_{\text{avg}} = u_\infty = 0.$$

Next, we compute the time derivatives of the L^2 norms and obtain:

$$\begin{aligned} \frac{d}{dt}\left(\|u_1(t)\|^2 + \|u_2(t)\|^2 + \left\|u_3(t)^2\right\|\right) = -\frac{1}{\pi}\int_0^{2\pi} \sigma(x) u_2(x,t)^2 dx \\ -\frac{1}{\pi}\int_0^{2\pi} \sigma(x) u_3(x,t)^2 dx. \end{aligned} \tag{5.6.8}$$

Continuing, we see that

$$\begin{aligned} \frac{d}{dt}\int_0^{2\pi} \partial_x^{-1} u_1(x,t) u_2(x,t) dx = 2\pi\sqrt{\frac{2}{3}}\left(\left(u_2(t)_{\text{avg}}\right)^2 - \|u_2(t)\|^2\right) + \frac{2\sqrt{2}\pi}{\sqrt{3}}\|u_1(t)\|^2 \\ + \frac{1}{\sqrt{3}}\int_0^{2\pi} u_1(x,t) u_3(x,t) dx - \int_0^{2\pi} \sigma(x) \partial_x^{-1} u_1(x,t) u_2(x,t) dx, \end{aligned}$$

where we used Lemma 5.3.2. As such, together with (5.6.8), we conclude that

$$\begin{aligned} \frac{d}{dt}\mathfrak{E}_\theta(u_1(t), u_2(t), u_3(t)) = -\frac{1}{2\pi}\int_0^{2\pi}\left(2\sigma(x) - \sqrt{\frac{2}{3}}\theta\right) u_2(x,t)^2 dx \\ -\frac{1}{\pi}\int_0^{2\pi} \sigma(x) u_3(x,t)^2 dx - \sqrt{\frac{2}{3}}\theta\|u_1(t)\|^2 - \sqrt{\frac{2}{3}}\theta\left(u_2(t)_{\text{avg}}\right)^2 \\ -\frac{\theta}{2\sqrt{3}\pi}\int_0^{2\pi} u_1(x,t) u_3(x,t) dx + \frac{\theta}{2\pi}\int_0^{2\pi} \sigma(x)\partial_x^{-1} u_1(x,t) u_2(x,t) dx. \end{aligned} \tag{5.6.9}$$

Thus

$$\frac{d}{dt}\mathfrak{E}_\theta\left(u_1(t),u_2(t),u_3(t)\right)=-\alpha\mathfrak{E}_\theta\left(u_1(t),u_2(t),u_3(t)\right)+R_{\theta,\alpha,\sigma}(t)$$

with

$$\begin{aligned}R_{\theta,\alpha,\sigma}(t):=&-\frac{1}{2\pi}\int_0^{2\pi}\left(2\sigma(x)-\sqrt{\frac{2}{3}}\theta-\alpha\right)u_2(x,t)^2dx\\ &-\frac{1}{2\pi}\int_0^{2\pi}\left(2\sigma(x)-\alpha\right)u_3(x,t)^2dx-\left(\sqrt{\frac{2}{3}}\theta-\alpha\right)\|u_1(t)\|^2-\sqrt{\frac{2}{3}}\theta\left(u_2(t)_{\text{avg}}\right)^2\\ &-\frac{\theta}{2\sqrt{3}\pi}\int_0^{2\pi}u_1(x,t)u_3(x,t)dx+\frac{\theta}{2\pi}\int_0^{2\pi}\left(\sigma(x)-\alpha\right)\partial_x^{-1}u_1(x,t)u_2(x,t)dx.\end{aligned}\tag{5.6.10}$$

To conclude the proof it is enough to show that under conditions (5.6.6) and (5.6.7) we have that $R_{\theta,\alpha,\sigma}(t)\le 0$. We will, in fact, show the stronger statement:

$$\begin{aligned}&\left|-\frac{\theta}{2\sqrt{3}\pi}\int_0^{2\pi}u_1(x,t)u_3(x,t)dx+\frac{\theta}{2\pi}\int_0^{2\pi}\left(\sigma(x)-\alpha\right)\partial_x^{-1}u_1(x,t)u_2(x,t)dx\right|\\ &\qquad\le\frac{1}{2\pi}\int_0^{2\pi}\left(2\sigma(x)-\sqrt{\frac{2}{3}}\theta-\alpha\right)u_2(x,t)^2dx\\ &\qquad+\frac{1}{2\pi}\int_0^{2\pi}\left(2\sigma(x)-\alpha\right)u_3(x,t)^2dx+\left(\sqrt{\frac{2}{3}}\theta-\alpha\right)\|u_1(t)\|^2.\end{aligned}\tag{5.6.11}$$

Similarly to the techniques we have used in the proof of part (b) of Theorem 5.2.2, and using the positivity of the coefficients in the last two terms (which follows from (5.6.6)), we see that

$$\begin{aligned}&\left|\frac{\theta}{2\pi}\int_0^{2\pi}\left(\sigma(x)-\alpha\right)\partial_x^{-1}u_1(x,t)u_2(x,t)dx\right|\\ &\le\frac{\theta}{2\pi}\int_0^{2\pi}\frac{|\sigma(x)-\alpha|}{\sqrt{2\sigma(x)-\sqrt{\frac{2}{3}}\theta-\alpha}}\left|\partial_x^{-1}u_1(x,t)\right|\cdot\sqrt{2\sigma(x)-\sqrt{\frac{2}{3}}\theta-\alpha}\,|u_2(x,t)|dx\\ &\le\left(\sup_{x\in\mathbb{T}}\frac{\theta^2\left(\sigma(x)-\alpha\right)^2}{8\sigma(x)-4\sqrt{\frac{2}{3}}\theta-4\alpha}\right)\|u_1(t)\|^2+\frac{1}{2\pi}\int_0^{2\pi}\left(2\sigma(x)-\sqrt{\frac{2}{3}}\theta-\alpha\right)u_2(x,t)^2dx,\end{aligned}$$

and that

$$\begin{aligned}\left|\frac{\theta}{2\sqrt{3}\pi}\int_0^{2\pi}u_1(x,t)u_3(x,t)dx\right|&\le\frac{\theta}{2\pi}\int_0^{2\pi}\frac{|u_1(x,t)|}{\sqrt{6\sigma(x)-3\alpha}}\sqrt{2\sigma(x)-\alpha}\,|u_3(x,t)|dx\\ &\le\left(\sup_{x\in\mathbb{T}}\frac{\theta^2}{12\left(2\sigma(x)-\alpha\right)}\right)\|u_1(t)\|^2+\frac{1}{2\pi}\int_0^{2\pi}\left(2\sigma(x)-\alpha\right)u_3(x,t)^2dx.\end{aligned}$$

Thus, one sees that (5.6.11) holds when

$$\left(\sup_{x\in\mathbb{T}}\frac{\theta^2\left(\sigma(x)-\alpha\right)^2}{8\sigma(x)-4\sqrt{\frac{2}{3}}\theta-4\alpha}\right)+\left(\sup_{x\in\mathbb{T}}\frac{\theta^2}{12\left(2\sigma(x)-\alpha\right)}\right)\le\sqrt{\frac{2}{3}}\theta-\alpha,$$

which is (5.6.7). The proof is complete. □

While we have elected not to optimise the choice of α (as in §5.5), we can still infer the following, simpler yet far from optimal, corollary:

Corollary 5.6.3. Let $\theta > 0$ and $\alpha > 0$ be such that

$$\sqrt{\frac{2}{3}}\theta + \alpha < 2\sigma_{\min}, \quad \alpha \le \sqrt{\frac{2}{3}}\theta$$

and

$$\frac{\theta^2\sigma_{\max}^2}{8\sigma_{\min} - 4\sqrt{\frac{2}{3}}\theta - 4\alpha} + \frac{\theta^2}{12\,(2\sigma_{\min} - \alpha)} \le \sqrt{\frac{2}{3}}\theta - \alpha. \tag{5.6.12}$$

then

$$\mathfrak{E}_\theta\,(u_1(t) - u_\infty, u_2(t), u_3(t)) \le \mathfrak{E}_\theta\left(u_{1,0} - u_\infty, u_{2,0}, u_{3,0}\right)e^{-\alpha t}.$$

In particular, for

$$\alpha := \min\left(\frac{\sigma_{\min}}{2}, \frac{3\sigma_{\min}}{9\sigma_{\max}^2 + 1}\right)$$

we have that $\mathfrak{E}_{\sqrt{6}\alpha}$ decays exponentially to zero with rate α.

Proof. Since $\alpha < 2\sigma_{\min} \le \sigma_{\max} + \sigma_{\min}$ we see that

$$\alpha - \sigma_{\max} < \sigma_{\min} \le \sigma(x) < \sigma_{\max} + \alpha,$$

implying that $(\sigma(x) - \alpha)^2 \le \sigma_{\max}^2$ for any $x \in \mathbb{T}$. Using this with additional elementary estimation on the denominator of the expressions that appear in (5.6.7), we see that (5.6.6) and (5.6.7) are valid. As such the first statement of the corollary follows from Theorem 5.6.1.
To show the second part of the corollary we notice that with the choice $\theta_\alpha := \sqrt{6}\alpha$ and $\alpha \le \frac{\sigma_{\min}}{2}$

$$\sqrt{\frac{2}{3}}\theta_\alpha + \alpha = 3\alpha < 2\sigma_{\min}, \quad \alpha \le 2\alpha = \sqrt{\frac{2}{3}}\theta_\alpha,$$

giving us (5.6.6). Using the inequalities

$$8\sigma_{\min} - 4\sqrt{\frac{2}{3}}\theta_\alpha - 4\alpha \ge 2\sigma_{\min}, \quad \text{and} \quad 2\sigma_{\min} - \alpha \ge \frac{3}{2}\sigma_{\min}$$

for the l.h.s. of (5.6.12), we see that

$$\frac{\theta_\alpha^2\sigma_{\max}^2}{8\sigma_{\min} - 4\sqrt{\frac{2}{3}}\theta_\alpha - 4\alpha} + \frac{\theta_\alpha^2}{12\,(2\sigma_{\min} - \alpha)} \le \left(9\sigma_{\max}^2 + 1\right)\frac{\alpha^2}{3\sigma_{\min}}.$$

Thus, since $\sqrt{\frac{2}{3}}\theta_\alpha - \alpha = \alpha$, the desired condition (5.6.12) is valid when

$$\alpha \le \frac{3\sigma_{\min}}{9\sigma_{\max}^2 + 1},$$

which concludes the proof. □

5.7 Appendix: Lack of optimality

In this appendix we will briefly discuss the lack of optimality of our decay rate for non-homogeneous $\sigma(x)$ in comparison to that given in [8]. We will even go one step further and show how one can improve our general methodology in simple cases, though even this improvement will fall short of the optimal convergence rate.

As one simple example we will explore the following relaxation function:

$$\sigma(x) := \begin{cases} 1, & 0 < x \leq \pi \\ 4, & \pi < x \leq 2\pi \end{cases}, \tag{5.7.1}$$

which is motivated by the fact that for this function $\sigma_{\min} = \frac{\sigma_{\max}}{4}$, and so the choice of $\theta^* = 1$ in our main Theorem 5.2.2 (b) comes "from both directions".

Before we start with a more structured discussion, we would like to explain how one can improve the technique we developed in §5.5. A crucial point in the investigation of the behaviour of E_{θ^*} was to find, and close, a linear differential inequality for this entropy, as can be seen in the proof of Theorem 5.5.1. One of the final steps in this proof, appearing in (5.5.6), was to combine Poincaré inequality with an L^∞ estimation on the mixed term of $\partial_x^{-1}\left(u - u_{\text{avg}}\right)$ and v, to show the non-positivity of an appropriate "remainder". The use of these two inequalities is somewhat crude (yet due to that, quite general), and one can imagine that replacing these two estimations with an inequality that is more L^2 based would improve the range of validity of the theorem. One idea that comes to mind is a *weighted Poincaré inequality*, i.e. an inequality of the form

$$\int_0^{2\pi} \left(f(x) - f_{\text{avg}}\right)^2 \omega(x)dx \leq C^2 \int_0^{2\pi} \left(f'(x)\right)^2 dx. \tag{5.7.2}$$

for a given weight $\omega(x) \geq 0$ and constant C. Denoting by[6]

$$C_\omega := \inf\left\{C > 0 \,\Big|\, \int_0^{2\pi} \left(f(x) - f_{\text{avg}}\right)^2 \omega(x)dx \leq C^2 \int_0^{2\pi} \left(f'(x)\right)^2 dx, \quad \forall f \in H^1(\mathbb{T})\right\},$$

we can replace condition (5.5.2) of Theorem 5.5.1 with the improved condition

$$\frac{\theta^2}{4} C_\omega^2 \leq \theta - \alpha, \quad \text{where} \quad \omega(x) = \frac{(\sigma(x) - \alpha)^2}{2\sigma(x) - \theta - \alpha}. \tag{5.7.3}$$

(5.7.3) will be explicitly derived in §5.7.2.

From this point onwards the appendix will proceed as follows: First we will show how one can find the optimal weighted Poincaré constant, and compute it in some simple cases, which we will then use in the case were $\sigma(x)$ is given by (5.7.1) to obtain an improvement of our current rate of convergence to equilibrium. Next we will compute the optimal rate given by [8], and conclude a lack of optimality by comparing the rate we achieved in our main theorem, the improved rate we have found, and the optimal rate of [8].

[6] Note that by definition, and by Lemma 5.3.1, $C_\omega \leq \sqrt{\|\omega\|_\infty}$, and as such we will automatically get an improvement to condition (5.5.2).

5.7.1 Weighted Poincaré inequality

The problem of finding a weighted Poincaré inequality and its associated sharp constant can be recast as a constrained variational problem. We define the functional

$$\mathscr{F}: \quad \mathscr{D} := H^1(\mathbb{T}) \to \mathbb{R},$$

where $H^1(\mathbb{T})$ is the Sobolev space of real valued periodic functions, by

$$\mathscr{F}(u) := \int_0^{2\pi} \left(u'(x)\right)^2 dx,$$

and denote by

$$c_{\min} := \inf\left\{\mathscr{F}(u) \,\middle|\, u \in \mathscr{D},\ \int_0^{2\pi} u(x)^2 \omega(x)\, dx = 1,\ \int_0^{2\pi} u(x)\, dx = 0\right\}. \tag{5.7.4}$$

Even though the minimization set is not convex, standard techniques from Calculus of Variation (see for instance [14, §8] and [22]) show that if ω is bounded then the infimum is attained (the conditions on ω can be weakened).
One can easily check that in that case

$$C_\omega^2 = \frac{1}{c_{\min}}.$$

Finding a minimiser to the problem (5.7.4) amounts to solving the following constrained Euler-Lagrange equation on $\mathbb{T}$

$$u''(x) + \lambda u(x)\omega(x) - \tau = 0, \tag{5.7.5}$$

considered in weak form, with two Lagrange multipliers $\lambda > 0$ and $\tau \in \mathbb{R}$. Integrating (5.7.5) against u shows that

$$\lambda = \mathscr{F}(u), \tag{5.7.6}$$

which we will use shortly.

Since $\omega \in L^\infty(\mathbb{T})$, we find that $u \in H^2(\mathbb{T}) \hookrightarrow C^1(\mathbb{T})$. When ω is piecewise constant, the ODE (5.7.5), now in strong form, can be solved explicitly. This shows that in these cases the minimiser of $\mathscr{F}$ is actually unique.

As we shall see in §5.7.2 below, the relevant weight functions we require for our improved study are closely related to $\sigma(x)$. With (5.7.1) in mind, we shall consider weights of the form:

$$\omega(x) := \begin{cases} \omega_1, & 0 < x \le \pi \\ \omega_2, & \pi < x \le 2\pi \end{cases}.$$

Hence, the solution to the Euler-Lagrange equation (5.7.5) is given by

$$\begin{aligned} u(x) &= \begin{cases} c_1 \sin\left(\sqrt{\lambda\omega_1}\, x\right) + c_2 \cos\left(\sqrt{\lambda\omega_1}\, x\right) + \frac{\tau}{\lambda\omega_1}, & 0 < x < \pi \\ c_3 \sin\left(\sqrt{\lambda\omega_2}\, x\right) + c_4 \cos\left(\sqrt{\lambda\omega_2}\, x\right) + \frac{\tau}{\lambda\omega_2}, & \pi < x < 2\pi \end{cases} \\ &=: \begin{cases} u_1(x), & 0 < x < \pi \\ u_2(x), & \pi < x < 2\pi \end{cases}, \end{aligned} \tag{5.7.7}$$

and it satisfies the following C^1-matching conditions and constraints:

$$\begin{aligned} u_1(0) &= u_2(2\pi), \\ u_1(\pi) &= u_2(\pi), \\ u_1'(0) &= u_2'(2\pi), \\ u_1'(\pi) &= u_2'(\pi), \\ \int_0^\pi u_1(x)\,dx + \int_\pi^{2\pi} u_2(x)\,dx &= 0, \\ \int_0^\pi \omega_1 u_1(x)^2\,dx + \int_\pi^{2\pi} \omega_2 u_2^2(x)\,dx &= 1. \end{aligned} \tag{5.7.8}$$

The first five equations correspond to the linear set of equations:

$$\boldsymbol{M}(\lambda)\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ \tau \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

where the matrix $\boldsymbol{M}(\lambda)$ is

$$\begin{pmatrix} 0 & 1 & -\sin\left(2\pi\sqrt{\lambda\omega_2}\right) & -\cos\left(2\pi\sqrt{\lambda\omega_2}\right) & \frac{\omega_2-\omega_1}{\lambda\omega_1\omega_2} \\ \sin\left(\pi\sqrt{\lambda\omega_1}\right) & \cos\left(\pi\sqrt{\lambda\omega_1}\right) & -\sin\left(\pi\sqrt{\lambda\omega_2}\right) & -\cos\left(\pi\sqrt{\lambda\omega_2}\right) & \frac{\omega_2-\omega_1}{\lambda\omega_1\omega_2} \\ \sqrt{\omega_1} & 0 & -\sqrt{\omega_2}\cos\left(2\pi\sqrt{\lambda\omega_2}\right) & \sqrt{\omega_2}\sin\left(2\pi\sqrt{\lambda\omega_2}\right) & 0 \\ \sqrt{\omega_1}\cos\left(\pi\sqrt{\lambda\omega_1}\right) & -\sqrt{\omega_1}\sin\left(\pi\sqrt{\lambda\omega_1}\right) & -\sqrt{\omega_2}\cos\left(\pi\sqrt{\lambda\omega_2}\right) & \sqrt{\omega_2}\sin\left(\pi\sqrt{\lambda\omega_2}\right) & 0 \\ \frac{1-\cos(\pi\sqrt{\lambda\omega_1})}{\sqrt{\omega_1}} & \frac{\sin(\pi\sqrt{\lambda\omega_1})}{\sqrt{\omega_1}} & \frac{\cos(\pi\sqrt{\lambda\omega_2})-\cos(2\pi\sqrt{\lambda\omega_2})}{\sqrt{\omega_2}} & \frac{\sin(2\pi\sqrt{\lambda\omega_2})-\sin(\pi\sqrt{\lambda\omega_2})}{\sqrt{\omega_2}} & \frac{\pi(\omega_2+\omega_1)}{\sqrt{\lambda}\omega_1\omega_2} \end{pmatrix}.$$

As we are looking for a non-zero solution to the above equation, we must have that $\det(\boldsymbol{M}(\lambda)) = 0$. In (5.7.8), the last condition on u_1 and u_2 merely acts as normalisation, and doesn't help in finding λ. Hence, due to (5.7.6), we find that

$$c_{\min}(\omega_1, \omega_2) = \min\{\lambda > 0 \mid \det(\boldsymbol{M}(\lambda)) = 0\}.$$

This is how we can find $c_{\min}$, and consequently C_ω, explicitly (numerically in many cases).

5.7.2 Improved methodology

We return now to the proof of the differential inequality (5.5.4) that governs the evolution of E_θ, which is essentially based on the estimate (5.5.6). Choosing $\theta^* = 1$ and $\sigma(x)$ as in (5.7.1) and using the weight

$$\omega_\alpha(x) := \frac{(\sigma(x)-\alpha)^2}{2\sigma(x)-1-\alpha} = \begin{cases} 1-\alpha & 0 < x \le \pi \\ \frac{(4-\alpha)^2}{7-\alpha} & \pi < x \le 2\pi \end{cases},$$

which appears in the penultimate line of (5.5.6), we see that by using the previously discussed weighted Poincaré inequality instead of the last step of (5.5.6), we obtain from (5.5.4) and (5.5.6):

$$\begin{aligned} \frac{d}{dt}E_1\left(u(t)-u_{\text{avg}}, v(t)\right) \le{}& -\alpha E_1\left(u(t)-u_{\text{avg}}, v(t)\right) \\ & -\left(1-\alpha-\frac{C_{\omega_\alpha}^2}{4}\right)\left\|u(t)-u_{\text{avg}}\right\|^2. \end{aligned} \tag{5.7.9}$$

We will maximise the decay rate α, satisfying

$$0 < \alpha \leq 1 - \frac{C^2_{\omega_\alpha}}{4} < 1 \tag{5.7.10}$$

(so that the second term in (5.7.9) is non-positive) by a processes of iteration: Guessing the starting value $\alpha_0 := \alpha^*(1,4) = 2\left(2-\sqrt{3}\right)$ (the rate one obtains from our main Theorem 5.2.2, cf. (5.2.9)) we follow the process described in the previous subsection and find the weighted Poincaré constant $C^2_{\omega_{\alpha_0}} = 1.12013...$, which indeed satisfies (5.7.10).
We proceed and create a sequence $\{\alpha_n\}_{n\in\mathbb{N}_0}$, defined recursively, so that each α_n improves upon the previous step by taking its "optimal" value, i.e.

$$\alpha_n := 1 - \frac{C^2_{\omega_{\alpha_{n-1}}}}{4}, \qquad n \in \mathbb{N},$$

as long as (5.7.10) is still satisfied for this choice. A change of α implies a change of our weight function $\omega_\alpha(x)$, yet these new weights are still of the form given in our previous subsection. As such we are able to compute the appropriate $C_{\omega_{\alpha_n}}$'s, and to show that this sequence converges to the improved decay rate[7]

$$\alpha_{\max} \approx 0.7234.$$

5.7.3 Comparison of convergence rates

The optimal rate[8] of exponential convergence to the Goldstein-Taylor equation, (5.1.1), was found by Bernard and Salvarani in [8]. Taking into account the different scaling of the torus $\mathbb{T}$ in our paper, this convergence rate is given by

$$\alpha_{BS} := \frac{1}{\pi} \min\left(\|\tilde{\sigma}\|_{L^1\left(\frac{\mathbb{T}}{2\pi}\right)}, \tilde{D}(0) \right)$$

where

$$\tilde{\sigma}(\xi) := \pi\sigma(2\pi\xi), \qquad \xi \in \frac{\mathbb{T}}{2\pi},$$

and $\tilde{D}(0)$ is the spectral gap of the telegrapher's equation, see [8, Proposition 3.5], [19, Theorem 2]. More precisely,

$$\tilde{D}(0) := \inf\left\{ \operatorname{Re}\lambda_j \,|\, \lambda_j \in \left(\text{spectrum of } \boldsymbol{A}_{\tilde{\sigma}} = \begin{pmatrix} 0 & -1 \\ -\partial_{xx} & 2\tilde{\sigma} \end{pmatrix} \text{ in } H^2 \oplus H^1 \right) \setminus \{0\} \right\}.$$

We note that, for $\tilde{\sigma}$ constant and in Fourier space, the matrix $\begin{pmatrix} 0 & -1 \\ k^2 & 2\tilde{\sigma} \end{pmatrix}$ is related to $\mathbf{C}_k$ from (5.4.2) by a simple similarity transformation.

Following on our choice for $\sigma(x)$ from (5.7.1), we see that

$$\tilde{\sigma}(\xi) = \begin{cases} \sigma_1 := \pi, & 0 < \xi \leq \frac{1}{2} \\ \sigma_2 := 4\pi, & \frac{1}{2} < \xi \leq 1 \end{cases}, \tag{5.7.11}$$

[7] This process was dealt with numerically.
[8] at least for H^1-initial data

and as such $\|\widetilde{\sigma}\|_{L^1(\frac{\mathbb{T}}{2\pi})} = \frac{5\pi}{2}$.
The calculation of $\widetilde{D}(0)$ is more involved. According to [19], the spectrum of $\boldsymbol{A}_{\widetilde{\sigma}}$, besides potentially $\{0\}$, is discrete and the real part of its eigenvalues must lie in $(0, 2\|\widetilde{\sigma}\|_\infty]$. A more detailed investigation of the spectrum can be found in [11].
The eigenvalue problem

$$\boldsymbol{A}_{\widetilde{\sigma}} \begin{pmatrix} u \\ v \end{pmatrix} = \gamma \begin{pmatrix} u \\ v \end{pmatrix},$$

with $\gamma \neq 0$, is equivalent to the set of equations

$$v''(\xi) = \gamma\left(\gamma - 2\widetilde{\sigma}(\xi)\right) v(\xi), \quad v(\xi) = -\gamma u(\xi).$$

To find $\tilde{D}(0)$ it is sufficient to consider only eigenvalues with $\operatorname{Re}\gamma \in (0, 2\sigma_1) = (0, 2\pi)$, since this complex strip already includes one (real) eigenvalue as we shall see below. With the notation $\tau_{1,2}(\gamma) := \sqrt{\gamma(2\sigma_{1,2} - \gamma)}$, which may have to be considered as a complex root, the solution of the ODE is of the form

$$v(\xi) = \begin{cases} A_1 \cos(\tau_1(\gamma)\,\xi) + B_1 \sin(\tau_1(\gamma)\,\xi), & 0 < \xi \leq \frac{1}{2} \\ A_2 \cos(\tau_2(\gamma)\,\xi) + B_2 \sin(\tau_2(\gamma)\,\xi), & \frac{1}{2} < \xi \leq 1 \end{cases}.$$

With C^1–matching conditions at $\xi = 0$ and $\xi = \frac{1}{2}$, the coefficients satisfy the following system of linear equations:

$$\boldsymbol{M}(\gamma) \begin{pmatrix} A_1 \\ B_1 \\ A_2 \\ B_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

where the matrix $\boldsymbol{M}(\gamma)$ is given by

$$\begin{pmatrix} 1 & 0 & -\cos\left(\tau_2(\gamma)\right) & -\sin\left(\tau_2(\gamma)\right) \\ 0 & 1 & \frac{\tau_2(\gamma)}{\tau_1(\gamma)} \sin\left(\tau_2(\gamma)\right) & -\frac{\tau_2(\gamma)}{\tau_1(\gamma)} \cos\left(\tau_2(\gamma)\right) \\ \cos\left(\frac{\tau_1(\gamma)}{2}\right) & \sin\left(\frac{\tau_1(\gamma)}{2}\right) & -\cos\left(\frac{\tau_2(\gamma)}{2}\right) & -\sin\left(\frac{\tau_2(\gamma)}{2}\right) \\ \sin\left(\frac{\tau_1(\gamma)}{2}\right) & -\cos\left(\frac{\tau_1(\gamma)}{2}\right) & -\frac{\tau_2(\gamma)}{\tau_1(\gamma)} \sin\left(\frac{\tau_2(\gamma)}{2}\right) & \frac{\tau_2(\gamma)}{\tau_1(\gamma)} \cos\left(\frac{\tau_2(\gamma)}{2}\right) \end{pmatrix}.$$

The requirement that

$$\begin{aligned} \det\left(\boldsymbol{M}(\gamma)\right) &= -\sin\left(\frac{\tau_1(\gamma)}{2}\right) \sin\left(\frac{\tau_2(\gamma)}{2}\right) \left(1 + \left(\frac{\tau_2(\gamma)}{\tau_1(\gamma)}\right)^2\right) \\ &+ 2\frac{\tau_2(\gamma)}{\tau_1(\gamma)} \left(\cos\left(\frac{\tau_1(\gamma)}{2}\right) \cos\left(\frac{\tau_2(\gamma)}{2}\right) - 1\right) = 0 \end{aligned} \tag{5.7.12}$$

yields the wanted eigenvalues $\gamma \in \mathbb{C}$ with $\operatorname{Re}\gamma \in (0, 2\pi)$.
In our case, i.e. when $\tilde{\sigma}(x)$ is given by (5.7.11), we find (numerically) that the minimal real part of the non-zero eigenvalues found from (5.7.12) is approximately 2.72831, which implies that $\tilde{D}(0) \approx 2.72831$. Thus, the optimal decay rate given by [8] is

$$\alpha_{\mathrm{BS}} \approx \frac{1}{\pi} \min\left(\frac{5\pi}{2}, 2.72831\right) \approx 0.86845.$$

Summarising, we now have three convergence rates for the case

$$\sigma(x) = \begin{cases} 1, & 0 \le x \le \pi \\ 4, & \pi < x \le 2\pi \end{cases} \quad :$$

- Rate from our main Theorem 5.2.2: $\alpha^* = 4 - \sqrt{12} \approx 0.5359$.
- Rate from our improved technique in §5.7.2: $\alpha_{\text{max}} \approx 0.7234$.
- Rate from the work of Bertrand and Salvarani: $\alpha_{\text{BS}} \approx 0.86845$.

This shows, as expected, the lack of optimality in our technique.

Bibliography

[1] Achleitner, F., Arnold, A., 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., 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 behaviour 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., Jin, S., 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).

[8] 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).

[9] Bhatnagar, P.L., Gross, E.P., 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).

[10] Cercignani, C., Illner, R., Shinbrot, W.: *A boundary value problem for discrete-velocity models.* Duke Math. J. vol. 55 (4), 889–900 (1987).

[11] Cox, S., Zuazua E.: *The rate at which energy decays in a damped string.* Comm. Partial Differential Equations vol. 19 , no. 1-2, 213–243 (1994).

[12] Desvillettes, L., Salvarani, F.: *Asymptotic behaviour 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] Evans, L.C.: Partial Differential Equations, American Mathematical Soc., (2010).

[15] Goldstein, C.: *On diffusion by discontinuous movements, and on the telegraph equation.* Quart. J. Mech. Appl. Math. vol. 4, 129–156 (1951).

[16] Gosse, L., Toscani, G.: *An asymptotic-preserving well-balanced scheme for the hyperbolic heat equations.* Comptes Rendus Math. vol. 334 (4), 337–342 (2002).

[17] Jin, S.: *Efficient Asymptotic-Preserving (AP) Schemes For Some Multiscale Kinetic Equations.* SIAM J. Sc. Comp. vol. 21 (2), 441–454 (1999).

[18] Kawashima, S.: *Existence and Stability of Stationary Solutions to the Discrete Boltzmann Equation.* Japan J. Indust. Appl. Math. vol. 8, 389–429 (1991).

[19] Lebeau, G.: *Équations des ondes amorties*, Séminaire Équations aux dérivées partielles (École Polytechnique), talk no. 15, 1–14 (1993-1994).

[20] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, 2nd edition, (1992).

[21] 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).

[22] M. Struwe, *Variational Methods*, Springer-Verlag Berlin Heidelberg, (2008).

[23] Taylor G.I.: *Diffusion by Continuous Movements*, Proc. London Math. Soc., vol. S2-20 (1), 196–212 (1922).

[24] Tran, M.-B.: *Convergence to equilibrium of some kinetic models.* J. Diff. Eq. vol. 255, 405–440 (2013).

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