Technische Universität Wien
Vienna University of Technology

DISSERTATION

Entropy Method for Hypocoercive Fokker-Planck Type Equations

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

Univ.Prof. Dipl.-Ing. Dr. techn. Anton Arnold
E101 - Institut für Analysis und Scientific Computing

eingereicht an der Technischen Universität Wien
bei der Fakultät für Mathematik und Geoinformation

von

Dipl.-Math. Jan Erb

Matrikelnummer: 0728502
Kaschlgasse 2/7
1200 Wien

Datum

Unterschrift

Acknowledgements.

There are a number of people I want to thank for their support in the years I worked on this thesis.
First, I want to thank my supervisor, Prof.Dr. Arnold, without whom this work would not have been possible. I am also grateful to all of my colleagues, in particular Maia, Dominik, Franz, Lukas, Ilona and Jens. It was great working with you. I acknowledge the financial support of the Wissenschaftskolleg Differential Equations (supported by the Austrian Science Fund) and the TU Vienna.
Many thanks to my Pen&Paper-groups: Christoph, Volker, Andrew, Ole, Martin, Flanf, Chris and Markus, for their support and for providing a welcome distraction. I particularly want to thank Martin for also proofreading this thesis.
Last but not least I want to thank my wife Sigrid for believing in me all these years, and my daughter Merle for cheering me up whenever I got home after a long day.

Deutsche Kurzfassung

Evolutionsgleichungen spielen eine zentrale Rolle in vielen Anwendungsgebieten. Ein Beispiel für eine solche Evolutionsgleichung ist die Fokker-Planck-Gleichung, eine partielle Differentialgleichung für die zeitliche Entwicklung einer Wahrscheinlichkeitsdichte, das heißt einer Funktion, die zum Beispiel die Verteilung von Teilchen im Raum beschreibt. Die Fokker-Planck-Gleichung findet in vielen Naturwissenschaften Anwendung, unter anderem in der Festkörperphysik, Quantenoptik, chemischen Physik und theoretischen Biologie [57]. Sie wird auch in der Finanzmathematik verwendet, dort allerdings in stochastischer Formulierung unter dem Namen Ornstein-Uhlenbeck-Prozess, um Zinsraten, Währungswechselraten und die preisliche Entwicklung von Gütern zu modellieren. Neben den Standardfragestellungen bezüglich Existenz und Regularität von Lösungen sind im Umfeld von Evolutionsgleichungen typischerweise die Existenz und Eindeutigkeit von Stationärzuständen sowie das Langzeitverhalten der Lösungen (zum Beispiel Abklingen zum Stationärzustand und die zugehörige Abklingrate) interessant. In dieser Dissertation behandeln wir diese Fragestellungen für Fokker-Planck-Gleichungen vom Typ

$$
\begin{aligned}
\partial_t f = Lf &:= \operatorname{div}(D\nabla f + Ff) \text{ on } (0,\infty)\times\mathbb{R}^d,\\
f(t=0) &= f_0 \in L^1(\mathbb{R}^d),\\
\int_{\mathbb{R}^d} f_0 \,\mathrm{d}x &= 1, f_0 \geq 0.
\end{aligned}
$$

Hierbei ist $F:\mathbb{R}^d \to \mathbb{R}^d$ ein Vektorfeld und $D = D^T \geq 0$ eine positiv (semi-)definite Matrix in $\mathbb{R}^{d\times d}$. Ist D singulär, wie zum Beispiel in der kinetischen Fokker-Planck-Gleichung [57], gibt es eine mathematisches Problem, welches in den letzten Jahren zunehmend in den Fokus der Forschung gerückt ist: Im Term $\operatorname{div}(D\nabla f)$ zweiter Ordnung, der eine wesentliche Rolle für das Verhalten von Lösungen spielt, fehlen Ableitungen in einem Teil der Ortsvariablen. Daher ist die Gleichung nicht mehr voll parabolisch, sondern degeneriert parabolisch, und viele der bekannten Resultate über voll parabolische Gleichungen lassen sich nicht anwenden (oder zumindest nicht direkt). Es gibt zwei essentielle Eigenschaften für voll parabolische Gleichungen: Regularisierung, das heißt glatte Lösungen selbst bei unstetigen Anfangsdaten, und (unter Bedingungen an den Term erster Ordnung oder den Definitionsbereich) Koerzivität. Koerzivität impliziert ein Abklingen der Lösungen gegen einen Stationärzustand. Beide Eigenschaften sind im Allgemeinen für degeneriert parabolische Gleichungen aufgrund fehlender zweiter Ableitungen nicht zu erwarten. Voraussetzungen, unter denen Regularisierung auch für degeneriert parabolische Gleichungen gilt, wurden in den 60er- und 70er-Jahren primär von Hörmander [42] etabliert. Diese Ergebnisse, die unter das Konzept *Hypoelliptizität* fallen, werden wir in der vorliegenden Arbeit verwenden. Die Existenz von Stationärzuständen und das Abklingen der Lösungen im degenerierten Fall haben in den letzten Jahren wachsendes Interesse unter dem Begriff *Hypokoerzivität* geweckt, vor allem durch die Arbeit von Villani [67]. In der vorliegenden Dissertation werden wir Bedingungen beweisen, unter denen die Fokker-Planck-Gleichung hypokoerziv ist, und zudem einen eindeutigen (normalisierten) Stationärzustand sowie scharfe Abklingraten für Lösungen gegen diesen Stationärzustand berechnen.

Zu diesem Zweck werden wir eine Entropiemethode für lineare Fokker-Planck-Gleichungen mit linearen Driftkoeffizienten F entwickeln, die es erlaubt, scharfe Abklingraten für L^1-Anfangsdaten mit endlicher relativer Entropie (eine schwächere Bedingung als L^2-Anfangsdaten) zu berechnen. Dazu benötigen wir die Regularität und Positivität von Lösungen, die wir in Abschnitt 1.1 beweisen. Diese Eigenschaften folgen aus der Hypoelliptizität der Gleichung, die wir zusammen mit der Hypokoerzivität in Lemma 1.3 charakterisieren. Die erhaltenen Bedingungen verwenden wir in Abschnitt 1.2 um den eindeutigen (normalisierten) Stationärzustand f_∞ zu berechnen und den Operator L auf $L^2(\mathbb{R}^d, f_\infty^{-1})$, dem Standardraum für Fokker-Planck-Gleichungen, zu betrachten. Abschnitt 1.3 enthält das Hauptresultat der vorliegenden Arbeit, Satz 1.27: Mithilfe einer adaptierten Entropiemethode aus [6] berechnen wir eine scharfe Abklingrate für Lösungen in relativer Entropie. Die Schärfe der Rate wird in Satz 1.28 bewiesen. Außerdem diskutieren wir kurz das Verhalten der Fishermatrix, das bisher in diesem Kontext nicht untersucht wurde. In Abschnitt 1.4 berechnen wir das Spektrum von L auf dem gewichteten Raum $L^2(\mathbb{R}^d, f_\infty^{-1})$. Anhand dreier charakteristischer Beispiele diskutieren wir die Ergebnisse dann in Abschnitt 1.5. Abschnitt 1.6 enthält einen alternativen Beweis für die Hypokoerziviät von L beruhend auf einer Methode aus [67]. Abschließend diskutieren wir einige Ergebnisse und Probleme, die auftreten, wenn man nichtlineare Driftkoeffizienten F zulässt.

Contents

Introduction

Evolution equations play a central role in many applications. One such partial differential equation (PDE) is the Fokker-Planck equation, which describes the time evolution of a probability density (a function describing e.g. the likelyhood of finding particles under investigation in a certain region). The Fokker-Planck equation is employed in different fields of natural science, including solid-state physics, quantum optics, chemical physics and theoretical biology [57]. It also appears in financial mathematics as an Ornstein-Uhlenbeck process, used to model interest rates, currency exchange rates and commodity prices; though in this context it is usually found in a stochastic formulation. Besides the usual questions of existence and smoothness of solutions, typical questions in the context of evolution equations are existence and uniqueness of stationary states and long-term behaviour of solutions, such as decay towards the stationary state and the speed of such decay. In this thesis, we shall investigate these questions for Fokker-Planck equations of the form

$$\begin{aligned} \partial_t f = Lf &:= \operatorname{div}(D\nabla f + Ff) \text{ on } (0,\infty)\times\mathbb{R}^d, \qquad (1)\\ f(t=0) &= f_0 \in L^1(\mathbb{R}^d),\\ \int\limits_{\mathbb{R}^d} f_0 \,\mathrm{d}x &= 1, f_0 \geq 0. \end{aligned}$$

Here, $F : \mathbb{R}^d \to \mathbb{R}^d$ is a vector field and $D = D^T \geq 0$ is a positive (semi-)definite matrix in $\mathbb{R}^{d\times d}$. If D is not regular (such as in the case of the kinetic Fokker-Planck equation [57]), there is a mathematical difficulty that has received growing attention in recent years: The second order term $\operatorname{div}(D\nabla f)$, which plays a crucial role in the behaviour of solutions, does not contain derivatives in all space variables. Thus, the equation is not fully parabolic, but degenerate parabolic, and many of the results for fully parabolic equations do not apply (at least not directly). There are two fundamental properties one can expect from fully parabolic equations: regularisation, that is smooth solutions even for non-smooth initial data, and (under conditions on the domain or first-order terms) coercivity, which implies decay of solutions towards a stationary state. Both of these do not in general apply to degenerate parabolic equations due to the missing second-order terms. Conditions for the first property, i.e. regularisation, to apply to degenerate parabolic equations were established in the 60's and 70's, primarily by Hörmander. These run under the concept of *hypoellipticity*, and we will make use of these results in our calculations. Existence of stationary states and decay of solutions in the degenerate case have received growing attention in recent years, and the emerging umbrella term is *hypocoercivity*, introduced primarily by the works of Villani [67]. In this thesis, we shall characterise when (1) exhibits hypocoercivity, and additionally establish stationary states and sharp rates of decay for solutions towards the stationary state.

To do so, we extend the entropy method, a powerful tool in the large-time-analysis of fully parabolic equations, to the case of degenerate Fokker-Planck type equations. The central idea of the entropy method is using the physical relative entropy between a solution and the stationary state as a Lyapunov functional, that is, a measurement of distance that is monotonously decreasing in time, and thus can be used to gain decay estimates. Another often used candidate for Lyapunov functionals is the physical energy of the system, but since the energy contains derivatives of the solutions, using it as a Lyapunov functional mandates higher regularity for the initial conditions. Employing the entropy method, one can reduce the regularity requirement for initial data to finite relative entropy, a requirement "between" L^1 and L^2. Another approach to decay estimates is the spectral method, where one computes a lower bound on the real part of the spectrum of L (on the orthogonal complement of its kernel, of course). The spectral approach requires a fixed setting, usually $H^1 \subset L^2$ to make use of their Hilbert space properties, and decay estimates thus obtained do not directly transfer to initial data with less regularity. Furthermore, the spectral method is usually only applicable for linear PDEs, whereas Lyapunov functionals (and the entropy method) have been successfully employed in many non-linear models (e.g. [4], [16], [17], [24], [25], [23], [28], [55]).

We remark that, in (1), the matrix D only appears in front of the gradient, whereas in most of the literature, the operator is usually written in the form

$$\operatorname{div}(D(\nabla f + F f)).$$

This is equivalent for a regular D, but if D is singular, the above form will not allow an entropy method: The operator then only acts on some subset of $\mathbb{R}^d$ – to be precise, the complement of the kernel of D – and there will never be a unique steady state. While this difference in notation is necessary, it is important to keep in mind when comparing this thesis with literature on non-degenerate equations.

For most of this thesis, we make the additional assumption

$$F : \mathbb{R}^d \to \mathbb{R}^d,\ x \mapsto Cx \text{ with } C \in \mathbb{R}^{d\times d}.$$

So we consider the degenerate parabolic Fokker-Planck equation

$$\partial_t f = Lf := \operatorname{div}(D\nabla f + Cxf) = \operatorname{div}(D\nabla f) + x^T C^T \nabla f + \operatorname{Tr}(C)f, \qquad (2)$$

and analyse solutions that satisfy $f(t,\cdot) \in L^1(\mathbb{R}^d)$ along with $\int_{\mathbb{R}^d} f(t,x)\, \mathrm{d}x = 1$ for all $t > 0$.
A change of variables $y := Mx$ for some orthogonal M (i.e., $M^T = M^{-1}$), leads to the equation

$$\partial_t f = \operatorname{div}(MDM^T \nabla f + MCM^T yf).$$

Since D is symmetric and positive definite, there exists an M such that MDM^T is diagonal. Rescaling the space variable then yields

$$D = \operatorname{diag}\{\underbrace{1,\dots,1}_{k},\underbrace{0,\dots,0}_{d-k}\},$$

$k := \operatorname{rank} D$, $1 \le k \le d$. We will thus assume w.l.o.g. that D has this simple form. The case $k = d$ has been studied extensively, see for example [6].

In the case $k < d$, the operator L is not elliptic, and classical parabolic results will not apply for (2). This motivates investigating hypoellipticity and hypocoercivity for the operator L in (2). We shall see that in this case, hypocoercivity requires hypoellipticity. That is not always the case, as for example with the linear relaxation terms considered in [29].

A very good, broad discussion of hypocoercivity can be found in [67], which also contains a precise definition of hypocoercivity:

Definition 0.1. Let H be a Hilbert space, L an unbounded operator on H with kernel $\mathcal{K}$. Let $\tilde{H}$ be another Hilbert space, which is continuously and densely embedded in $\mathcal{K}^\perp$. Then $-L$ is said to be *hypocoercive* on $\tilde{H}$ if and only if there is $\lambda > 0$ and some constant c such that

$$\forall h \in \tilde{H} : \|e^{tL}h\|_{\tilde{H}} \le ce^{-\lambda t}\|h\|_{\tilde{H}}.$$

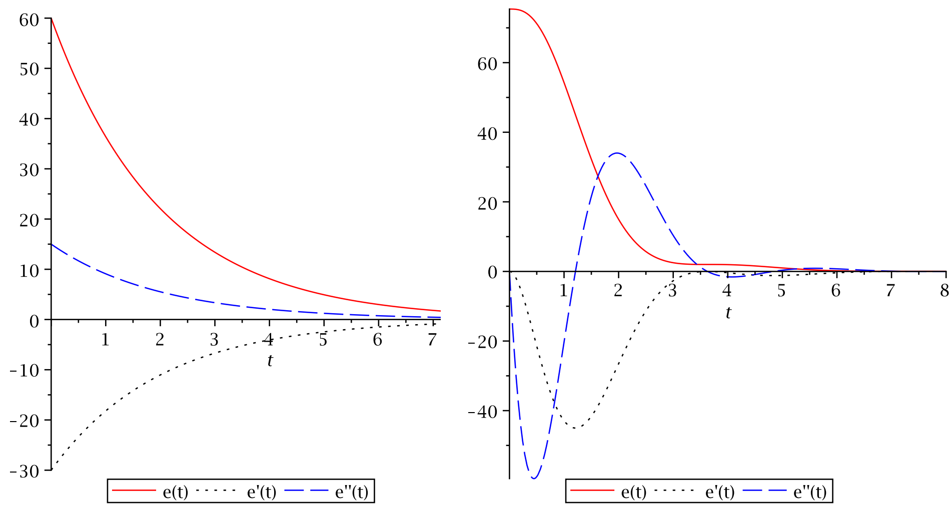
[67] also establishes a general criterion for exponential convergence of solutions for a class of hypocoercive evolution equations, based on a Lyapunov functional equivalent to a weighted H^1-norm. While the main theorem in [67] covers a wide class of problems, the price paid is in the estimate for the decay rate, which is off by orders of magnitude.

In the last few years several papers dealt with the large-time behaviour of hypocoercive equations. But to our knowledge, sharp decay rates were only considered in [34] for two specific toy models, using the spectral decomposition of their generators. In [52] several collisional kinetic models (including the Fokker-Planck equation, linearised Boltzmann and Landau) are analysed on the torus (in the spatial variable): Exponential convergence to the steady state is shown in the H^1–norm. In [51], a decay estimate is obtained for a 2-dimensional kinetic Fokker-Planck model using higher order time derivatives of the L^2-norm of solutions and their space derivative. Also [31] and [10] study dissipative kinetic models (i.e. with $k = \frac{d}{2}$) in H^1. While [31] uses a *macro-micro decomposition* of the models, [10] is based on an (augmented) Γ_2–calculus and local computations (in contrast to the integrated functionals used by most other authors), cf. also [13]. [30] and [10] also analyse much more general hypocoercive equations. Along with [31] they require the following restriction on the interaction between the degenerate dissipative part and the non-symmetric part of L: It is assumed that the map C^T does not map any subspace of the kernel of D into the kernel of D (which is equivalent to using only first order Hörmander-commutators to span all of $\mathbb{R}^d$, cf. §3 in [10]). But this condition is more restrictive than necessary. In this paper and in [67], only the weaker *condition (A)* (see Definition 1.1 below) shall be imposed.

The common approach to study the long-term behaviour of hypocoercive equations has been via a Lyapunov functional - usually on a weighted H^1-space, but [67] also contains (in Theorem 28) a Lyapunov functional based on the logarithmic entropy. In [30], the authors get rid of the H^1-regularity restriction on initial states and prove decay towards the steady state using a modified L^2-norm. In [67], it is shown that even for methods based upon H^1-functionals, one can often get rid of the regularity assumptions by using the regularisation of the semigroup e^{tL}. So far, there is no knowledge on the decay of general entropies “between” logarithmic and quadratic, nor on sharp decay rates for equations of types (2). In this paper we shall modify the entropy method (see [6], [13]- [15]) to achieve all three results for equations of type (2): no H^1-regularity requirement for the initial state, sharp decay rates, and decay for a wide class of relative entropies.

The strategy of the standard entropy method is to derive first a differential inequality between the first and second time derivative of the relative entropy (of the solution w.r.t. the equilibrium state). Their time evolution in a prototypic situation is shown in Fig. 1a. Integration in time of the inequality then allows to deduce exponential decay of the relative entropy. But this approach is not feasible for degenerate Fokker-Planck equations, since the entropy dissipation can vanish for states other than the equilibrium. Hence, the second time derivative of the entropy may change its sign along a trajectory, see Fig. 1b. Therefore, we shall introduce an auxiliary functional – structurally related to the entropy dissipation, but in general larger than the latter. A Bakry-Emery-type estimate then yields exponential decay of this auxiliary functional, and consequently also of the entropy dissipation. A convex Sobolev inequality with the auxiliary functional as its relative Fisher information [6] finally yields the exponential decay of the relative entropy. Initially, this approach shall need an additional regularity assumption for the initial state. But this can then be removed using the regularisation of the parabolic equation (2), as in [67].

Figure 1: Prototypical behaviour of the relative entropy $e(t)$, its first and second derivatives.



(a) Non-degenerate case: The inequalities $e' \leq -\mu e$, $e'' \geq -\mu e'$ can be obtained.

(b) Degenerate case: The inequalities $e' \leq -\mu e$, $e'' \geq -\mu e'$ are wrong, in general.

There is a well understood connection between convex Sobolev inequalities [36], [37] related to the measure $\mu = f_\infty \, \mathrm{d}x$, decay of solutions towards the unique stationary state $f_\infty = c_V \exp(-V)$ in relative entropy, and a Bakry-

Émery condition of the form $\frac{\partial^2 V}{\partial x^2} \geq \lambda_1 D^{-1}$. In the case $k = d$, a convex Sobolev inequality implies decay in the corresponding relative entropy. Conversely, decay in relative entropy implies a convex Sobolev inequality (see [6], §3). The Bakry-Émery condition implies both, but it is not a necessary requirement. For $k < d$, the (classical) Bakry-Émery condition on V can not hold due to the singularity of D. As can be seen from figure 1b, (2) will also not give rise to such inequalities along solution trajectories, as for $k = d$. There is still a convex Sobolev inequality related to the measure $\mu = f_\infty \,\mathrm{d}x$, but it no longer directly implies decay of relative entropy. It still relates the relative entropy to a modified entropy production, and we shall make use of that in §1.3.

This thesis is split in two parts. First, in §1, we establish an entropy method for (2) that proves sharp decay rates for L^1-initial data with finite relative entropy (a weaker condition than L^2-initial data). We start by deriving regularity and positivity of solutions in §1.1. In this case, they derive from the hypoellipticity of L, which we characterise in Lemma 1.3. §1.2 follows this up by explicitly giving the unique (up to normalisation) steady state f_∞ and discussing the operator L in $L^2(\mathbb{R}^d, f_\infty^{-1})$, the standard space for Fokker-Planck equations. In §1.3, we state our main result in Theorem 1.27: a modified entropy method from [6] allows to compute an explicit decay rate for solutions of (2) in relative entropy. We also briefly discuss decay of the Fisher information matrix, which has not yet been done in this context. The sharpness of the decay rate will be shown in Theorem 1.28. In §1.4 we compute the spectrum of L on the weighted space $L^2(\mathbb{R}^d, f_\infty^{-1})$. In §1.5, we discuss our result for three archetypical examples. §1.6 provides an alternative proof for the hypocoercivity of L, adapted from [67]. Finally, we will discuss a few results and problems that appear when generalising to nonlinear drift coefficients F in §1.7. Most of the results in §1 will be published in a separate paper [33].

In the second part, §2, we investigate a possible entropy method for open quantum systems in Lindblad form. Details and an introduction are given in the outline of the same chapter, §2.1.

Chapter 1

Hypocoercive Fokker-Planck equations

1.1 Hypoellipticity of L

If D is not regular, the operator L is neither coercive nor elliptic. In general, such an operator does not lead to a unique normalised stationary state in (2). We thus need additional assumptions on the parameters in L, which shall be assumed throughout §1.1-1.5:

Definition 1.1. The operator L from (2) fulfils ***condition (A)*** if and only if

(i) there is no nontrivial C^T-invariant subspace of $\ker D$,

(ii) the matrix $C \in \mathbb{R}^{d\times d}$ is positively stable[1].

Condition (A) is stricter than necessary for the existence of solutions to (2): the extra condition, positive stability of C, means that the drift part acts as a confinement potential. While there are solutions even without condition (A.ii), there will be no steady state (compare the heat equation on $\mathbb{R}^d$). Indeed, Theorems 1.12 and 1.27 show that condition (A) is both sufficient and necessary for the existence of a unique normalised steady state for (2) and exponential convergence of solutions to the steady state.

[1] A matrix is positively stable iff all eigenvalues have real part greater than zero.

1.1.1 Existence of solutions

Proposition 1.2. *Let $f_0 \in L^1(\mathbb{R}^d)$. Then there is a unique solution $f \in C^\infty(\mathbb{R}^+ \times \mathbb{R}^d)$ of (2) iff no non-trivial subspace of* $\ker D$ *is invariant under C^T.*

Proof: The assumption that no non-trivial subspace of $\ker D$ is invariant under C^T is equivalent to the hypoellipticity of L. We refer to page 148 of [42] for a proof. □

If the condition in Proposition 1.2 does not hold, (2) clearly also has a unique solution, but it would be less regular. Due to the special form of D, we conclude that C cannot be diagonal unless $k = d$. A heuristic explanation of this condition is that the solution cannot stay in the kernel of the dissipative part, and therefore the evolution under (2) acts dissipative in all space directions: If one considers merely the drift part of the equation,

$$f_t = (Cx) \cdot \nabla f, \tag{1.1}$$

the solution is $f(t, x) = f_0(e^{Ct}x)$. So, for the dissipative part to "extend" to the whole space, one needs that $e^{Ct}x$ reaches the whole space $\mathbb{R}^d$ for all $x \in \operatorname{im} D$ ($\operatorname{im} D$ being the image of D), or conversely, that $e^{C^T t}x$ evolves into $\operatorname{im} D$ for all $x \in \ker D$. This is, in fact, an alternative characterisation of the hypoellipticity of L, as shown in Lemma 1.3.

We recall that some approaches from the literature require a stricter condition than in Proposition 1.2: That no subspace of the kernel of D be mapped into the kernel of D by C^T. To illustrate this restriction, consider the examples

$$D_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad C_1^T = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

and

$$D_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad C_2^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

In both cases, there exists a unique steady state and all solutions converge exponentially to the steady state. For the case of D_1 and C_1, the condition given in [30], [31] and [10] hold:

$$\forall\, 0 \neq U \subset \ker D: \; C^T U \not\subset \ker D. \tag{1.2}$$

In the case of D_2 and C_2, the same condition does not hold, even though the problems seem very 'similar'. The difference can be seen as follows: consider a vector of the form $(0,0,0,a)^T$. If we apply C_1^T to this vector, it is moved out of the kernel of D. However, if we apply C_2^T, it is not. In order to move it out of the kernel of D, we need to apply C_2^T twice (i.e. multiply by $(C_2^T)^2$). More precisely, the following weaker version of (1.2) holds:

$$\forall\, 0 \neq U \subset \ker D: \; C^T U \not\subset U.$$

So the condition given in [67], [42] and here is less strict. As will be shown in §1.2.1, condition (A) is equivalent to the existence of a unique normalised steady state.

In the following lemma, we shall give four equivalent characterisations of the hypoellipticity of L. This will allow us to use either characterisation where required. For example, we shall use (iv) for a proof of the positivity of solutions in §1.1.2, and (iii) in the proof of the regularisation for the semigroup e^{Lt} in Theorem 1.26.

Lemma 1.3. *The following four statements are equivalent:*

(i) No non-trivial subspace of $\ker D$ *is invariant under* C^T.

(ii) No eigenvector v *of* C^T *fulfils* $Dv = 0$.

(iii) There exists $\tau \in \{1, \dots, d-k\}$ *and* $\kappa > 0$ *such that*

$$\sum_{j=0}^{\tau} C^j D (C^T)^j \geq \kappa \operatorname{Id}, \tag{1.3}$$

where $k = \operatorname{rank} D$.

(iv) For any $t \in \mathbb{R}$, $h > 0$, *it holds that*

$$\forall 0 \neq \xi \in \ker D \; \exists s \in [t, t+h] \; \exists \eta \in \operatorname{im} D: \langle e^{C^T s} \xi, \eta \rangle = 1.$$

Proof: (i)⇒(ii): If $0 \neq v \notin i\mathbb{R}^d$ is an eigenvector of C^T with $Dv = 0$, then $\operatorname{span}\{v + \overline{v}, v - \overline{v}\}$ is a non-trivial subspace of $\ker D$ invariant under C^T.
(ii)⇒(i): If $V \subset \ker D$ is invariant under C^T, then so is its extension to a

subspace in $\mathbb{C}^d$. But in $\mathbb{C}^d$, any non-trivial subspace invariant under C^T contains at least one eigenvector of C^T.
(i)$\Rightarrow$(iii): All matrices $C^j D(C^T)^j$ are symmetric and positive semidefinite, since D is symmetric and positive semidefinite. It suffices to show that for any vector $v \neq 0$, there exists $j \leq d-k$ with $D^{\frac{1}{2}}(C^T)^j v = D(C^T)^j v \neq 0$, since then $\sum_{j=0}^{\tau}(C)^j D(C^T)^j$ is regular for $\tau := \max_{v\neq 0}\min_{j\in\mathbb{N}}\{j|D^{\frac{1}{2}}(C^T)^j v \neq 0\}$.
If $v \notin \ker D$, we choose $j=0$, and hence $Dv \neq 0$. So let now $0 \neq v \in \ker D$. Then either $C^T v \notin \ker D$, in which case $DC^T v \neq 0$, or $C^T v \in \ker D$. Repeating this procedure, we see that either there is $j \leq d-k$ such that $(C^T)^j v \notin \ker D$ or $\forall 0 \leq j \leq d-k : (C^T)^j v \in \ker D$. Assume the latter. Since the dimension of $\ker D$ is $d-k$, the $d-k+1$ vectors $(C^T)^j v$, $0 \leq j \leq d-k$ are not linearly independent. Thus, $\exists\, l \in \{1, ..., d-k\}$ such that $\operatorname{span}\{v, \dots, (C^T)^l v\} = \operatorname{span}\{v, \dots, (C^T)^{l-1} v\}$. Hence, $\operatorname{span}\{v, \dots, (C^T)^l v\}$ is a C^T-invariant subspace of $\ker D$, which has to be trivial due to condition (A). But then $v = 0$, which is a contradiction.
(iii)$\Rightarrow$(i): If $0 \neq v \in \ker D$, then by (ii) there is a $j \in \{1, \dots, \tau\}$ such that $D^{\frac{1}{2}}(C^T)^j v \neq 0$, i.e. $(C^T)^j v \notin \ker D$. Thus, no non-trivial subspace of $\ker D$ can be invariant under C^T.
(i)$\Rightarrow$(iv): Let $0 \neq \xi \in \ker D$. To proceed by contradiction we assume

$$\forall s \in [t, t+h]\ \forall \eta \in \operatorname{im} D : \langle e^{C^T s}\xi, \eta\rangle = 0. \tag{1.4}$$

This implies

$$\forall s \in [t, t+h]\ : e^{C^T s}\xi \in \ker D,$$

and therefore in particular $\nu := e^{C^T t}\xi \in \ker D$. Differentiating (1.4) with respect to s yields

$$\forall s \in [t, t+h]\ \forall \eta \in \operatorname{im} D : \langle e^{C^T s} C^T \xi, \eta\rangle = 0. \tag{1.5}$$

But this implies $C^T \nu \in \ker D$. Differentiating (1.5) repeatedly with respect to s yields $(C^T)^j \nu \in \ker D$ for any $0 \leq j \leq d-1$. Hence $\operatorname{span}\{\nu, \dots, (C^T)^{d-1}\nu\} \subset \ker D$ is a C^T-invariant subspace of $\ker D$. That is a contradiction to condition (A).
(iv)$\Rightarrow$(i): Let $\xi \neq 0$ be in a C^T-invariant subspace of $\ker D$, i.e. $(C^T)^j \xi \in \ker D$ for any $j \in \mathbb{N}_0$. Since $C^T \in \mathbb{R}^{d\times d}$, $e^{C^T s}$ is a polynomial in C^T (albeit with s-dependent coefficients) and (iv) immediately gives a contradiction. $\square$

Remark 1.4. If τ is the minimal constant for which Lemma 1.3 holds, then L

fulfils the finite rank Hörmander condition of order τ (see [42] Theorem 1.1). Using $\tau = d - k$ summands in (1.3) actually covers the worst-case scenario. But in many examples, $\sum_{j=0}^{\tau}(C)^j D(C^T)^j$ with $\tau < d - k$ is already positive definite. This is the case in the kinetic approaches discussed in [10], and [31], which require $\tau = 1$ and $k = \frac{d}{2}$. Also in [30], $\tau = 1$ is assumed.

We also want our solution to be in $C(\mathbb{R}_0^+, L^1(\mathbb{R}^d))$, since then the normalisation

$$\int_{\mathbb{R}^d} f(t,x)\ \mathrm{d}x = 1$$

holds for all t due to the divergence form of the operator. This fact follows from the existence of a Green's function for (2), which we will construct in the following lemma. We note that this construction has already been done in [42], with slightly different notation.

Lemma 1.5. *Let condition (A.i) be fulfilled. Then the Green's function g to (2) is given by*

$$g(t) = \frac{1}{(2\pi)^{\frac{d}{2}}\det(W(t))}\exp(-x^T W(t)^{-1}x),$$

where

$$W(t) = \int_0^t e^{C(s-t)} D e^{C^T(s-t)}\ \mathrm{d}s$$

is positive definite for all $t > 0$.

Proof: The Fourier transform of (2) is

$$\hat{g}_t = -(\xi^T D\xi)\hat{g} - (\xi^T C\nabla)\hat{g}. \tag{1.6}$$

We are looking for a solution to the initial condition $\hat{g}_0 \equiv 1$. As an ansatz, we take

$$\hat{g}(t,\xi) := \exp(-\xi^T W(t)\xi)$$

with a symmetric matrix W that is positive definite for all $t > 0$ and fulfills $W(t = 0) = 0$. Inserting this into (1.6), we get

$$-(\xi^T W_t\xi)\hat{g} = (-\xi^T D\xi)\hat{g} + 2(\xi^T CW\xi)\hat{g},$$

which implies due to $\tilde{g} > 0$

$$\xi^T W_t \xi = \xi^T D\xi - \xi^T (CW + WC^T)\xi.$$

This can only be fulfilled if

$$W_t = D - CW - WC^T. \tag{1.7}$$

Equation (1.7) together with the initial condition $W(t = 0) = 0$ has the unique solution

$$W(t) = \int_0^t e^{C(s-t)} D e^{C^T(s-t)} \, \mathrm{d}s.$$

We need to prove: $g(t, \cdot) \in L^1(\mathbb{R}^d)$ and $g(t, x) > 0$ for $t > 0$, $x \in \mathbb{R}^d$. To this end it remains to show that $W(t) > 0$ for all $t > 0$. Clearly $W(t) \geq 0$, since it is an integral over positive semidefinite matrices. So assume that $W(t_0)$ is singular for some $t_0 > 0$, i.e. there exists $0 \neq \xi \in \mathbb{R}^d$ such that

$$0 = \xi^T W(t_0)\xi = \int_0^{t_0} \xi^T e^{c(s-t_0)} D e^{C^T(s-t_0)} \xi \, \mathrm{d}s,$$

where the integrand is non-negative. Due to the continuity of the matrix exponential, this can only hold if

$$\xi^T (e^{C(s-t_0)} D e^{C^T(s-t_0)} \xi = 0$$

for all $s \in [0, t_0]$. Since $D = D^2$, this implies

$$De^{-C^T r}\xi = 0$$

for all $r \in [0, t_0]$, in particular $\xi \in \ker D$. But this is a contradiction to Lemma 1.3. Hence, $W(t) > 0$ for all $t > 0$, and an inverse Fourier transformation of $\hat{g}$ gives

$$g(t, x) = \frac{1}{(2\pi)^{\frac{d}{2}} \det(W(t))} \exp(-x^T W(t)^{-1} x).$$

$\square$

We now state an existence result on solutions in L^p, which is similar to Corollary 3.1 from [62]:

Corollary 1.6. *Let condition (A.i) be fulfilled. Let* $f_0 \in L^1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$, *for a* $p \in [1, \infty]$. *Then there exists a unique classical solution* f *to (2) with* $f \in C([0, \infty), L^p(\mathbb{R}^d)) \cup C^\infty(\mathbb{R}^+ \times \mathbb{R}^d)$. *If* $\int_{\mathbb{R}^d} f_0 \, \mathrm{d}x = 1$, *it follows* $\int_{\mathbb{R}^d} f(t) \, \mathrm{d}x = 1$ *for all* $t > 0$.

Proof: From Proposition 1.2 we already have that a solution f is smooth for any $t > 0$. With the Green's function from Lemma 1.5 we obtain

$$f(t, \cdot) = g(t, \cdot) * f_0.$$

Applying Young's inequality yields

$$\|f(t)\|_{L^p(\mathbb{R}^d)} = \|g(t) * f_0\|_{L^p(\mathbb{R}^d)} \le \|f_0\|_{L^p(\mathbb{R}^d)} \|g(t)\|_{L^1(\mathbb{R}^d)}.$$

From the proof of Lemma 1.5 we obtain $\|g(t)\|_{L^1(\mathbb{R}^d)} = 1$. The desired normalisation then follows from the divergence form of the operator. □

This answers the question of existence of solutions.

1.1.2 Global positivity

For a *non-degenerate* Fokker-Planck equation, the solution is globally positive for any $t > 0$. This follows from a strong maximum principle supplied by the fully parabolic operator. In our degenerate case, a strong classic maximum principle does not hold. From a weak maximum principle, we obtain that the solution is nonnegative. However, global positivity is important, since most admissible entropies are only defined on positive functions. Under condition (A.i), we can derive global positivity of solutions from the hypoellipticity of L:

Theorem 1.7. *Let condition (A.i) hold and* $f_0 \in L^1_+(\mathbb{R}^d)$. *Let* f *be the solution to (2). Then*

$$\forall t > 0 \ \forall x \in \mathbb{R}^d : f(t, x) > 0.$$

This theorem follows directly from the strict positivity of the Green's function g from Lemma 1.5. However, we give a second proof via a *sharp maximum principle* from [41]. As the second approach is more general, it is more promising for an extension of these results to a nonlinear coefficient F. We need to introduce some notation.

First, we rewrite our operator in degenerate elliptic form:

$$\tilde{L}f := \left[\begin{pmatrix} \partial_t \\ \nabla \end{pmatrix}^T \tilde{D} \begin{pmatrix} \partial_t \\ \nabla \end{pmatrix} \right] f + b \cdot \begin{pmatrix} \partial_t \\ \nabla \end{pmatrix} f,$$

where

$$\tilde{D} := \begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix} \in \mathbb{R}^{(d+1)\times(d+1)},$$
$$b(x) := \begin{pmatrix} -1 \\ Cx \end{pmatrix} \in \mathbb{R}^{d+1}.$$

Comparing this with our original operator L, we have

$$\tilde{L}f = Lf - f_t - \mathrm{Tr}(C)f. \tag{1.8}$$

Due to the special form of D, the rows d_j of $\tilde{D}$ are of the form

$$(d_j)_l = \delta_{jl}$$

for $2 \le j \le k+1$ ($k = \operatorname{rank} D$) and $d_j = 0$ for $j = 1$, $k+1 < j \le d+1$. With this notation, we shall now introduce drift and diffusion trajectories:

Definition 1.8. Let Ω be a connected open set in $\mathbb{R}^{d+1}$, $p_0 \in \Omega$.

- If $p(s)$ is the solution to

$$\tfrac{\mathrm{d}}{\mathrm{ds}}p(s) = d_j,$$
$$p(0) = p_0,$$

  with some $2 \le j \le k+1$, and if $p(s) \in \Omega$ for $s_1 \le s \le s_2$ with some $s_1 < 0 < s_2$, then we call $\Gamma := \{p(s) | s_1 \le s \le s_2\}$ a *diffusion trajectory* running through p_0.

- If $p(s)$ is the solution to

$$\tfrac{\mathrm{d}}{\mathrm{ds}}p(s) = b(p(s)),$$
$$p(0) = p_0,$$

with $b(p(s)) \neq 0$, and if $p(s) \in \Omega$ for $0 \leq s \leq s'$ with some $s' > 0$, then we call $\Gamma := \{p(s)|0 \leq s \leq s'\}$ a *drift trajectory* starting at p_0.

Remark: Drift trajectories are oriented at p_0 in the direction $b(p_0)$; they do not run both ways. Diffusion trajectories are not oriented, they run in both directions. In our special case of a diagonal D, each diffusion trajectory moves along one of the canonical unit vectors in $\operatorname{im} D$.

Next, we introduce the propagation set:

Definition 1.9. Let $\Omega \subset \mathbb{R}^{d+1}$. Two points $p, q \in \mathbb{R}^{d+1}$ are connected by a diffusion trajectory in Ω iff there is some diffusion trajectory $\Gamma \subset \Omega$ with $p, q \in \Gamma$. q is connected to p by a drift trajectory in Ω iff there is a drift trajectory $\Gamma \subset \Omega$ starting at p with $q \in \Gamma$.
For any point $p \in \Omega$, the *propagation set* $S(p, \Omega)$ consists of all $q \in \Omega$ that are connected to p by a finite series of drift and diffusion trajectories.

Again, note that drift trajectories are oriented and cannot connect points in the 'backward direction'. Therefore, it is possible that $q \in S(p, \Omega)$ while $p \notin S(q, \Omega)$.
With this notation, we can restate the interior maximum principle from Theorem 1 of [41]:

Theorem 1.10. *Let $p = (t, x) \in \Omega \subset \mathbb{R}^{d+1}$, $t > 0$. Let the function $f \in C^2(\Omega)$ satisfy $\tilde{L}f \leq 0$ on the propagation set $S(p, \Omega)$ and*

$$\inf_{S(p,\Omega)} f \geq 0.$$

If $f(p) = 0$, then $f = 0$ in $\overline{S(p, \Omega)}$.

The propagation set corresponding to equation (2) can be characterised as follows:

Lemma 1.11. *Let $p = (t, x) \in \mathbb{R}^{d+1}$. Then $S(p, \mathbb{R}^{d+1}) = [0, t) \times \mathbb{R}^d \cup \{(t, x_0)\} \times \mathbb{R}^k$, where x_0 is the orthogonal projection of x onto the kernel of D.*

Proof: First, note that only drift-trajectories are non-constant in time, since the first row of $\tilde{D}$ is zero. A drift trajectory $\xi(s) = (t(s), v(s))$ starting at $\xi_0 = (t_0, v_0)$ has the form

$$\frac{\mathrm{d}}{\mathrm{ds}}\xi = \begin{pmatrix} -1 \\ Cv \end{pmatrix},$$
$$\xi(0) = \xi_0.$$

The solution to this equation is

$$\xi(s) = \begin{pmatrix} t_0 - s \\ e^{Cs} v_0 \end{pmatrix}.$$

This means that drift trajectories move backwards in time linearly. Thus, for a point $q = (t', y)$ to be connected to p in the time-variable, it is necessary that $t' \leq t$. This is to be expected, as it is also the case for the classical maximum principle for parabolic equations.
Since the diffusion trajectories span the subspace $\mathbb{R}^\tau = \operatorname{im} D \subset \mathbb{R}^d$, we write $p = (t, x_D, x_0)$ and $q = (t', y_D, y_0)$, where x_0 and y_0 are the projection of x and y onto the kernel of D. Without moving backwards in time, we can only connect via diffusion trajectories. This implies

$$S(p, \mathbb{R}^{d+1}) \cap \{(\tilde{t}, x) \in \mathbb{R}^{d+1} | \tilde{t} = t\} = \{(t, x_0)\} \times \mathbb{R}^\tau.$$

It remains to show that any point $q = (t', y)$ with $t' < t$ can be connected to p. The strategy here is the following: Since we can freely move around in $\operatorname{im} D$, we only need to connect q and p in the kernel of D and in time. To achieve this, we employ Lemma 1.3, (iv). We will proceed in a series of trajectories: A number of drift trajectories (equal to $\mu := \dim\ker D + 1 = d - k + 1$), each of them followed by up to $k = \operatorname{rank} D$ diffusion trajectories. Starting at $\xi_0 = (t, x)$, such a series of two drift and $2k$ diffusion trajectories will arrive at

$$\big(t - s_1 - s_2, e^{Cs_2}[e^{Cs_1}x + z_1] + z_2\big),$$

where $z_1, z_2 \in \operatorname{im} D$ are the results of shifts by diffusion trajectories and $0 \leq s_1, s_2$. Thus, a series of μ trajectories will arrive at

$$\big(t - \sum_{j=1}^{\mu} s_j, \exp(C\sum_{j=1}^{\mu} s_j)x + \sum_{j=1}^{\mu-1} \exp(C\sum_{l=1+j}^{\mu} s_l)z_j + z_\mu\big),$$

where $z_j \in \operatorname{im} D$, $1 \leq j \leq \mu$. Setting this equal to our target point $q = (t', y)$ and rearranging terms, we obtain the following requirements:

$$\sum_{j=1}^{\mu} s_j \overset{!}{=} t - t',$$
$$\sum_{j=1}^{\mu-1} e^{Cr_j} z_j + z_\mu \overset{!}{=} y - e^{C(t-t')}x,$$

with $r_j \in [0, t - t']$, $r_j = \sum\limits_{l=j+1}^{\mu} s_l$, $s_j \geq 0$. The projection of this equation onto

$\operatorname{im} D$ can always be solved by setting z_μ accordingly. For the projection onto $\ker D$, we get

$$(\operatorname{Id} - D) \sum_{j=1}^{\mu-1} e^{Cr_j} z_j \stackrel{!}{=} (\operatorname{Id} - D)(y - e^{C(t-t')}x) =: v_0 \in \ker D.$$

The left-hand side can be seen as a linear mapping from $(\operatorname{im} D)^{\mu-1}$ to $\ker D$, since each of the matrix exponentials can take an arbitrary argument $z_j \in \operatorname{im} D$. So we need to show that

$$\begin{aligned}((\operatorname{Id} - D)e^{Cr_j})_{1\le j\le\mu} : (\operatorname{im} D)^{\mu-1} &\to \ker D, \\ (z_j)_{1\le j\le \mu-1} &\mapsto (\operatorname{Id} - D)\sum_{j=1}^{\mu-1} e^{Cr_j} z_j\end{aligned} \tag{1.9}$$

is surjective for some choice of $0 \le r_{\mu-1} < r_{\mu-2} < \cdots < r_1 \le t - t'$. Let $r_1 \in [\frac{t-t'}{2}, t - t']$. Then either

$$(\operatorname{Id} - D)e^{Cr_1} : \operatorname{im} D \to \ker D$$

is surjective, or there is $\xi \in \ker D$ with $\xi \perp (\operatorname{Id} - D)e^{Cr_1} \operatorname{im} D$ (since the image of a linear map is always a linear subspace). But then from Lemma 1.3 there is $r_2 \in (0, r_1)$ and $\eta \in \operatorname{im} D$ with

$$\langle (\operatorname{Id} - D)e^{Cr_2}\eta, \xi\rangle = \langle \eta, e^{C^T r_2}\xi\rangle = 1.$$

Now, since $\xi \not\perp (\operatorname{Id} - D)e^{Cr_2} \operatorname{im} D$, we have

$$\dim \operatorname{span}\big[(\operatorname{Id} - D)e^{Cr_1} \operatorname{im} D, (\operatorname{Id} - D)e^{Cr_2} \operatorname{im} D\big] > \dim(\operatorname{Id} - D)e^{Cr_1} \operatorname{im} D.$$

Then either

$$\big((\operatorname{Id} - D)e^{Cr_1}, (\operatorname{Id} - D)e^{Cr_2}\big) : \operatorname{im} D \times \operatorname{im} D \to \ker D.$$

is surjective, or we repeat the process. Each repetition increases the dimension of the reachable subspace of $\ker D$ by at least one. Thus, we will need at most $\mu - 1 = \dim \ker D$ iterations, and hence (1.9) is surjective. □

To see that the solution f of (2) fulfils $f \ge 0$, one can employ the same method used for the classical maximum principle for non-degenerate parabolic equations. Now we give the proof of Theorem 1.7 via the sharp maximum principle from [41]:

Proof (of Theorem 1.7): Let $f \in C^2(\mathbb{R}^+ \times \mathbb{R}^d)$ be a solution to (2) for some $f_0 \geq 0$ with $\int_{\mathbb{R}^d} f_0 \, \mathrm{d}x = 1$. Then

$$g(t) := e^{-\beta t} f(t),$$

with $\beta > |\mathrm{Tr}(C)|$, solves

$$\begin{aligned} g_t - \mathrm{div}(D\nabla g) - (Cx)^T \nabla g + (\beta - \mathrm{Tr}(C))g = 0, \\ g(t=0) = f_0 \geq 0, \end{aligned} \tag{1.10}$$

where $\beta - \mathrm{Tr}(C) > 0$. The classical maximum principle then shows that $g(x,t) \geq 0$ for $t \geq 0$. From (1.8), we have

$$\tilde{L}g = Lg - g_t - \mathrm{Tr}(C)g = (-\mathrm{Tr}(C) - \beta)g \leq 0,$$

since $\beta > |\mathrm{Tr}(C)|$. Assume $g(t', x') = 0$ for some $t' > 0$, $x' \in \mathbb{R}^d$. Then Theorem 1.10 gives $g = 0$ on $[0, t') \times \mathbb{R}^d$ and in particular $g(0) = f_0 = 0$. But this is a contradiction to

$$\int_{\mathbb{R}^d} f_0(x) \, \mathrm{d}x = 1.$$

Hence, $g(x,t) > 0$ and also $f(t,x) > 0$ for all $t > 0, x \in \mathbb{R}^d$. $\square$

1.2 Steady State, weighted L^2-space

1.2.1 Existence of a steady state

In light of Theorem 1.7, we are looking for a steady state f_∞ of (2) that fulfils the conditions

$$\int_{\mathbb{R}^d} f_\infty(x) \, \mathrm{d}x = 1, \quad \forall x \in \mathbb{R}^d : \ f_\infty(x) > 0. \tag{1.11}$$

In fact, the existence of such a steady state is equivalent to condition (A):

Theorem 1.12. *There exists a unique steady state $f_\infty \in L^1(\mathbb{R}^d)$ of (2) fulfilling (1.11) iff condition (A) holds.*
Moreover, this steady state is of the (non-isotropic) Gaussian form

$$f_\infty(x) = c_K \exp(-\tfrac{x^T K^{-1} x}{2}),$$

where K is the unique, symmetric, and positive definite solution to the continuous Lyapunov equation

$$2D = CK + KC^T, \tag{1.12}$$

and c_K is the normalisation constant.

For the proof of Theorem 1.12 we consider the Fourier transform of (2):

$$\begin{aligned} \hat{f}_t(t,\xi) &= -(\xi^T D\xi)\hat{f}(t,\xi) - (C^T\xi)\cdot(\nabla_\xi \hat{f}(t,\xi)), \\ \hat{f}(t=0) &= \hat{f}_0. \end{aligned} \tag{1.13}$$

A steady state $f_\infty \in L^1(\mathbb{R}^d)$ implies $\hat{f}_\infty \in C_0(\mathbb{R}^d)$. Also note that

$$\hat{f}_\infty(0) = \int_{\mathbb{R}^d} f_\infty(x)\ \mathrm{d}x = 1$$

for the normalised steady state.

Thus, the steady state equation in Fourier space reads

$$\begin{aligned} 0 &= (\xi^T D\xi)\hat{f}_\infty(\xi) + (C^T\xi)\cdot\nabla_\xi \hat{f}_\infty(\xi), \\ \hat{f}_\infty(0) &= 1. \end{aligned} \tag{1.14}$$

The problem at hand is closely related to the stationary Fokker-Planck equation in section 2.2 in [4]. But for $k < d$, the singularity of D requires a more careful analysis.

We will split the proof of Theorem 1.12 into three lemmas: In Lemmas 1.13 and 1.15 we establish that existence of a steady state is equivalent to condition (A). Lemma 1.14 establishes that the steady state is Gaussian. In the proof of these Lemmas, we shall switch between the equivalent characterisations (i) and (ii) for condition (A.i) in Lemma 1.3.

Lemma 1.13. *Let (1.14) have a unique solution $\hat{f}_\infty \in C_0(\mathbb{R}^d)$. Then condition (A) holds.*

Proof: First, we shall show that C^T is regular: if C^T has a non-trivial kernel, (1.14) restricted to the kernel of C^T reads

$$\forall \xi \in \ker C^T : (\xi^T D\xi)\hat{f}_\infty(\xi) = 0. \tag{1.15}$$

Now, either $\ker C^T \subset \ker D$, which would mean that both drift and diffusion in (2) only act on a proper subspace of $\mathbb{R}^d$ and there would be no unique steady state; or (1.15) implies

$$\exists v \in \mathbb{R}^d : \ \forall s \in \mathbb{R}\backslash\{0\} : \hat{f}_\infty(sv) = 0.$$

Hence, $f_\infty(0) = 0$ by continuity, which is a contradiction to $\hat{f}_\infty(0) = 1$. So C^T is regular.

Next, we will show that C is positively stable, i.e. that all eigenvalues have a strictly positive real part. The characteristic equations for (1.14) are

$$\begin{aligned} \dot{\xi}(s) &= C^T\xi(s), \quad s \in \mathbb{R}, \\ \dot{z}(s) &= -(\xi(s)^T D\xi(s))z(s), \quad s \in \mathbb{R}, \\ (z(0), \xi(0)) &= (z_0, \xi_0) \in \mathbb{R}^{d+1}. \end{aligned} \tag{1.16}$$

The solutions to these equations are

$$\begin{aligned} \xi(s) &= e^{C^T s}\xi_0, \\ z(s) &= z_0 \exp\left(-\int_0^s \xi(\tau)^T D\xi(\tau)\ \mathrm{d}\tau \right). \end{aligned}$$

Assume that C has an eigenvalue λ with $\Re\{\lambda\} < 0$. Let v be a corresponding eigenvector of C^T, i.e. $C^T v = \lambda v$, chosen such that $v \notin i\mathbb{R}^d$. Consider the characteristic curve starting at $\xi_0 := v + \bar{v} \neq 0$:

$$\xi(s) = e^{\lambda s}v + e^{\bar{\lambda}s}\bar{v}.$$

Then

$$|\xi(s)| = e^{\Re\{\lambda\}s}|v + e^{2i\Im\{\lambda\}s}\bar{v}| \to \infty,\ s \to -\infty,$$

which implies

$$\forall s \leq 0 : |z(s)| = |z_0| \exp\left(\int_s^0 \xi(\tau)^T D\xi(\tau)\ \mathrm{d}\tau \right) \geq |z_0|$$

due to D being positive semidefinite. If $z_0 \neq 0$, this is a contradiction to $|\hat{f}_\infty(\xi)| \to 0$, $|\xi| \to \infty$. If $z_0 = 0$, we can take the limit $s \to \infty$ and obtain a contradiction to $\hat{f}_\infty(0) = 1$ and the continuity of $\hat{f}_\infty$. So C cannot have eigenvalues with negative real part.

Now assume that C has a purely imaginary eigenvalue. Then there exist characteristics $\xi(s)$ which form circles. Due to

$$z(s) = z_0 \exp\left(- \int_0^s \xi(\tau)^T D \xi(\tau) \, \mathrm{d}\tau \right)$$

and the continuity of $\hat{f}_\infty$, one of the following statements has to hold on any such characteristic curve:

(a) $\forall s \in \mathbb{R} : \xi(s) \in \ker D$,

(b) $z_0 = 0$.

If (a) holds, then we have $z(s) = z_0$ on this characteristic. Since the characteristic is closed, there will be no uniqueness of $\hat{f}_\infty$.

So (b) holds, and for any ε we can find such a characteristic starting at a vector ξ_0 with $|\xi_0| < \varepsilon$. But then $\hat{f}_\infty(\xi(s)) = z_0 = 0$, which is a contradiction to the continuity of $\hat{f}_\infty$ at 0.

This shows that C has to be positively stable. It remains to show the second part of condition (A). So assume C^T has an eigenvector $v \notin i\mathbb{R}^d$ with $Dv = 0$. Then $D\bar{v} = 0$, and for the characteristic starting at $\xi(0) = v + \bar{v}$, we have

$$\begin{aligned}
\xi(s) &= e^{\lambda s} v + e^{\bar{\lambda} s} \bar{v}, \\
z(s) &= z(0) \exp\left(- \int_0^s (e^{\lambda\tau} v + e^{\bar{\lambda}\tau} \bar{v})^T D (e^{\lambda\tau} v + e^{\bar{\lambda}\tau} \bar{v}) \, \mathrm{d}\tau \right) = z(0).
\end{aligned}$$

This means that z is constant on the characteristic ξ. Now, since C is positively stable,

$$\begin{aligned}
\lim_{s \to \infty} |\xi(s)| &= \infty, \\
\lim_{s \to -\infty} |\xi(s)| &= 0.
\end{aligned}$$

So we would need $z(0) = 1$ because of the continuity in 0, and $z(0) = 0$ because $\hat{f}_\infty \in C_0(\mathbb{R}^n)$. That is a contradiction, so there can be no eigenvector v of C^T with $Dv = 0$. $\square$

Lemma 1.14. *Let C be positively stable. Then, the function*

$$\hat{f}_\infty(\xi) := \exp(-\tfrac{\xi^T K \xi}{2})$$

is a solution to (1.14), where $K \geq 0$ is the unique solution of (1.12). Furthermore, K is regular iff no eigenvector v of C^T satisfies $Dv = 0$. In this case, f_∞ is Gaussian and hence in $L^1(\mathbb{R}^d)$.

Proof: We insert the ansatz

$$\hat{f}_\infty(\xi) = \exp(-\tfrac{\xi^T K \xi}{2})$$

with a symmetric matrix $K \in \mathbb{R}^{d\times d}$ into (1.14) and obtain

$$\forall \xi \in \mathbb{R}^d : 0 = \Big(\xi^T D\xi - (C^T\xi)\cdot(K\xi)\Big)\hat{f}_\infty,$$

which holds iff

$$\forall \xi \in \mathbb{R}^d : 0 = \xi^T(D - CK)\xi.$$

This in turn holds iff $D - CK$ is antisymmetric, which is equivalent to

$$D - CK = KC^T - D,$$

and thus to (1.12). This continuous Lyapunov equation has a unique, symmetric and positive semidefinite solution K since C is positively stable (see Theorem 2.2 in [61], Theorem 2.2.3 in [44]).
Now assume that K is not regular. Then there is a $v \neq 0$ with $Kv = 0$ and (1.12) implies

$$\begin{aligned} 2v^T Dv &= v^T CKv + v^T KC^T v = 0. \\ \Rightarrow 0 = 2Dv &= CKv + KC^T v = KC^T v, \end{aligned}$$

so $C^T v$ is also an eigenvector of K to the eigenvalue 0. Since $v \neq 0$ and C^T is regular, $C^T v \neq 0$. Repeating this calculation with $C^T v$ instead of v, we can see that $C^T v$ is in the kernel of D, and thus $(C^T)^2 v$ is in the kernel of K. A proof by induction then gives $(C^T)^k v \in \ker D \cap \ker K$ for all $k \in \mathbb{N}$. Therefore, the space

$$V := \operatorname{span}[v, \ldots, (C^T)^{d-1} v]$$

is a C^T-invariant subspace of $\ker D$. So K is regular if there is no eigenvector v of C^T with $Dv = 0$.

For the reversed implication, assume that there is an eigenvector v of C^T (corresponding to the eigenvalue λ_v) with $Dv = 0$. This implies

$$\begin{aligned} 0 &= 2v^T Dv = v^T CKv + v^T KC^T v = \overline{\lambda_v} v^T Kv + \lambda_v v^T Kv \\ &= 2\Re\{\lambda_v\} v^T Kv. \end{aligned}$$

Since $\Re\{\lambda_v\} > 0$ for all eigenvalues of C^T, it follows that $v^T Kv = 0$ and thus, since K is symmetric, it is not regular. $\square$

Lemma 1.15. *Let condition (A) hold. Then the steady state f_∞ from Lemma 1.14 is unique.*

Proof: We will show that the characteristic equations (1.16) have a unique solution fulfilling (1.11). As the starting manifold for the characteristics, we take $\Gamma := \{\xi_0 \in \mathbb{R}^d : |\xi_0| = 1\}$, which is admissible since C is positively stable. The characteristic curve starting at ξ_0 is

$$\xi(s) = e^{C^T s} \xi_0.$$

This implies

$$|\xi(s)|^2 = \langle \xi_0, e^{Cs} e^{C^T s} \xi_0 \rangle,$$

which yields with the positive stability of C:

$$e^{\eta} |\xi_0| \le |\xi(s)| \le e^{\tilde{\eta}} |\xi_0|.$$

Here, $0 < \eta$ is the smallest real part of the eigenvalues of C, and $\tilde{\eta}$ the largest. Hence

$$\begin{aligned} \lim_{s \to \infty} |\xi(s)| &= \infty, \\ \lim_{s \to -\infty} |\xi(s)| &= 0, \end{aligned}$$

and the characteristic curves cover all of $\mathbb{R}^d$.
The value of solutions along the characteristics is

$$z(s) = z(0) \exp\left(- \int_0^s \xi(\tau)^T D \xi(\tau) \, \mathrm{d}\tau \right).$$

So taking

$$z(0) = \exp\left(-\int\limits_{-\infty}^{0} \xi(\tau)^T D \xi(\tau) \, \mathrm{d}\tau \right)$$

as initial condition implies $1 = \lim_{s \to -\infty} z(s) = \hat{f}_\infty(0)$. Since $\xi(s)$ decays exponentially for $s \to -\infty$, $z(0)$ is always finite and there is a unique solution $z(s)$. □

This lemma completes the proof of Theorem 1.12.

1.2.2 Decomposition of the generator L

In analogy to the entropy method for linear, nondegenerate Fokker-Planck equations presented in [6], we now consider (2) in the weighted space $L^2 := L^2(\mathbb{R}^d, f_\infty^{-1})$ with inner product $\langle \cdot, \cdot \rangle$. On this space, the operator $L = \operatorname{div}(D\nabla \cdot +Cx \cdot)$ can be decomposed very naturally.

Theorem 1.16. *Let (2) fulfil condition (A). Consider L on the weighted space L^2. Then L can be decomposed into its symmetric part L_s and its antisymmetric part L_{as} as*

$$\begin{aligned} L_s f &= \operatorname{div}(D\nabla f + DK^{-1}xf) = \operatorname{div}(D\nabla(\tfrac{f}{f_\infty})f_\infty), \\ L_{as} f &= x^T T \nabla f = \operatorname{div}(R\nabla(\tfrac{f}{f_\infty})f_\infty). \end{aligned} \tag{1.17}$$

Here, $R := \frac{1}{2}(CK - KC^T)$ is antisymmetric, K is the covariance matrix of f_∞ from Theorem 1.12, and $T := -K^{-1}R = \frac{1}{2}(C^T - K^{-1}CK) \in \mathbb{R}^{d\times d}$ with $\operatorname{Tr}(T) = 0$.

Remark: 1. Note that the steady state f_∞ fulfils both $L_s f_\infty = 0$ and $L_{as} f_\infty = 0$.
2. $R \neq 0$ and hence (2) is non-symmetric in L^2. Otherwise (1.12) would imply $D = KC^T$ and $\ker D = \ker C^T$, which contradicts condition (A).

Proof (of Theorem (1.16):
We compute

$$\begin{aligned} \langle Lf, g\rangle &= \int\limits_{\mathbb{R}^d} (Lf) g \exp(\tfrac{x^T K^{-1} x}{2}) \, \mathrm{d}x \\ &= -\int\limits_{\mathbb{R}^d} [D\nabla f + Cxf] \cdot [\nabla g + K^{-1}xg] \exp(\tfrac{x^T K^{-1} x}{2}) \, \mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^d} f \operatorname{div}\left[(D\nabla g + DK^{-1}xg)\exp(\tfrac{x^TK^{-1}x}{2})\right] \mathrm{d}x \\
&- \int_{\mathbb{R}^d} f x^T C^T \left[\nabla g + K^{-1}xg\right] \exp(\tfrac{x^TK^{-1}x}{2}) \mathrm{d}x \\
&= \int_{\mathbb{R}^d} f \left[\operatorname{div}(D\nabla g + DK^{-1}xg) + x^TK^{-1}D\nabla g\right] \exp(\tfrac{x^TK^{-1}x}{2}) \mathrm{d}x \\
&+ \int_{\mathbb{R}^d} f \left[x^TK^{-1}DK^{-1}xg - x^TC^T\nabla g - x^TC^TK^{-1}xg\right] \exp(\tfrac{x^TK^{-1}x}{2}) \mathrm{d}x.
\end{aligned}
$$

Using (1.12), we have $K^{-1}DK^{-1} - C^TK^{-1} = K^{-1}RK^{-1}$. Since R is antisymmetric it follows that $x^T(K^{-1}DK^{-1} - C^TK^{-1})x = 0$ and hence

$$L^*g = \operatorname{div}(D\nabla g + DK^{-1}xg) + x^TK^{-1}D\nabla g - x^TC^T\nabla g.$$

Furthermore, $\operatorname{Tr}(DK^{-1} - C) = \operatorname{Tr}((D - CK)K^{-1}) = 0$, since $D - CK = -R$ is antisymmetric and K^{-1} is symmetric. Thus we can write (using (1.12) in the last step)

$$
\begin{aligned}
L^*g &= \operatorname{div}(D\nabla g + DK^{-1}xg + (DK^{-1} - C)xg) \\
&= \operatorname{div}(D\nabla g + (2DK^{-1} - C)xg) \\
&= \operatorname{div}(D\nabla g + (KC^TK^{-1}x)g).
\end{aligned}
$$

So we get, again using (1.12),

$$
\begin{aligned}
L_s f &= \tfrac{L+L^*}{2} f \\
&= \operatorname{div}(D\nabla f + \tfrac{1}{2}(C + KC^TK^{-1})xf) = \operatorname{div}(D\nabla f + DK^{-1}xf) \\
&= \operatorname{div}(D\nabla(\tfrac{f}{f_\infty})f_\infty); \\
L_{as} f &= \tfrac{L-L^*}{2} f \\
&= \operatorname{div}(\tfrac{1}{2}(C - KC^TK^{-1})xf) = \operatorname{div}(RK^{-1}xf) \\
&= \operatorname{div}(R\nabla(\tfrac{f}{f_\infty})f_\infty),
\end{aligned}
$$

where we have used $\operatorname{div}(R\nabla f) = 0$ for the last equality. $\square$

1.3 Entropy Method

In this section, we will prove decay of solutions f of (2) to f_∞ in relative entropy under condition (A). We will also compute a sharp rate for this decay. To do so, we consider relative entropies, as in [6]. We will see that, unlike in the fully parabolic case, a direct entropy-entropy dissipation estimate cannot be obtained. Instead, we establish an auxiliary functional that bounds the entropy dissipation (in §1.3.1). We then prove decay of this modified functional in §1.3.3 by establishing a replacement for the classical Bakry-Émery condition. By a convex Sobolev inequality, this still implies a decay rate for the relative entropy, initially at the price of additional regularity requirements on the initial state f_0. In §1.3.4, we adapt a regularisation result from [67], which is then employed in §1.3.5 to obtain the sharp decay rate for solutions with finite initial entropy. The sharpness of this rate is shown by establishing special solutions in §1.3.6. As for the classical method, the Csiszár-Kullback inequality ([26], [46])

$$\|f_1 - f_2\|^2_{L^1(\mathbb{R}^d)} \le \frac{2}{\psi''(1)} e_\psi(f_1|f_2)$$

yields L^1-decay of solutions from decay in relative entropy (Theorem 1.27).

With the notations of §1.2.1 we introduce the relative entropy:

Definition 1.17. Let $0 \not\equiv \psi \in C(\mathbb{R}_0^+) \cap C^4(\mathbb{R}^+)$, $\psi(1) = \psi'(1) = 0$, $\psi'' \ge 0$ on $\mathbb{R}^+$, $(\psi''')^2 \le \frac{1}{2}\psi''\psi^{IV}$ on $\mathbb{R}^+$. Let $f \in L^1_+(\mathbb{R}^d)$ with $\int f \,\mathrm{d}x = 1$. Then

$$e_\psi(f|f_\infty) := \int\limits_{\mathbb{R}^d} \psi(\tfrac{f}{f_\infty}) f_\infty \,\mathrm{d}x$$

is called an *admissible relative entropy* with *generating function* ψ.

The entropy method is based on computing a bound on the first two time-derivatives of the relative entropy $e_\psi(f(t)) := e_\psi(f(t)|f_\infty)$ with f the solution to (2). Formally,

$$\tfrac{\mathrm{d}}{\mathrm{dt}} e_\psi(f(t)) = -\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \nabla(\frac{f}{f_\infty})^T D \nabla(\frac{f}{f_\infty}) f_\infty \,\mathrm{d}x =: -I_\psi(f) \le 0. \quad (1.18)$$

However, there may be a technical problem if $f(t,x) = 0$ (which can happen at the initial state f_0). For example, $\psi''(s) = \frac{1}{s}$ for the logarithmic entropy $\psi(s) = s\ln s - s + 1$, and this would lead to a division by zero. For this reason, we use a trick from [6] (see Remark 2.12) to rewrite (1.18):

Definition 1.18. Let ψ generate an admissible entropy, and let $f_0 \in L^1_+(\mathbb{R}^d)$ (or $f_0 \in L^1(\mathbb{R}^d)$ for quadratic ψ) with $\int_{\mathbb{R}^d} f_0 \, \mathrm{d}x = 1$. Define

$$w(x) := \int_1^{\frac{f_0}{f_\infty}(x)} \sqrt{\psi''(s)} \, \mathrm{d}s. \tag{1.19}$$

Then we call f_0 a *ψ-compatible initial state* iff $\nabla w \in L^2(\mathbb{R}^d, f_\infty)$.

With Definition 1.18, (1.18) can be written as

$$I_\psi(f) = \int_{\mathbb{R}^d} (\nabla w)^T D (\nabla w) f_\infty \, \mathrm{d}x. \tag{1.20}$$

Whenever $f \neq 0$, this is equivalent to (1.18). However, now there is no longer a problem when $f = 0$, since the integral in (1.19) is Hölder continuous at 0 with exponent $\frac{1}{2}$. The assumptions of Definition 1.18 clearly imply that a ψ-compatible initial state has finite entropy dissipation. It also has finite relative entropy, as we shall prove in Corollary 1.21 below.

Remark: The integral in Definition 1.18 can be calculated explicitly for the most common entropies:
For the quadratic entropy, $\psi(s) = \alpha(s-1)^2$ for some $\alpha > 0$, and thus

$$w = \sqrt{2\alpha}\left(\tfrac{f_0}{f_\infty} - 1\right). \tag{1.21}$$

For the logarithmic entropy, with

$$\psi(s) = \alpha(s+\beta)\ln\left(\tfrac{s+\beta}{1+\beta}\right) - \alpha(s-1) \tag{1.22}$$

for some $\alpha > 0$, $\beta \geq 0$, we have

$$w = 2\sqrt{\alpha}\left(\sqrt{\tfrac{f_0}{f_\infty} + \beta} - \sqrt{1+\beta}\right). \tag{1.23}$$

For the p-entropies, $1 < p < 2$, $\psi(s) = \alpha[(s+\beta)^p - (1+\beta)^p - p(1+\beta)^{p-1}(s-1)]$ for some $\alpha > 0$, $\beta \geq 0$, and thus

$$w = 2\sqrt{\tfrac{\alpha(p-1)}{p}}\left(\left(\tfrac{f_0}{f_\infty} + \beta\right)^{\frac{p}{2}} - (1+\beta)^{\frac{p}{2}}\right).$$

1.3.1 Modified entropy dissipation

There is another, in fact systematic problem with the entropy dissipation (1.18): Since D is singular for $k < d$, this functional is "lacking information" on some partial derivatives of $\frac{f}{f_\infty}$. But this information would be vital for the (standard) entropy method to work. More precisely, the functional I_ψ vanishes *not only* for $f = f_\infty$. As shown in Corollary 1.31, for any $t^* \geq 0$ there are initial conditions such that $I_\psi(f(t^*)) = 0$. Also, due to the monotonicity of $e_\psi(f(t))$, $I_\psi(f(t^*)) = 0$ for some $t^* \geq 0$ implies $I'_\psi(f(t^*)) = 0$. So, for degenerate Fokker-Planck equations, $e_\psi(f(t))$ is not a convex function of t – in contrast to the non-degenerate case from [6]. The possibility of having $I_\psi(f(t^*)) = I'_\psi(f(t^*)) = 0$ for $f(t^*) \neq f_\infty$ also shows that the standard entropy method cannot be carried over directly to the degenerate case in (2).
We therefore introduce the modified functional

$$S_\psi(f) := \int_{\mathbb{R}^d} (\nabla w)^T P (\nabla w) f_\infty \,\mathrm{d}x = \int_{\frac{f}{f_\infty}>0} \psi''(\frac{f}{f_\infty}) \nabla(\frac{f}{f_\infty})^T P \nabla(\frac{f}{f_\infty}) f_\infty \,\mathrm{d}x, \tag{1.24}$$

where we replace the matrix D in I_ψ with a regular, symmetric matrix P. P will be chosen in such a way that it provides an estimate between $\frac{\mathrm{d}}{\mathrm{dt}} S_\psi(f(t))$ and $S_\psi(f(t))$ for solutions f to (2), as shown later in this section. Moreover, since P is positive definite, there is a constant $c_P > 0$ with $P \geq c_P D$, and hence $S_\psi \geq c_P I_\psi$. The choice of P has to be done carefully: Simply choosing any positive definite matrix will retain information on all derivatives, but will in general not give a decay estimate (see §1.5, Example 1).

Remark: Introducing the functional S_ψ differs from the modified entropy dissipation approach in [27]. There one considers an "intermediate functional" $K(f)$, which measures the distance of f to the set of stationary states of the symmetric part (L_s in our case). It is constructed in such a way that, whenever K becomes small, but the relative entropy e does not, there is a mechanism that increases K again. One then aims for an inequality like

$$\frac{\mathrm{d}}{\mathrm{dt}} e(f(t)|f_\infty) \leq -K(f(t)).$$

While the right-hand side is still zero for some $f \neq f_\infty$, due to the construction of K this can no longer happen along trajectories $f(t)$.

Choosing the matrix P is the crucial ingredient for the definition of our modified entropy dissipation S_ψ:

Lemma 1.19. *Let $Q := KC^T K^{-1}$. Let $\mu := \min\{\Re\{\lambda\} | \lambda$ is an eigenvalue of $C\}$. Due to condition (A), $\mu > 0$. Let $\{\lambda_m | 1 \leq m \leq m_0\}$ be all the eigenvalues of C with $\mu = \Re\{\lambda_m\}$, only counting their geometric multiplicity.*

(i) If λ_m is non-defective[2] for all $m \in \{1, \ldots, m_0\}$, then there exists a symmetric, positive definite matrix $P \in \mathbb{R}^{d\times d}$ with

$$QP + PQ^T \geq 2\mu P. \tag{1.25}$$

(ii) If λ_m is defective for at least one $m \in \{1, \ldots, m_0\}$, then for any $\varepsilon > 0$ there exists a symmetric, positive definite matrix $P = P(\varepsilon) \in \mathbb{R}^{d\times d}$ with

$$QP + PQ^T \geq 2(\mu - \varepsilon)P. \tag{1.26}$$

(iii) For any such matrix P, and for any ψ-compatible function f_0, $S_\psi(f_0) < \infty$.

Proof: The idea behind the construction of P is the following: If Q is not defective and $w_1, \ldots, w_d$ are its eigenvectors, then one can choose P as the weighted sum

$$P := \sum_{j=1}^{d} b_j \, w_j \otimes \overline{w_j} \,, \tag{1.27}$$

with $b_j \in \mathbb{R}^+$, $j = 1, \ldots, d$. As $\{w_j\}_{j=1,\ldots,d}$ is a basis of $\mathbb{C}^d$, P is positive definite. If any w_j is complex, its complex conjugate $\overline{w_j}$ is also an eigenvector of Q, since Q is real. By taking the same coefficient b_j for both, we obtain a real matrix P since

$$w_j \otimes \overline{w_j} + \overline{w_j} \otimes w_j$$

is real. For P from (1.27), we obtain

$$\begin{aligned} QP + PQ^T &= \sum_{j=1}^{d} b_j \big(Qw_j \otimes \overline{w_j} + w_j \otimes \overline{w_j} Q^T\big) \\ &= \sum_{j=1}^{d} b_j (\lambda_j + \overline{\lambda_j}) w_j \otimes \overline{w_j} = \sum_{j=1}^{d} 2\Re\{\lambda_j\} b_j \, w_j \otimes \overline{w_j} \\ &\geq 2\mu \sum_{j=1}^{d} b_j \, w_j \otimes \overline{w_j} = 2\mu P \,. \end{aligned}$$

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

This also implies that (1.25) is an equality iff all eigenvalues of Q are non-defective and have real part μ.
If at least one of the eigenvalues of Q is defective, then there is no basis of $\mathbb{C}^d$ consisting of eigenvectors of Q, and the sum in (1.27) will have less summands and not be regular. One can still construct P in a similar fashion using the basis of $\mathbb{C}^d$ consisting of generalised eigenvectors of Q, but the computations are no longer as straightforward. To this end, we consider the Jordan normal form J of Q^T, given by the similarity transformation $AQ^TA^{-1} = J$ with some $A \in \mathbb{C}^{d\times d}$. Let J have N Jordan blocks, each of length l_n; $n = 1, \ldots, N$.
(i) By assumption, all Jordan blocks corresponding to eigenvalues with $\Re\{\lambda_n\} = \mu$ are trivial, i.e. of length 1. Corresponding to the structure of J, we define

$$B_n := \operatorname{diag}(b_n^{l_n}, \ldots, b_n^1)\,, \quad n = 1, \ldots, N,$$

and the positive diagonal matrix

$$B := \operatorname{diag}(B_1, \ldots, B_N),$$

where $c_1 := 1$, $c_j := 1 + (c_{j-1})^2$; $j = 2, \ldots, l_n$,

$$b_n^j := c_j\,(\tau_n)^{2(1-j)}\,, \quad j = 1, \ldots, l_n\,, \tag{1.28}$$

and $\tau_n := 2(\Re\{\lambda_n\} - \mu) \geq 0$ for $n = 1, \ldots, N$. This yields for the n-th Jordan block J_n in the case $l_n = 1$: $B_n = I$ and

$$J_n^H B_n + B_n J_n = (\overline{\lambda_n} + \lambda_n) B_n \geq 2\mu B_n.$$

Here, J_n^H denotes the Hermitian adjoint of J_n. In the case $l_n > 1$, we have $\tau_n > 0$ and

$$\begin{aligned}
&J_n^H B_n + B_n J_n - 2\mu B_n \\
&= \begin{pmatrix} 2(\Re\{\lambda_n\} - \mu) b_n^{l_n} & b_n^{l_n - 1} & & \\ b_n^{l_n-1} & 2\Re(\{\lambda_n\} - \mu) b_n^{l_n-1} & \ddots & \\ & \ddots & \ddots & b_n^1 \\ & & b_n^1 & 2(\Re\{\lambda_n\} - \mu) b_n^1 \end{pmatrix} \geq 0.
\end{aligned}$$

The last inequality follows from

$$M_m := \begin{pmatrix} \tau^{3-2m}c_m & \tau^{4-2m}c_{m-1} & & \\ \tau^{4-2m}c_{m-1} & \tau^{5-2m}c_{m-1} & \ddots & \\ & \ddots & \ddots & c_1 \\ & & c_1 & \tau c_1 \end{pmatrix} \geq 0, \quad m = 1, \ldots, \max_n(l_n),$$

for any $\tau > 0$, which can be verified by induction over $m := l_n$ using the principal minor test:
<u>$m = 2$</u>:

$$M_2 = \begin{pmatrix} 2\tau_n^{-1} & 1 \\ 1 & \tau_n \end{pmatrix} > 0.$$

<u>$m > 2$</u>: Let $M_n \geq 0$ hold for $l_n \leq m$ with $\det M_n = \tau_n^{1-(l_n-1)^2}$. Then

$$M_{m+1} = \begin{pmatrix} \tau_n b_1^n & b_2^n \\ b_2^n & M_m \end{pmatrix}$$

and we compute, using Laplace's formula for the first column

$$\begin{aligned} \det M_{m+1} &= \tau_n b_1^n \det M_m - (b_2^n)^2 \det M_{m-1} \\ &= \tau_n^{1-2m}[1 + (c_2^n)^2]\tau_n^{1-(m-1)^2} - (c_2^n)^2 \tau_n^{4-4m} \tau_n^{1-(m-2)^2} \\ &= \tau_n^{1-m^2} > 0. \end{aligned}$$

In total, we have $J^H B + BJ \geq 2\mu B$, and hence

$$(A^{-1})^H Q A^H B + BAQ^T A^{-1} \geq 2\mu B,$$

which implies

$$QA^H BA + A^H BAQ^T \geq 2\mu A^H BA.$$

The claim then follows with $P := A^H BA$.
(ii) In this case, there exists a non-trivial Jordan block $J_{\tilde{n}}$ corresponding to an eigenvalue with $\Re\{\lambda_{\tilde{n}}\} = \mu$. In the above computation, we choose $\tau_{\tilde{n}} := 2(\Re\{\lambda_{\tilde{n}}\} - \mu + \varepsilon) > 0$ for some $\varepsilon > 0$. Hence, $J_{\tilde{n}}^H B_{\tilde{n}} + B_{\tilde{n}} J_{\tilde{n}} \geq 2(\mu - \varepsilon) B_{\tilde{n}}$ and the result follows. However, in this case P depends on ε.
(iii) This follows just like for (1.24). □

Remarks:

(i) In general, the matrix P in Lemma 1.19 is not uniquely determined.

(ii) From (1.28), we see that for a defective eigenvalue $\lambda_{\tilde{n}}$ (i.e. $P = P(\varepsilon)$),

$$\forall 1 < j \leq l_{\tilde{n}} : \quad \lim_{\varepsilon \to 0} b_{\tilde{n}}^j = \infty$$

and thus

$$\lim_{\varepsilon \to 0} S_\psi(f_0, \varepsilon) = \infty,$$

with $S_\psi(f_0, \varepsilon) := \int\limits_{\mathbb{R}^d} (\nabla w)^T P(\varepsilon) \nabla w f_\infty \, \mathrm{d}x$.

(iii) (1.25) can be rewritten as $(Q - \mu)P + P(Q^T - \mu) \geq 0$, which bears a close resemblance to the continuous Lyapunov equation from Theorem 1.12. If we assume equality in (1.25) and if $Q - \mu$ were positively stable, then there would be a unique solution $P = 0$, see e.g. [44]. But since μ is the real part of an eigenvalue of Q, $Q - \mu$ is not positively stable. This explains why we can find a non-trivial solution of (1.25) at the price of uniqueness. There is equality in (1.25) iff all eigenvalues of Q have the same real part μ and are non-degenerate. For additional details, we refer to [44], [61].

We shall later make use of convex Sobolev inequalities related to the measure $\mu = f_\infty \, \mathrm{d}x$:

Lemma 1.20. *Let $f \in L^1_+$ ($f \in L_1$ for quadratic ψ). Then*

$$e_\psi(f|f_\infty) \leq \frac{1}{2\lambda_P} S_\psi(f), \tag{1.29}$$

where both sides may be infinite and λ_P is the largest constant such that

$$K^{-1} \geq \lambda_P P^{-1}$$

holds.

Proof: Consider the Fokker-Planck operator

$$L_P f := \mathrm{div}(P \nabla (\frac{f}{f_\infty}) f_\infty)$$

on L^2. Then L_P is symmetric due to the symmetry of P, and f_∞ spans the kernel of L_P. One easily checks that

$$\tfrac{\mathrm{d}}{\mathrm{dt}} e_\psi(\tilde{f}(t)|f_\infty) = -S_\psi(\tilde{f}(t))$$

for a solution $\tilde{f}(t)$ to $\tilde{f}_t = L_P\tilde{f}$. As shown in Corollary 2.17, [6], this symmetric, non-degenerate Fokker-Planck equation leads to an exponential decay of the relative entropy, and in parallel to a convex Sobolev inequality: Using the notation $f_\infty(x) := c_K e^{-V(x)}$, $V(x) := \frac{x^T K^{-1} x}{2}$, we have the Bakry-Émery condition

$$\frac{\partial^2 V}{\partial x^2} = K^{-1} \geq \lambda_P P^{-1},$$

where the constant $\lambda_P > 0$ is chosen as large as possible. Hence, all $g \in L^1_+(\mathbb{R}^d)$ with $\int\limits_{\mathbb{R}^d} g \,\mathrm{d}x = 1$ satisfy the convex Sobolev inequality

$$e_\psi(g|f_\infty) \leq \frac{1}{2\lambda_P} S_\psi(g),$$

where both sides may be infinite. This completes the proof. □

Lemma 1.20 implies that any ψ-compatible f also has finite relative entropy generated by ψ:

Corollary 1.21. *Let f be ψ-compatible. Then it holds that*

$$e_\psi(f|f_\infty) < \infty.$$

Proof: Since f is ψ-compatible, we have $S_\psi(f) < \infty$, $f \in L^1_+(\mathbb{R}^d)$ and $\int\limits_{\mathbb{R}^d} f \,\mathrm{d}x = 1$. Applying (1.29) completes the proof. □

1.3.2 Fisher information matrix

While it is perfectly possible to derive a decay estimate for S and directly show its decay, we take a detour at this point and introduce the Fisher matrix. For a direct estimate on S, we refer to [33].

Definition 1.22. For any ψ generating an admissible entropy (see Definition 1.17), the *Fisher matrix* Σ_ψ of f with respect to f_∞ is defined as

$$\Sigma_\psi := \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u \otimes u f_\infty \,\mathrm{d}x,$$

where $u := \nabla \frac{f}{f_\infty}$.

The Fisher matrix (sometimes also Fisher information matrix or Fisher information, though the latter also applies to the scalar version) is for example used to predict the precision of measurements (e.g. particle production rate in some physical system) taking into account covariance (the non-diagonal entries) of observables. For a brief introduction from a physics standpoint, we refer to [68] and references therein.

For the entropy method, we are interested in the time-dependent Fisher matrix $\Sigma_\psi(t)$ along a solution trajectory $f(\cdot)$. We compute

$$\begin{aligned}
\mathrm{Tr}(D\Sigma_\psi(t)) &= \int_{\mathbb{R}^d} \psi''(\nabla\frac{f(t)}{f_\infty})\mathrm{Tr}(D(u(t)\otimes u(t)))f_\infty \,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} \psi''(\nabla\frac{f(t)}{f_\infty}) \sum_{j,l=1}^{d} D_{jl}u_l(t)u_j(t)f_\infty \,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} \psi''(\nabla\frac{f(t)}{f_\infty})(\nabla\frac{f(t)}{f_\infty})^T D\nabla\frac{f(t)}{f_\infty} f_\infty \,\mathrm{d}x = I_\psi(f(t)),
\end{aligned}$$

so the entropy dissipation $I_\psi(f(t))$ can be recovered from $\Sigma_\psi(t)$. But $\Sigma_\psi(t)$ contains more information than the entropy dissipation. Since $u(t)\otimes u(t)$ is a positive semidefinite matrix, it follows that $\Sigma_\psi(t) \geq 0$, and hence

$$S_\psi(t) = \mathrm{Tr}(P\Sigma_\psi) \geq 0 \tag{1.30}$$

for any $P \geq 0$. Here, we have used that the trace of a product of two positive semidefinite matrices is non-negative. In Lemma 1.23, we compute a bound on $\frac{\mathrm{d}}{\mathrm{dt}}\Sigma_\psi$. Since

$$\tfrac{\mathrm{d}}{\mathrm{dt}}\mathrm{Tr}(P\Sigma_\psi) = \mathrm{Tr}(P\tfrac{\mathrm{d}}{\mathrm{dt}}\Sigma_\psi),$$

this has two applications: First, if one can prove the estimate $\frac{\mathrm{d}}{\mathrm{dt}}\Sigma_\psi(t) \leq -2\mu\Sigma_\psi(t)$, decay of the Fisher matrix with rate 2μ follows. The Fisher matrix has not yet been studied in the context of Fokker-Planck equations, and conditions for its decay - even though they are very restrictive, see Theorem 1.24 - are new. Second, even if there is no decay estimate on Σ_ψ itself, one can try to find a matrix P such that $\mathrm{Tr}(P\frac{\mathrm{d}}{\mathrm{dt}}\Sigma_\psi) \leq -2\mu\mathrm{Tr}(P\Sigma_\psi)$ holds. In fact, we will prove in §1.3.3 that the matrix P from Lemma 1.19 has this property.

Lemma 1.23. *Let f_0 be a ψ-compatible initial state, and let $f(t)$ be the corresponding solution to (2). Let $u := \nabla\frac{f}{f_\infty}$. Then the estimate*

$$\tfrac{d}{dt}\Sigma_\psi(t) \leq -\int_{\mathbb{R}^d} \psi''(\nabla\frac{f(t)}{f_\infty})\left(Q^T\left[u(t)\otimes u(t)\right] + \left[u(t)\otimes u(t)\right]Q\right) f_\infty \,\mathrm{d}x$$

holds for all $t > 0$, where

$$Q := (D - R)\frac{\partial^2 V}{\partial x^2}.$$

Proof: By Corollary 1.6, the solutions of (2) have the required regularity to perform all the computations in this proof for $t > 0$. We compute, leaving out the argument t for ease of reading,

$$u_t = -\frac{\partial^2 V}{\partial x^2}(D + R)u - \frac{\partial u}{\partial x}(D - R)\nabla V + (\nabla^T D \nabla)u.$$

Hence,

$$\tfrac{\mathrm{d}}{\mathrm{dt}}\Sigma_\psi = \underbrace{\int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) f_t u \otimes u \;\mathrm{d}x}_{=:(I)} \underbrace{\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})\left(u_t \otimes u + u \otimes u_t\right) f_\infty \;\mathrm{d}x}_{=:(II)}.$$

We introduce the short-hand notations $f_{,j} := \frac{\partial f}{\partial x_j}$ and $f_{,js} := \frac{\partial^2 f}{\partial x_j \partial x_s}$. For the (r,s)-entry of (I), $1 \le r, s \le d$, we obtain

$$\int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) \operatorname{div}(f_\infty (D+R)u) u_r u_s \;\mathrm{d}x = \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_r u_s (f_\infty (D+R)_{lj} u_j)_{,l} \;\mathrm{d}x$$
$$= -\int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_r u_s (D+R)_{lj} u_j V_{,l} f_\infty \;\mathrm{d}x + \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_r u_s (D+R)_{lj} u_{j,l} f_\infty \;\mathrm{d}x.$$

Examining the last term, we compute, using the antisymmetry of R and the symmetry of $\frac{\partial u}{\partial x}$,

$$\begin{aligned}
&\int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_r u_s (D+R)_{lj} u_{j,l} f_\infty \;\mathrm{d}x \\
&= \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_r u_s D_{lj} u_{j,l} f_\infty \;\mathrm{d}x \\
&= -\int\limits_{\mathbb{R}^d} u_j D_{lj} \left(\psi'''(\frac{f}{f_\infty}) u_r u_s f_\infty\right)_{,l} \;\mathrm{d}x \\
&= -\int\limits_{\mathbb{R}^d} \psi^{IV} u_j D_{lj} u_l u_r u_s f_\infty \;\mathrm{d}x - \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_j D_{lj} u_{r,l} u_s f_\infty \;\mathrm{d}x \\
&\quad - \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_j D_{lj} u_r u_{s,l} f_\infty \;\mathrm{d}x + \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_j D_{lj} u_r u_s V_{,l} f_\infty \;\mathrm{d}x.
\end{aligned}$$

Hence,

$$
\begin{aligned}
(I) = &-\int_{\mathbb{R}^d} \psi^{IV}(\frac{f}{f_\infty}) \left[u^T D u\right] [u \otimes u] f_\infty \,\mathrm{d}x \\
&-\int_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) \left[\left(\frac{\partial u}{\partial x} D u\right) \otimes u + u \otimes \left(\frac{\partial u}{\partial x} D u\right)\right] f_\infty \,\mathrm{d}x \\
&-\int_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) \left[(\nabla V)^T R u\right] [u \otimes u] f_\infty \,\mathrm{d}x. \qquad (1.31)
\end{aligned}
$$

Next, we compute

$$
\begin{aligned}
&\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_{r,t} u_s f_\infty \,\mathrm{d}x \\
&= -\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) V_{,rl} (D+R)_{lj} u_j u_s f_\infty \,\mathrm{d}x - \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_{r,l} (D+R)_{jl} V_{,j} u_s f_\infty \,\mathrm{d}x \\
&+ \underbrace{\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) D_{lj} u_{r,lj} u_s f_\infty \,\mathrm{d}x}_{=:(III)}\,. \qquad (1.32)
\end{aligned}
$$

Integrating (III) by parts yields

$$
\begin{aligned}
(III) &= -\int_{\mathbb{R}^d} D_{lj} u_{r,l} (\psi''(\frac{f}{f_\infty}) u_s f_\infty)_{,j} \,\mathrm{d}x \\
&= -\int_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) D_{lj} u_{r,l} u_j u_s f_\infty \,\mathrm{d}x - \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) D_{lj} u_{s,j} u_{r,l} f_\infty \,\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) D_{lj} u_{r,l} u_s V_{,j} f_\infty \,\mathrm{d}x
\end{aligned}
$$

and thus, reinserting into (1.32)

$$
\begin{aligned}
&\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_{r,t} u_s f_\infty \,\mathrm{d}x \\
&= -\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) V_{,rl} (D+R)_{lj} u_j u_s f_\infty \,\mathrm{d}x \qquad (1.33) \\
&\quad -\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_{r,l} \left((D+R) - D\right)_{jl} V_{,j} u_s f_\infty \,\mathrm{d}x \\
&\quad -\int_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) D_{lj} u_{r,l} u_j u_s f_\infty \,\mathrm{d}x - \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) D_{lj} u_{s,j} u_{r,l} f_\infty \,\mathrm{d}x.
\end{aligned}
$$

Again using $R = -R^T$, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_t \otimes u f_\infty \,\mathrm{d}x \\
&= -\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \left[\left(\frac{\partial^2 V}{\partial x^2}(D+R)u \right) \otimes u \right] f_\infty \,\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) \left[\left(\frac{\partial u}{\partial x} R \nabla V \right) \otimes u \right] f_\infty \,\mathrm{d}x \\
&\quad - \int_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) \left[\left(\frac{\partial u}{\partial x} D u \right) \otimes u \right] f_\infty \,\mathrm{d}x - \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \frac{\partial u}{\partial x} D \frac{\partial u}{\partial x} f_\infty \,\mathrm{d}x
\end{aligned} \tag{1.34}$$

Next, we consider

$$\begin{aligned}
&- \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_{r,l} R_{jl} V_{,j} u_s f_\infty \,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} u_r R_{jl} \left(\psi''(\frac{f}{f_\infty}) u_s V_{,j} f_\infty \right)_{,l} \mathrm{d}x \\
&= \int_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_r u_l R_{jl} V_{,j} u_s f_\infty \,\mathrm{d}x + \int_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_r R_{jl} u_{s,l} V_{,j} f_\infty \,\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_r R_{jl} V_{,lj} u_s f_\infty \,\mathrm{d}x - \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_r u_s R_{jl} V_{,l} V_{,j} f_\infty \,\mathrm{d}x.
\end{aligned}$$

Since R is antisymmetric, the sums $R_{jl}V_{,lj}$ and $R_{jl}V_{,l}V_{,j}$ are 0, and we obtain for (1.34):

$$\begin{aligned}
&\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_t \otimes u f_\infty \,\mathrm{d}x \\
&= -\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \left[\left(\frac{\partial^2 V}{\partial x^2}(D+R)u \right) \otimes u \right] f_\infty \,\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) \left[(\nabla V)^T R u \right] \left[u \otimes u \right] f_\infty \,\mathrm{d}x \\
&\quad - \int_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) \left[\left(\frac{\partial u}{\partial x} D u \right) \otimes u \right] f_\infty \,\mathrm{d}x - \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \frac{\partial u}{\partial x} D \frac{\partial u}{\partial x} f_\infty \,\mathrm{d}x \qquad (1.35) \\
&\quad - \int_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) \left[u \otimes \left(\frac{\partial u}{\partial x} R \nabla V \right) \right] f_\infty \,\mathrm{d}x.
\end{aligned}$$

The computations for the term with $u \otimes u_t$ are the same; combining (1.34) and

(1.35) for both yields

$$
\begin{aligned}
(II) &= \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \left[u_t \otimes u + u \otimes u_t\right] f_\infty \,\mathrm{d}x \\
&= -\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \left[\left(\frac{\partial^2 V}{\partial x^2}(D+R)u\right) \otimes u + u \otimes \left(\frac{\partial^2 V}{\partial x^2}(D+R)u\right)\right] f_\infty \,\mathrm{d}x \\
&\quad - \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) \left[\left(\frac{\partial u}{\partial x} D u\right) \otimes u + u \otimes \left(\frac{\partial u}{\partial x} D u\right)\right] f_\infty \,\mathrm{d}x \\
&\quad - 2\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \frac{\partial u}{\partial x} D \frac{\partial u}{\partial x} f_\infty \,\mathrm{d}x + \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) \left[(\nabla V)^T R u\right] \left[u \otimes u\right] f_\infty \,\mathrm{d}x.
\end{aligned}
\tag{1.36}
$$

Finally, we can add (1.31) and (1.36) and obtain

$$
\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{dt}}\Sigma_\psi &= -\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \left[\left(\frac{\partial^2 V}{\partial x^2}(D+R)u\right) \otimes u + u \otimes \left(\frac{\partial^2 V}{\partial x^2}(D+R)u\right)\right] f_\infty \,\mathrm{d}x \\
&\quad - 2\int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) \left[\left(\frac{\partial u}{\partial x} D u\right) \otimes u + u \otimes \left(\frac{\partial u}{\partial x} D u\right)\right] f_\infty \,\mathrm{d}x \\
&\quad - 2\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \frac{\partial u}{\partial x} D \frac{\partial u}{\partial x} f_\infty \,\mathrm{d}x - \int\limits_{\mathbb{R}^d} \psi^{IV}(\frac{f}{f_\infty}) \left[u^T D u\right] \left[u \otimes u\right] f_\infty \,\mathrm{d}x.
\end{aligned}
\tag{1.37}
$$

Now, let

$$
\begin{aligned}
\Lambda &:= \begin{pmatrix} 2\psi''(\frac{f}{f_\infty})\,\mathrm{Id} & 2\psi'''(\frac{f}{f_\infty})\,\mathrm{Id} \\ 2\psi'''(\frac{f}{f_\infty})\,\mathrm{Id} & \psi^{IV}(\frac{f}{f_\infty})\,\mathrm{Id} \end{pmatrix} \in \mathbb{R}^{2d\times 2d}, \\
U &:= \begin{pmatrix} D\frac{\partial u}{\partial x} \\ D\left[u \otimes u\right] \end{pmatrix} \in \mathbb{R}^{d\times 2d}.
\end{aligned}
$$

Then we obtain, using $D = D^2$ and the symmetry of $u \otimes u$,

$$
\begin{aligned}
U^T \Lambda U &= 2\psi''(\frac{f}{f_\infty}) \frac{\partial u}{\partial x} D \frac{\partial u}{\partial x} + 2\psi'''(\frac{f}{f_\infty}) \left[u \otimes u\right] \left[D\frac{\partial u}{\partial x}\right] \\
&\quad + 2\psi'''(\frac{f}{f_\infty}) \left[\frac{\partial u}{\partial x} D\right] \left[u \otimes u\right] + \psi^{IV} \left[u \otimes u\right] D \left[u \otimes u\right].
\end{aligned}
$$

We compute

$$
\begin{aligned}
\left[u \otimes u\right] D \left[u \otimes u\right] &= \left[u^T D u\right] \left[u \otimes u\right], \\
\left[\frac{\partial u}{\partial x} D\right] \left[u \otimes u\right] &= \left(\frac{\partial u}{\partial x} D u\right) \otimes u,
\end{aligned}
$$

and hence (1.37) implies:

$$\tfrac{\mathrm{d}}{\mathrm{dt}}\Sigma_\psi = -\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \left[\left(\frac{\partial^2 V}{\partial x^2}(D+R)u\right) \otimes u + u \otimes \left(\frac{\partial^2 V}{\partial x^2}(D+R)u\right)\right] f_\infty \,\mathrm{d}x$$
$$- \int\limits_{\mathbb{R}^d} U^T \Lambda U \,\mathrm{d}x.$$

Since $\psi'' \geq 0$, $\psi''\psi^{IV} - 2(\psi''')^2 \geq 0$, we obtain $\Lambda \geq 0$, and hence

$$\tfrac{\mathrm{d}}{\mathrm{dt}}\Sigma_\psi \leq -\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \left[\left(\frac{\partial^2 V}{\partial x^2}(D+R)u\right) \otimes u + u \otimes \left(\frac{\partial^2 V}{\partial x^2}(D+R)u\right)\right] f_\infty \,\mathrm{d}x \tag{1.38}$$
$$= -\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) \left(\left[\frac{\partial^2 V}{\partial x^2}(D+R)\right][u \otimes u] + [u \otimes u]\left[(D-R)\frac{\partial^2 V}{\partial x^2}\right]\right) f_\infty \,\mathrm{d}x.$$

□

Lemma 1.23 motivates to define a Bakry-Émery-type condition that leads to exponential decay of the Fisher information matrix.

Theorem 1.24. *The following two conditions are equivalent:*

(i) There exists $\mu > 0$ such that for all ψ-compatible f_0 and corresponding solutions $f(t)$ to (2),

$$\tfrac{d}{dt}\Sigma(t) \leq -2\mu\Sigma(t) \tag{1.39}$$

holds.

(ii) There exists $\mu > 0$ such that

$$\frac{\partial^2 V}{\partial x^2}(D+R) = \mu \,\mathrm{Id}\,.$$

Proof: (ii)⇒(i) follows immediately from Lemma 1.23.
To see (i)⇒(ii), we first consider the case $d = 2$. From Lemma 1.23, we have

$$\tfrac{\mathrm{d}}{\mathrm{dt}}\Sigma_\psi(t) \leq -\int\limits_{\mathbb{R}^d} \psi''(\frac{f(t)}{f_\infty}) \left(Q^T [u(t) \otimes u(t)] + [u(t) \otimes u(t)]\, Q\right) f_\infty \,\mathrm{d}x.$$

If we want to bound this above by a negative multiple of Σ_ψ for any possible solution trajectory, we need to investigate the estimate

$$Q^T(u \otimes u) + (u \otimes u)Q \geq \lambda(u \otimes u)$$

for some $Q \in \mathbb{R}^{2\times 2}$ and *all* $u \in \mathbb{R}^2$. Q does not need to be symmetric; in fact, in the hypocoercive case it is not. So let

$$Q := \begin{pmatrix} r & t_2 \\ t_1 & s \end{pmatrix}, \quad u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

Then

$$Q(u \otimes u) + (u \otimes u)Q^T = \begin{pmatrix} 2ra^2 + 2t_1 ab & (r+s)ab + t_1 b^2 + t_2 a^2 \\ (r+s)ab + t_1 b^2 + t_2 a^2 & 2sb^2 + 2t_2 ab \end{pmatrix}.$$

Note that the right hand side of (1.39) is negative semidefinite. Thus, to have any chance of obtaining (1.39), we need that $Q(u \otimes u) + (u \otimes u)Q^T$ is positive (semi-)definite. A necessary (but not sufficient) condition for that is $\det(Q(u \otimes u) + (u \otimes u)Q^T) \geq 0$ for all u, i.e. for all $a, b \in \mathbb{R}$. We compute

$$\begin{aligned} &\det\left(Q(u \otimes u) + (u \otimes u)Q^T\right) \\ &= 4\left[(ra^2 + t_1 ab)(sb^2 + t_2 ab)\right] - \left((r+s)ab + t_2 a^2 + t_1 b^2\right)^2 \\ &= 4\left[rsa^2b^2 + t_1 t_2 a^2 b^2 + (rt_2 a^2 + st_1 b^2)ab\right] \\ &\quad - \left[(r+s)^2 a^2 b^2 + (t_2 a^2 + t_1 b^2)^2 + 2(r+s)ab(t_2 a^2 + t_1 b^2)\right] \\ &= -(r-s)^2 a^2 b^2 - (t_2 a^2 - t_1 b^2)^2 + 2\left[(r-s)t_2 a^2 + (s-r)t_1 b^2\right] ab. \end{aligned}$$

If $a = 0$ or $b = 0$, we immediately see that $t_1 = t_2 = 0$ is necessary, or the determinant is negative. But if $t_1 = t_2 = 0$, we see from the case $a = b$ that $r = s$ is also necessary. So we need that $Q = \mu\,\mathrm{Id}$ for some $\mu \in \mathbb{R}$. From (i), it then follows that $\mu > 0$. This concludes the second implication for $d = 2$.
The case $d > 2$ directly follows from the case $d = 2$: for $1 \leq \alpha < \beta \leq d$, choose $u_\alpha = a, u_\beta = b$ and $u_j = 0$ for $\alpha \neq j \neq \beta$. Then one gets the same expression as above with $r = Q_{\alpha\alpha}$, $s = Q_{\beta\beta}$, $t_1 = Q_{\alpha\beta}$ and $t_2 = Q_{\beta\alpha}$. Repeating this for all choices of α, β yields (i)$\Rightarrow$(ii).
As an interesting side note, if Q is not a multiple of Id, then $Q(u\otimes u)+(u\otimes u)Q^T$ is indefinite for most u. If, for example, Q is diagonal ($t_1 = t_2 = 0$), then if $Q(u\otimes u)+(u\otimes u)Q^T$ is (semi-)definite for any $u \neq 0$, it follows that either $r = s$ or the determinant of $Q(u\otimes u)+(u\otimes u)Q^T$ is negative, which is a contradiction to definiteness. $\square$

Under our assumption of linear drift coefficients, $\frac{\partial^2 V}{\partial x^2}(D+R) = K^{-1}CK = \mu\,\mathrm{Id}$ would imply $C = \mu\,\mathrm{Id}$, which would be a contradiction to condition (A). So Theorem 1.24 is not applicable in the case $k < d$. In fact, a decay estimate for Σ_ψ implies a decay estimate for I_ψ, which as demonstrated in the introduction can not hold under (2), so this result should not be surprising. From (1.30), it even follows that a decay estimate for Σ_ψ would imply a decay estimate for the norm of any directional derivative $v \cdot \nabla$, $v \in \mathbb{R}^d$. So it is not surprising that such an estimate only holds in the most simple case.

1.3.3 Decay of the modified entropy dissipation

We now return to the modified entropy dissipation S_ψ. Since S_ψ can be recovered from the Fisher matrix, see (1.30), it is straightforward to utilize Lemma 1.23 for a decay estimate on the modified entropy production S_ψ (Lemma 1.19):

Proposition 1.25. *Assume condition (A). Let ψ generate an admissible entropy and let f be the solution to (2) with a ψ-compatible initial state f_0, $\mu := \min\{\Re\{\lambda\}|\lambda$ is an eigenvalue of $C\}$. Let P, $S_\psi(f_0)$ be defined as in Lemma 1.19, $\{\lambda_m | 1 \le m \le m_0\}$ be the eigenvalues of C with $\mu = \Re\{\lambda_m\}$.*

(i) If all λ_m, $1 \le m \le m_0$, are non-defective, then

$$S_\psi(f(t)) \le S_\psi(f_0)e^{-2\mu t}, \quad t \ge 0.$$

(ii) If λ_m is defective for at least one $m \in \{1, \dots, m_0\}$, then

$$S_\psi(f(t)) \le S_\psi(f_0, \varepsilon)e^{-2(\mu-\varepsilon)t}, \quad t \ge 0,$$

for any $\varepsilon \in (0, \mu)$.

Proof: For P from Definition 1.19, we compute

$$\mathrm{Tr}(\Sigma_\psi(t)P) = \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})u^T P u f_\infty \,\mathrm{d}x = S_\psi(t).$$

With Lemma 1.23, this yields

$$\tfrac{\mathrm{d}}{\mathrm{dt}}S_\psi(t) = \mathrm{Tr}(P\tfrac{\mathrm{d}}{\mathrm{dt}}\Sigma_\psi(t)) \qquad \le -\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})u^T(QP + PQ^T)u f_\infty \,\mathrm{d}x,$$

where $Q = KC^TK^{-1}$. By the definition of P (see (1.25)), we obtain

$$\tfrac{\mathrm{d}}{\mathrm{dt}}S_\psi(t) \le -2\eta S_\psi(t),$$

where $\eta = \mu$ in case (i), and $\eta = \mu - \varepsilon$, $\varepsilon \in (0, \mu)$ in case (ii). Applying Gronwall's Lemma completes the proof. $\square$

Remark: This result holds for all matrices P chosen according to Lemma 1.19. Clearly, the rate μ is independent of the choice of P.
Using (1.29), this result already implies exponential decay of the relative entropy – but under the (too strong) assumption that $S_\psi(f_0) < \infty$. This will be improved in Theorem 1.27 below.

In the standard entropy method for fully parabolic equations, one derives decay of the relative entropy from the decay of the entropy dissipation by integrating the inequality $\frac{\mathrm{d}^2}{\mathrm{d}s^2} e_\psi(f(s)) \geq -\frac{\mathrm{d}}{\mathrm{d}s} e_\psi(f(s))$ over (t, ∞). This requires a-priori knowlegde that $e_\psi(f(t = \infty)) = 0$, which, as shown in [3], can be derived from the decay of S (which is the entropy dissipation functional for fully parabolic equations). However, since the inequality $\frac{\mathrm{d}^2}{\mathrm{d}s^2} e_\psi(f(s)) \geq -\frac{\mathrm{d}}{\mathrm{d}s} e_\psi(f(s))$ is in general wrong for the degenerate case, this approach won't work.

1.3.4 Regularisation in relative entropy

We will now prove a regularisation result that allows us to extend the result of Proposition 1.25 to initial states with (only) finite relative entropy. The fundamental concept is that hypoelliptic operators regularise, though at a slower rate than fully elliptic ones. Local estimates of this sort first appeared in the proof by Hörmander [42] as well as in [45], [59]. Our result generalises Theorems A.12, A.15 in [67] (expressed for quadratic and logarithmic entropies) to all admissible ψ-entropies. Those results, in turn, used an idea developed by Hérau [40]. The regularisation depends on the order τ of the finite rank Hörmander condition for L (cf. Remark 1.4).

Theorem 1.26. *Let condition (A) hold, $f_0 \in L^1_+(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} f_0 \,\mathrm{d}x = 1$ and $e_\psi(f_0|f_\infty) < \infty$. Let $f(t)$ be the solution of (2) with initial condition f_0, and let τ be the minimal constant such that Lemma 1.3 holds. Then there is a positive constant $c_r > 0$ such that*

$$\forall t \in (0, 1]: \; S_\psi(f(t)) \leq c_r t^{-(2\tau+1)} e_\psi(f_0|f_\infty). \tag{1.40}$$

Proof: The idea of the proof is to construct a decaying-in-time functional $\mathcal{F}$ that is a (positive) linear combination of both sides of (1.40) – multiplied by $t^{2\tau+1}$.

Step 1 (construction of $\mathcal{F}$): With $Q := KC^T K^{-1}$ from Lemma 1.19, let

$$M_j := Q^j D (Q^T)^j \geq 0, \quad N_j := Q^j D (Q^T)^{j+1} + Q^{j+1} D (Q^T)^j$$

for $j = 0, \ldots, \tau+2$. Since $Q^T = K^{-1}CK = 2K^{-1}D - C^T$, we can apply Lemma 1.3 (iii) to $\sum\limits_{j=0}^{\tau} M_j$ and obtain

$$\sum_{j=0}^{\tau} M_j \geq c_0 \, \mathrm{Id}$$

for some $c_0 > 0$. Thus there is $c_1 > 0$ such that

$$M_{\tau+2} \leq c_1 \sum_{j=0}^{\tau} M_j. \tag{1.41}$$

We compute

$$QM_j + M_j Q^T = N_j, \tag{1.42}$$
$$QN_j + N_j Q^T = 2M_{j+1} + Q^j D (Q^T)^{j+2} + Q^{j+2} D (Q^T)^j. \tag{1.43}$$

Using $D^2 = D$, we have for any $\varepsilon > 0$:

$$\begin{aligned} 0 &\leq \left(\frac{1}{\sqrt{\varepsilon}} Q^j D \pm \sqrt{\varepsilon} Q^{j+2} D \right) \left(\frac{1}{\sqrt{\varepsilon}} D (Q^T)^j \pm \sqrt{\varepsilon} D (Q^T)^{j+2} \right) \\ &= \frac{1}{\varepsilon} M_j + \varepsilon M_{j+2} \pm \left(Q^j D (Q^T)^{j+2} + Q^{j+2} D (Q^T)^j \right). \end{aligned} \tag{1.44}$$

Then (1.42) and the analogue of (1.44) with $j+2$ replaced by $j+1$ yield the estimate

$$\pm N_j \leq \frac{1}{\varepsilon} M_j + \varepsilon M_{j+1}. \tag{1.45}$$

Further, (1.44) yields

$$\pm \left(Q^j D (Q^T)^{j+2} + Q^{j+2} D (Q^T)^j \right) \leq \frac{1}{\varepsilon} M_j + \varepsilon M_{j+2}. \tag{1.46}$$

Now let

$$P(t) := \sum_{j=0}^{\tau+1} \left(a_j t^{2j+1} M_j \right) + \sum_{j=0}^{\tau} \left(b_j t^{2j+2} N_j \right),$$

with $P(0)=0$. As (positive) coefficients, we first choose $a_{\tau+1} := \frac{1}{c_1}$,

$$b_\tau := \frac{2}{3}\left[1 + a_{\tau+1}(2\tau+4)\right], \quad a_\tau := 2\frac{b_\tau^2}{a_{\tau+1}}.$$

Then we choose iteratively, starting with $j=\tau$ and finishing with $j=1$:

$$b_{j-1} := \frac{2}{3}\left[2 + c_1 + a_j(2j+1) + b_j^2 + \frac{2(b_j(2j+2)-a_j)^2}{b_j}\right], \quad a_{j-1} := 8\frac{b_{j-1}^2}{a_j}. \tag{1.47}$$

Using (1.45) with $\varepsilon = \frac{2b_j t}{a_j}$, $0 \le j \le \tau$, we obtain

$$\forall j = 0,\dots,\tau: \quad b_j t^{2j+2} N_j \ge -\frac{a_j}{2} t^{2j+1} M_j - \frac{2b_j^2}{a_j} t^{2j+3} M_{j+1},$$

and thus

$$\begin{aligned}\sum_{j=0}^{\tau} b_j t^{2j+2} N_j &\ge -\frac{a_0}{2} t M_0 - \sum_{j=1}^{\tau}\left([\frac{a_j}{2} + \frac{2b_{j-1}^2}{a_{j-1}}] t^{2j+1} M_j \right) - \frac{2b_\tau^2}{a_\tau} t^{2\tau+3} M_{\tau+1} \\ &= -\frac{a_0}{2} t M_0 - \sum_{j=1}^{\tau}\left(\frac{3a_j}{4} t^{2j+1} M_j \right) - a_{\tau+1} t^{2\tau+3} M_{\tau+1},\end{aligned}$$

where we have used (1.47). Inserting this into $P(t)$ yields

$$P(t) \ge \frac{a_0}{2} t M_0 + \sum_{j=1}^{\tau} \frac{a_j}{4} t^{2j+1} M_j\,.$$

Writing $c_3 := \min\{\frac{a_0}{2}, \frac{a_1}{4}, \dots, \frac{a_\tau}{4}\}$, this implies for $t \in [0,1]$:

$$P(t) \ge t^{2\tau+1} c_3 \sum_{j=0}^{\tau} M_j \ge c_0 c_3 t^{2\tau+1}\,\mathrm{Id}\,. \tag{1.48}$$

So $P(t)$ is positive definite for all $t>0$, and we define the functional

$$\mathcal{F}(t) := \gamma e_\psi(f(t)|f_\infty) + \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u^T P(t) u f_\infty \,\mathrm{d}x \ge 0,$$

with some $\gamma > 0$ to be chosen later.

Step 2 (decay of $\mathcal{F}$): For $\mathcal{F}$, we can repeat all the computations in the proof of Proposition 1.25 and arrive at

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{dt}}\mathcal{F}(t) &\le -\gamma I_\psi(f(t)|f_\infty) + \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u^T \left[\dot{P}(t) - \left(QP(t) + P(t)Q^T\right)\right] u f_\infty \,\mathrm{d}x \\
&= \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u^T \left[\dot{P}(t) - \left(QP(t) + P(t)Q^T\right) - \gamma M_0\right] u f_\infty \,\mathrm{d}x.
\end{aligned}$$

We compute

$$\dot{P}(t) = \sum_{j=0}^{\tau+1} \left(a_j(2j+1)t^{2j}M_j\right) + \sum_{j=0}^{\tau} \left(b_j(2j+2)t^{2j+1}N_j\right)$$

and further, using (1.42), (1.43), and (1.46) with $\varepsilon := \frac{t^2}{b_j}$:

$$\begin{aligned}
-\left(QP(t) + P(t)Q^T\right) &= -\sum_{j=0}^{\tau+1} \left(a_j t^{2j+1} N_j\right) - 2\sum_{j=0}^{\tau} \left(b_j t^{2j+2} M_{j+1}\right) \\
&\quad - \sum_{j=0}^{\tau} \left(b_j t^{2j+2} [Q^j D (Q^T)^{j+2} + Q^{j+2} D (Q^T)]\right) \\
&\le -\sum_{j=0}^{\tau+1} \left(a_j t^{2j+1} N_j\right) - 2\sum_{j=0}^{\tau} \left(b_j t^{2j+2} M_{j+1}\right) \\
&\quad + \sum_{j=0}^{\tau} b_j t^{2j+2} \left(\frac{b_j}{t^2} M_j + \frac{t^2}{b_j} M_{j+2}\right) \\
&= -\sum_{j=0}^{\tau+1} \left(a_j t^{2j+1} N_j\right) - 2\sum_{j=0}^{\tau} \left(b_j t^{2j+2} M_{j+1}\right) \\
&\quad + \sum_{j=0}^{\tau} \left(t^{2j} b_j^2 M_j\right) + \sum_{j=2}^{\tau+2} \left(t^{2j} M_j\right).
\end{aligned}$$

This implies

$$\begin{aligned}
&\dot{P}(t) - \left(QP(t) + P(t)Q^T\right) - \gamma M_0 \\
&\le \left(a_0 + b_0^2 - \gamma\right) M_0 + \left(3a_1 + b_1^2 - 2b_0\right) t^2 M_1 \\
&\quad + \sum_{j=2}^{\tau} \left([a_j(2j+1) + 1 + b_j^2 - 2b_{j-1}] t^{2j} M_j\right) \\
&\quad + (a_{\tau+1}(2\tau+3) + 1 - 2b_\tau) t^{2\tau+2} M_{\tau+1} \\
&\quad + \sum_{j=0}^{\tau+1} \left(\alpha_j t^{2j+1} N_j\right) + t^{2(\tau+2)} M_{\tau+2},
\end{aligned}$$

where $\alpha_j := -a_j + b_j(2j+2)$, $0 \le j \le \tau$; $\alpha_{\tau+1} := -a_{\tau+1}$. Using

$$\forall j = 0, \dots, \tau: \quad \pm N_j \le \frac{2|\alpha_j|}{b_j t} M_j + \frac{b_j t}{2|\alpha_j|} M_{j+1},$$
$$N_{\tau+1} \le \frac{1}{t} M_{\tau+1} + t M_{\tau+2},$$

we obtain

$$\begin{aligned}
&\sum_{j=0}^{\tau+1} \left(\alpha_j t^{2j+1} N_j\right) \\
&\le \sum_{j=0}^{\tau} \left(\frac{2\alpha_j^2}{b_j} t^{2j} M_j + \frac{b_j}{2} t^{2j+2} M_{j+1} \right) + a_{\tau+1} t^{2\tau+2} M_{\tau+1} + a_{\tau+1} t^{2\tau+4} M_{\tau+2} \\
&= \frac{2\alpha_0^2}{b_0} M_0 + \sum_{j=1}^{\tau} \left(\frac{2\alpha_j^2}{b_j} + \frac{b_{j-1}}{2} \right) t^{2j} M_j + (\frac{b_\tau}{2} + a_{\tau+1}) t^{2\tau+2} M_{\tau+1} + a_{\tau+1} t^{2\tau+4} M_{\tau+2}.
\end{aligned}$$

Thus, we finally arrive at

$$\begin{aligned}
&\dot{P}(t) - \left(QP(t) + P(t)Q^T\right) - \gamma M_0 \\
&\le \left(a_0 + b_0^2 + \frac{2\alpha_0^2}{b_0} - \gamma \right) M_0 + \left(3a_1 + b_1^2 + \frac{2\alpha_1^2}{b_1} + \frac{b_0}{2} - 2b_0 \right) t^2 M_1 \\
&\quad + \sum_{j=2}^{\tau} \left(a_j(2j+1) + 1 + b_j^2 + \frac{2\alpha_j^2}{b_j} + \frac{b_{j-1}}{2} - 2b_{j-1} \right) t^{2j} M_j \\
&\quad + (a_{\tau+1}(2\tau+4) + 1 + \frac{b_\tau}{2} - 2b_\tau) t^{2\tau+2} M_{\tau+1} + (a_{\tau+1} + 1) t^{2\tau+4} M_{\tau+2}.
\end{aligned}$$

We use (1.41) and obtain for $t \in [0,1]$:

$$\begin{aligned}
&\dot{P}(t) - \left(QP(t) + P(t)Q^T\right) - \gamma M_0 \\
&\le \left[c_1(a_{\tau+1} + 1) + a_0 + b_0^2 + \frac{2\alpha_0^2}{b_0} - \gamma \right] M_0 \\
&\quad + \left[c_1(a_{\tau+1} + 1) + 3a_1 + b_1^2 + \frac{2\alpha_1^2}{b_1} - \frac{3b_0}{2} \right] t^2 M_1 \\
&\quad + \sum_{j=2}^{\tau} \left(\left[c_1(a_{\tau+1} + 1) + a_j(2j+1) + 1 + b_j^2 + \frac{2\alpha_j^2}{b_j} - \frac{3b_{j-1}}{2} \right] t^{2j} M_j \right) \\
&\quad + \left[a_{\tau+1}(2\tau+4) + 1 - \frac{3b_\tau}{2} \right] t^{2\tau+2} M_{\tau+1}.
\end{aligned}$$

Using (1.47), we obtain that all the coefficients in square brackets are non-positive for sufficiently large γ, and thus

$$\dot{P}(t) - \left(QP(t) + P(t)Q^T\right) - \gamma M_0 \le 0$$

This implies that $\mathcal{F}(t)$ is monotonously decreasing, and thus $\mathcal{F}(t) \le \mathcal{F}(0) = \gamma e_\psi(f_0|f_\infty)$ for all t in $[0,1]$. Together with (1.48), we obtain

$$c_0 c_3 t^{2\tau+1} \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})|u|^2 f_\infty \,\mathrm{d}x \le \gamma e_\psi(f_0|f_\infty),$$

which completes the proof using Lemma 1.19 (iii). □

1.3.5 Decay of admissible relative entropies

With this regularisation result, we can finally prove exponential decay of the relative entropy:

Theorem 1.27. *Assume condition (A). Let ψ generate an admissible relative entropy and let f be the solution to (2) with initial state $f_0 \in L^1_+(\mathbb{R}^d)$ such that $e_\psi(f_0|f_\infty) < \infty$. Let $\{\lambda_m | 1 \le m \le m_0\}$ be the eigenvalues of C with $\mu = \Re\{\lambda_m\}$, and let*

$$e(t) := e_\psi(f(t)|f_\infty).$$

Then

(i) If all λ_m, $1 \le m \le m_0$, are non-defective, then there is a constant $c > 0$ such that

$$\forall t \ge 0 : \quad e(t) \le c e^{-2\mu t} e_\psi(f_0|f_\infty).$$

(ii) If λ_m is defective for at least one $m \in \{1, \dots, m_0\}$, then for all $\varepsilon \in (0, \mu)$, there is $c_\varepsilon > 0$ such that

$$\forall t \ge 0 : \quad e(t) \le c_\varepsilon e^{-2(\mu-\varepsilon)t} e_\psi(f_0|f_\infty).$$

Proof: Let P, $S_\psi(f_0)$ be defined as in Lemma 1.19. Let $\kappa = \mu$ in case (i), and $\kappa = \mu - \varepsilon$ in case (ii). Let $\delta > 0$. Using (1.29), Proposition 1.25 and Theorem 1.26, we compute for $t \ge \delta$:

$$\begin{aligned} e_\psi(t) &\le \frac{1}{2\lambda_P} S_\psi(f(t)) \le \frac{1}{2\lambda_P} S_\psi(f(\delta)) e^{-2\kappa(t-\delta)} \\ &\le e^{2\kappa\delta} \frac{c_r}{2\lambda_P \delta^{2\tau+1}} e_\psi(0) e^{-2\kappa t}. \end{aligned} \tag{1.49}$$

For $t \leq \delta$, it follows from the monotonicity of e_ψ (cf. (1.18)) that

$$e_\psi(t) \leq e_\psi(0)\,. \tag{1.50}$$

Writing $c_\delta := e^{2\kappa\delta} \max\{1, \frac{c_\tau}{2\lambda_P \delta^{2\tau+1}}\}$ and combining (1.49), (1.50) yields

$$\forall t \geq 0: \quad e_\psi(t) \leq c_\delta e_\psi(0) e^{-2\kappa t}.$$

c_δ can now be optimized for $\delta > 0$, completing the proof. □

Remark: Let us add a remark on an alternative approach to the above inequality

$$e_\psi(t) \leq \frac{1}{2\lambda_P} S_\psi(t). \tag{1.51}$$

Since $S_\psi(f(t)) \geq c_P I_\psi(f(t))$, we could also integrate the inequality

$$-\tfrac{\mathrm{d}}{\mathrm{dt}} e_\psi(f(t)) \leq \frac{1}{c_P} S_\psi(f_0) e^{-2\kappa t}$$

on (t, ∞), which resembles the procedure in the standard entropy method [6]. This would yield a constant $\frac{1}{2c_P\kappa}$ instead of $\frac{1}{2\lambda_P}$ in (1.51). However, this constant is never better. To prove this, we observe that

$$\begin{aligned} CK + KC^T &= 2D \leq \frac{2}{c_P} P \\ &\leq \frac{1}{c_P\kappa}\left(QP + PQ^T\right) = \frac{1}{c_P\kappa}\left(KC^TK^{-1}P + PK^{-1}CK\right). \end{aligned}$$

Multiplying by $K^{-1} > 0$ from left and right yields

$$K^{-1}C + C^TK^{-1} \leq \frac{1}{c_P\kappa}\left(C^TK^{-1}PK^{-1} + K^{-1}PK^{-1}C\right),$$

which implies (since $c_P\kappa > 0$)

$$0 \leq \left(K^{-1}PK^{-1} - c_P\kappa K^{-1}\right) C + C^T \left(K^{-1}PK^{-1} - c_P\kappa K^{-1}\right) =: W.$$

This means that $K^{-1}PK^{-1} - c_P\kappa K^{-1}$ is a solution to the continuous Lyapunov equation

$$MC + C^TM = W. \tag{1.52}$$

But since C is positively stable and $W \geq 0$, (1.52) has a unique positive semidefinite solution M (see Theorem 2.2 in [61], Theorem 2.2.3 in [44]). So

$$\left(K^{-1}PK^{-1} - c_P\kappa K^{-1}\right) \geq 0,$$

which implies

$$P \geq c_P\kappa K.$$

Since λ_P is chosen as the largest constant such that

$$P \geq \lambda_P K,$$

it follows that $\lambda_P \geq c_P\kappa$ and thus

$$\frac{1}{2c_P\kappa} \geq \frac{1}{2\lambda_P}.$$

□

1.3.6 Special solutions and sharp decay rate

In this section, we investigate the sharpness of the decay rate obtained in Theorem 1.27 under condition (A). In particular, we show that the rate is optimal for both the quadratic entropy e_2 and the logarithmic entropy e_1. As shown in [6], all admissible entropies are bounded below by a logarithmic entropy and above by a quadratic one. Thus, the rate we obtained is optimal for all admissible entropies.

Theorem 1.28. *Let $\mu := \min\{\Re\{\lambda\}|\lambda \in \sigma(C)\}$, where $\sigma(C)$ denotes the spectrum of C.*

(i) If μ is a (real) eigenvalue of C, then there exist initial conditions f_0, g_0 (different from f_∞) such that for the corresponding solutions $f(t)$, $g(t)$ of (2), it holds that

$$e_1(f(t)) = e^{-2\mu t}e_1(f_0), \quad e_2(g(t)) = e^{-2\mu t}e_2(g_0), \quad t \geq 0.$$

(ii) If C has a complex conjugate eigenvalue pair with $\Re\{\lambda_{1,2}\} = \mu$, then there are initial conditions f_0, g_0 (different from f_∞) such that for the corresponding solutions $f(t)$, $g(t)$ of (2), it holds that

$$e_1(f(t)) \leq ce^{-2\mu t}e_1(f_0), \quad e_2(g(t)) \leq ce^{-2\mu t}e_2(g_0), \quad t \geq 0, \tag{1.53}$$

with some $c > 0$, and equality holds for $t = t_0 + n\tau$, $t_0 \geq 0$, $\tau > 0$, $n \in \mathbb{N}_0$. So the right hand sides of 1.53 are the sharp exponential envelope functions for the entropy decay.

(iii) *If C has a defective eigenvalue λ with $\Re\{\lambda\} = \mu$, then there are initial conditions f_0, g_0 (different from f_∞) such that for the corresponding solutions $f(t)$, $g(t)$ of (2), it holds that*

$$e_1(f(t)) = c_0 e^{-2\mu t}(e_1(f_0) + \frac{c_1}{2}t + \frac{c_2}{2}t^2),$$
$$e_2(g(t)) = c_0 e^{-2\mu t}(e_2(g_0) + c_1 t + c_2 t^2)$$

for $t \geq 0$ and some $c_0, c_2 > 0$, $c_1 \in \mathbb{R}$.

In all cases, f_0 is ψ_1-compatible and g_0 is ψ_2-compatible.

Remark: In the defective case (iii), the decay rate is indeed reduced to $2(\mu - \varepsilon)$ for an arbitrarily small $\varepsilon > 0$ – as announced in Theorem 1.27.

The proof of Theorem 1.28 is based on special solutions of (2), which will be computed in the next two lemmas. These computations are inspired by Theorem 3.11 in [6], where the sharpness of the convex Sobolev inequality (1.29) is discussed.

Lemma 1.29. *Let $v_0 \in \mathbb{R}^d$. Then*

(i)

$$f_0(x) := c_K \exp(-V(x) + v_0^T x - \frac{v_0^T K v_0}{2})$$

is in L^1_+ with $\int_{\mathbb{R}^d} f_0 \, \mathrm{d}x = 1$. Here, $V(x) = \frac{x^T K^{-1} x}{2}$ from Theorem 1.12, $f_\infty = c_K e^{-V}$. Furthermore, f_0 is ψ-compatible for the logarithmic entropies (1.22).

(ii) *If $v(t)$ solves*

$$\dot{v}(t) = -K^{-1}CKv(t), \quad v(t=0) = v_0,$$

then

$$f(t,x) =: \exp(-V(x) + v(t)^T x - \frac{v(t)^T K v(t)}{2})$$

is a solution to (2) with initial condition f_0.

(iii) For the relative logarithmic entropy $e_1(t) := e_1(f(t)|f_\infty)$*, it holds that*

$$e_1(t) = \frac{v(t)^T K v(t)}{2}.$$

Proof:

(i): We have from (1.23) that

$$w = 2\sqrt{\alpha}\left(\sqrt{\frac{f_0}{f_\infty} + \beta} - \sqrt{1+\beta}\right) = 2\sqrt{\alpha}\left(\sqrt{\exp(v_0^T x + c_g) + \beta} - \sqrt{1+\beta}\right)$$

for some α, $\beta > 0$. This implies

$$\nabla w = \frac{\sqrt{\alpha} v_0 \exp(v_0^T x + c_g)}{\sqrt{\exp(v_0^T x + c_g) + \beta}} \in L^2(\mathbb{R}^d, f_\infty),$$

so f_0 is ψ-compatible for logarithmic ψ by Definition 1.18. $\int\limits_{\mathbb{R}^d} f_0 \,\mathrm{d}x = 1$ is easily checked.

(ii): We insert $f(t,x)$ into (2) and obtain

$$\begin{aligned} f_t &= \left(x^T \dot{v}(t) - v(t)^T K \dot{v}(t)\right) f(t), \\ \operatorname{div}(f_\infty (D+R) \nabla \frac{f}{f_\infty}) &= \operatorname{div}\left[f_\infty C K \nabla \exp(v(t)^T x - \frac{v(t)^T K v(t)}{2})\right] \\ &= \operatorname{div}[f(t) C K v(t)] = (\nabla f(t)) \cdot C K v(t) \\ &= \left(-x^T K^{-1} C K v(t) + v(t)^T C K v(t)\right) f(t), \end{aligned}$$

where we have used $D + R = CK$ and the symmetry of K.

(iii): Setting $\alpha = 1$, $\beta = 0$ for ease of computation, it is $\psi_1(s) = s\ln(s) - s + 1$. We compute

$$\begin{aligned} e_1(f(t)) &= \int\limits_{\mathbb{R}^d} \left[\frac{f(t)}{f_\infty} \ln(\frac{f(t)}{f_\infty}) - \frac{f(t)}{f_\infty} + 1\right] f_\infty \,\mathrm{d}x \\ &= \int\limits_{\mathbb{R}^d} f(t) \ln(\frac{f(t)}{f_\infty}) \,\mathrm{d}x \\ &= \int\limits_{\mathbb{R}^d} \ln\left(\exp\left(v(t)^T x - \frac{v(t)^T K v(t)}{2}\right)\right) f(t) \,\mathrm{d}x \\ &= \int\limits_{\mathbb{R}^d} \left(v(t)^T x - \frac{v(t)^T K v(t)}{2}\right) f(t) \,\mathrm{d}x. \end{aligned}$$

For fixed $t \geq 0$, it holds that

$$\begin{aligned}\partial_{Kv(t)} f(t,x) &:= v(t)^T K \nabla f(t,x) = v(t)^T K \left(-K^{-1}x + v(t)\right) f(t,x) \\ &= \left(-v(t) \cdot x + v(t)^T K v(t)\right) f(t,x).\end{aligned}$$

Since

$$\int\limits_{\mathbb{R}^d} \partial_{Kv(t)} f(t,x) \, \mathrm{d}x = 0,$$

we obtain

$$e_1(t) = \frac{v(t)^T K v(t)}{2} \int\limits_{\mathbb{R}^d} f(t) \, \mathrm{d}x = \frac{v(t)^T K v(t)}{2}.$$

□

Lemma 1.30. *Let $v_0 \in \mathbb{R}^d$. Then*

(i)

$$f_0(x) = (1 + x^T v_0) f_\infty$$

is in $L^1(\mathbb{R}^d)$ with $\int\limits_{\mathbb{R}^d} f_0 \, \mathrm{d}x = 1$. Furthermore, f_0 is ψ-compatible for the quadratic case $\psi(s) = (s-1)^2$.

(ii) If $v(t)$ solves

$$\dot{v}(t) = -K^{-1} C K v(t), \quad v(t=0) = v_0,$$

then

$$f(t,x) := (1 + x^T v(t)) f_\infty$$

is a solution to (2) with initial condition f_0.

(iii) For the quadratic relative entropy $e_2(t) := e_2(f(t)|f_\infty)$, it holds that

$$e_2(t) = v(t)^T K v(t).$$

Proof: First, we note that $f_0 \geq 0$ does not hold here. This is not a problem, since we don't need positivity of the solution to define the quadratic entropy.

(i): Since $f_\infty(x) = f_\infty(-x)$, we have

$$\int_{\mathbb{R}^d} v_0^T x f_\infty \,\mathrm{d}x = 0,$$

and $\int_{\mathbb{R}^d} f_0 \,\mathrm{d}x = 1$ follows from the normalisation of f_∞. We recall from (1.21) that for quadratic ψ,

$$w = \sqrt{2\alpha}(\frac{f_0}{f_\infty} - 1) = \sqrt{2\alpha} v_0^T x$$

for some $\alpha > 0$. Then

$$\nabla w = \sqrt{2\alpha} v_0 \in L^2(\mathbb{R}^d, f_\infty),$$

and thus f_0 is ψ-compatible for quadratic ψ by Definition 1.19.

(ii): We insert $f(t,x)$ into (2) and obtain

$$f_t(t,x) = x^T \dot{v}(t) f_\infty,$$
$$\operatorname{div}(f_\infty (D+R)\nabla \frac{f(t,x)}{f_\infty}) = \operatorname{div}(f_\infty C K v(t)) = x^T K^{-1} C K v(t) f_\infty,$$

where we again used $D + R = CK$.

(iii): For a quadratic entropy with $\psi_2(s) = \alpha(s-1)^2$, we compute

$$e_2(f(t)) = \alpha \int_{\mathbb{R}^d} (\frac{f(t,x)}{f_\infty} - 1)^2 f_\infty \,\mathrm{d}x = \alpha \int_{\mathbb{R}^d} (x^T v(t))^2 f_\infty \,\mathrm{d}x.$$

For fixed $t \geq 0$, it holds that

$$\partial_{Kv(t)} f_\infty = -v(t)^T K K^{-1} x f_\infty = -x^T v(t) f_\infty.$$

It follows that

$$\begin{aligned} e_2(f(t)) &= -\alpha \int_{\mathbb{R}^d} (x^T v(t)) \partial_{Kv(t)} f_\infty \,\mathrm{d}x = \alpha \int_{\mathbb{R}^d} f_\infty \partial_{Kv(t)} (x^T v(t)) \,\mathrm{d}x \\ &= \alpha v(t)^T K v(t) \int_{\mathbb{R}^d} f_\infty \,\mathrm{d}x = v(t)^T K v(t). \end{aligned}$$

□

From Lemmas 1.29 and 1.30, we see that we can reduce the discussion of sharp decay rates for relative entropies to discussing the term $v(t)^T K v(t)$, where $v_0 \in \mathbb{R}^d$ and

$$\dot{v}(t) = -K^{-1}CKv(t), \quad v(t=0) = v_0. \tag{1.54}$$

A direct consequence is

Corollary 1.31. *Let condition (A) hold, and let $t^* \in \mathbb{R}_0^+$. Then there is an initial condition f_0 [g_0] distinct from f_∞ such that for the solution $f(t)$ [$g(t)$] to (2), the entropy dissipation I_{ψ_1} [I_{ψ_2}] (see (1.18)) for the logarithmic [quadratic] entropy vanishes at t^*, i.e. $I_{\psi_1}(f(t^*)) = 0$ [$I_{\psi_2}(g(t^*)) = 0$].*

Proof: We take the time derivative of $v(t)^T K v(t)$, where v fulfils (1.54), and obtain

$$\tfrac{\mathrm{d}}{\mathrm{dt}} \left[v(t)^T K v(t) \right] = -v(t)^T K C^T v(t) - v(t)^T C K v(t) = -2v(t)^T D v(t),$$

where we have used (1.12). Let $0 \neq w \in \ker D$. Setting $v_0 := \exp(K^{-1}CKt^*)w$ implies $v(t^*) = w$, which completes the proof. $\square$

We will now use Lemmas 1.29 and 1.30 to prove Theorem 1.28.

Proof (of Theorem 1.28):

(i): There is $0 \neq v_0 \in \mathbb{R}^d$ with $K^{-1}CKv_0 = \mu v_0$. So the solution of (1.54) is $v(t) = e^{-\mu t} v_0$, and thus

$$v(t)^T K v(t) = e^{-2\mu t} v_0 K v_0.$$

(ii): There is $0 \neq w \in \mathbb{C}^d$ with $K^{-1}CKw = \lambda w$, $\lambda \in \mathbb{C}$, $\Re\{\lambda\} = \mu > 0$, $\Im\{\lambda\} = \omega \neq 0$. We can choose $w \notin i\mathbb{R}^d$. Then $\overline{w}$ fulfils $K^{-1}CK\overline{w} = \overline{\lambda}\overline{w}$, since $K^{-1}CK$ is real. Moreover $v_0 := w + \overline{w} \in \mathbb{R}^d$, and $v_1 := i(\overline{w} - w) \in \mathbb{R}^d$. One easily verifies that $v(t) := e^{-\mu t} \left(\cos(\omega t) v_0 + \sin(\omega t) v_1 \right)$ is the solution to (1.54). We define

$$c := \sup_{t \in \mathbb{R}_0^+} \left(\cos(\omega t) v_0 + \sin(\omega t) v_1 \right)^T K \left(\cos(\omega t) v_0 + \sin(\omega t) v_1 \right) > 0,$$

since K is positive definite. Since $v(t)$ is $\frac{2\pi}{\omega}$-periodic, the value c is attained for

$t = t_0 + k\frac{2\pi}{\omega}$, $t_0 \in \mathbb{R}_0^+$. It follows that

$$v(t)^T K v(t) = e^{-2\mu t} \left(\cos(\omega t) v_0 + \sin(\omega t) v_1\right)^T K \left(\cos(\omega t) v_0 + \sin(\omega t) v_1\right) \leq c e^{-2\mu t},$$

with equality for $t = t_0 + k\frac{2\pi}{\omega}$.

(iii): We confine ourselves here to the case $\lambda = \mu \in \mathbb{R}$; the general case can be obtained by an extension of (ii). So, let $w, h \in \mathbb{R}^d$ with $K^{-1}CKw = \mu w$, $K^{-1}CKh = \mu h + w$. Let $v_0 := h$, then $v(t) := e^{-\mu t}(h - tw)$ is the solution to (1.54), and

$$v(t)^T K v(t) = e^{-2\mu t}(h - tw)^T K (h - tw) = \exp(-2\mu t)(v_0^T K v_0 + c_1 t + c_2 t^2).$$

□

From the proof of Theorem 1.28, we see that the constant c in the estimate $e_\psi(f(t)) \leq c e^{-2\mu t}$ does not derive from the initial state in a straightforward way, unless all eigenvalues of C are real and non-defective. For case (ii), if $|v_1| \gg |v_0|$, then c can be very large in comparison to $e_\psi(f_0)$; for case (iii), the same holds for $|w| \gg |h|$.

1.4 Spectral analysis

We now give a characterisation of the spectrum of L in L^2. Let $\lambda_1, \ldots, \lambda_d$ be the eigenvalues of C, counted with their algebraic multiplicity. Let $\mathcal{P}(\mathbb{R}^d)$ denote the polynomials over $\mathbb{R}^d$ (with complex coefficients) and let $\mathcal{Q} := \mathcal{P}(\mathbb{R}^d) f_\infty$. $\mathcal{Q}$ is dense in $L^2(\mathbb{R}^d, f_\infty^{-1})$, and it is the natural space for eigenfunctions of the (symmetric) Fokker-Planck operator (see for example [39] or [57]).

Theorem 1.32. *Assume condition (A). Then the following holds:*

(i) The spectrum of L in L^2 is given by

$$\sigma(L) = \sigma_p(L) = \Big\{ -\sum_{j=1}^{d} \alpha_j \lambda_j \Big| \alpha = (\alpha_j) \in \mathbb{N}_0^d \Big\} \subset \{0\} \cup \left(\mathbb{R}^- \times i\mathbb{R}\right).$$

(ii) The eigenspace to 0 is one-dimensional and spanned by f_∞. For each eigenvalue ν of L with $\Re\{\nu\} < 0$, the corresponding eigenfunctions and generalised eigenfunctions span a finite dimensional subspace of $\mathcal{Q}$.

(iii) If C is not defective, then the eigenfunctions of L form a basis of $\mathcal{Q}$.

(iv) If C is defective, then the eigenfunctions and generalised eigenfunctions of L form a basis of $\mathcal{Q}$.

Remark: As we will see in Lemma 1.34, all eigenfunctions and generalised eigenfunctions of L can be computed explicitly.

The proof of Theorem 1.32 is split into three lemmas. First, we show in Lemma 1.33 that L only has a point spectrum by proving the compactness of the resolvent of L. In Lemma 1.34, we explicitly compute the (generalised) eigenfunctions of L in $\mathcal{Q}$. Finally, in Lemma 1.36 we establish an orthogonal decomposition of L^2 into finite dimensional subspaces, which allows us to prove that the (generalised) eigenfunctions from Lemma 1.34 are indeed all the (generalised) eigenfunctions of L.
The main difficulty here is that the eigenfunctions of L will are not orthogonal, in contrast to the symmetric, fully parabolic case. They do, however, generate L–invariant and mutually orthogonal subspaces of $\mathcal{Q}$, and this fact can be exploited in the proof of Lemma 1.36 (see also [34], [5]).

In the next lemma we shall need a weighted H^1-space:

$$\mathcal{H} := \{f \in L^2 | \nabla(\tfrac{f}{f_\infty}) \in (L^2(\mathbb{R}^d, f_\infty))^d\},$$
$$\|f\|^2_{\mathcal{H}} := \int\limits_{\mathbb{R}^d} |f|^2 f_\infty^{-1}\, \mathrm{d}x + \int\limits_{\mathbb{R}^d} |\nabla \frac{f}{f_\infty}|^2 f_\infty\, \mathrm{d}x = \|\frac{f}{f_\infty}\|^2_{H^1(\mathbb{R}^d, f_\infty)}\,.$$

Lemma 1.33. *Under condition (A), the operator L has a compact resolvent on $L^2(\mathbb{R}^d, f_\infty^{-1})$.*

Proof: For a (uniformly) elliptic operator, compactness of the resolvent can be shown by establishing that the embedding $\mathcal{H} \hookrightarrow L^2$ is compact. For a degenerate elliptic operator, the resolvent will in general not map into $\mathcal{H}$, so one has to work in spaces with fractional derivatives. For this proof, we shall therefore proceed in three steps. First we establish the space we work in, then we extend the regularisation result from Theorem 1.26 for the solution semigroup e^{Lt} on L^2. Finally, we use these two results to show compactness of the resolvent of L.

Step 1 (interpolation spaces $\mathcal{H}_r$): We start by introducing the spaces $\mathcal{H}_r$, $0 < r < 1$, between L^2 and $\mathcal{H}$. We remark that, for any $f \in L^2$,

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

By definition of $\mathcal{H}$, this implies that for any $f \in \mathcal{H}$,

$$\|f\|_{\mathcal{H}} = \|\frac{f}{f_\infty}\|_{H^1(\mathbb{R}^d, f_\infty)}.$$

An orthonormal basis $\{z_j | j \in \mathbb{N}_0^d\}$ of $L^2(\mathbb{R}^d, f_\infty)$ is given by the "polynomial part" of the eigenfunctions of the (uniformly) elliptic Fokker-Planck operator

$$L_{\mathrm{Id}} f := \mathrm{div}(\nabla(\frac{f}{f_\infty}) f_\infty)$$

in L^2. It satisfies

$$L_{\mathrm{Id}}(z_j f_\infty) = -|j| z_j f_\infty,$$

with $|j|$ the degree of the multi-index j. In the case of gaussian f_∞, the z_j are the standard Hermite polynomials. For $f \in \mathcal{H}$ it holds

$$\|f\|_{L^2}^2 = \sum_{j \in \mathbb{N}_0^d} |c_j|^2, \quad \|f\|_{\mathcal{H}}^2 = \sum_{j \in \mathbb{N}_0^d} (1 + |j|) |c_j|^2,$$

where c_j is the coefficient of $\frac{f}{f_\infty}$ along z_j. We thus define

$$\mathcal{H}_r := \{f \in L^2 | \sum_{j \in \mathbb{N}_0^d} (1 + |j|)^r |c_j|^2 < \infty\}, \tag{1.55}$$

Using Hölder's inequality with $p = \frac{1}{r}$, $q = \frac{1}{1-r}$, we obtain

$$\begin{aligned} \|f\|_{\mathcal{H}_r}^2 &= \sum_{j \in \mathbb{N}_0^d} (1 + |j|)^r |c_j|^{2r} |c_j|^{2-2r} \\ &\le \left(\sum_{j \in \mathbb{N}_0^d} (1 + |j|) |c_j|^2 \right)^r \left(\sum_{j \in \mathbb{N}_0^d} |c_j|^2 \right)^{1-r}. \end{aligned}$$

This yields the interpolation inequality

$$\|f\|_{\mathcal{H}_r} \le \|f\|_{\mathcal{H}}^r \|f\|_{L^2}^{1-r}. \tag{1.56}$$

<u>Step 2 (regularisation from L^2 to $\mathcal{H}_r$)</u>: Since L generates a contraction semigroup on L^2, we have

$$\forall t \ge 0 : \quad \|e^{Lt} f\|_{L^2} \le \|f\|_{L^2}. \tag{1.57}$$

In the following estimate, we shall use the L^2-orthogonal decomposition $f =$

$\tilde{f} + f_\infty \int_{\mathbb{R}^d} f \, \mathrm{d}x$ with $\int_{\mathbb{R}^d} \tilde{f} \, \mathrm{d}x = 0$, and the scaled version of (1.40) for quadratic ψ:

$$\int_{\mathbb{R}^d} (\nabla \frac{f(t)}{f_\infty})^T P \nabla \frac{f(t)}{f_\infty} f_\infty \, \mathrm{d}x \leq ct^{-(2\tau+1)} \int_{\mathbb{R}^d} \Big(f - f_\infty \int_{\mathbb{R}^d} f \, \mathrm{d}x \Big)^2 f_\infty^{-1} \, \mathrm{d}x.$$

We have:

$$\begin{aligned}
\|e^{Lt} f\|_{\mathcal{H}}^2 &= \|e^{Lt} f\|_{L^2}^2 + \Big\| \nabla \frac{e^{Lt} f}{f_\infty} \Big\|_{L^2(\mathbb{R}^d, f_\infty)}^2 \\
&= \|e^{Lt} f\|_{L^2}^2 + \Big\| \nabla \frac{e^{Lt} \tilde{f}}{f_\infty} \Big\|_{L^2(\mathbb{R}^d, f_\infty)}^2 \\
&\leq \|f\|_{L^2}^2 + ct^{-(2\tau+1)} \|\tilde{f}\|_{L^2}^2 \\
&\leq (1 + ct^{-(2\tau+1)}) \|f\|_{L^2}^2,
\end{aligned}$$

where we have used the L^2-contractivity of e^{Lt} and that P is positive definite. We thus get

$$\forall 0 < t \leq 1: \quad \|e^{Lt} f\|_{\mathcal{H}} \leq \tilde{c} t^{-(\tau + \frac{1}{2})} \|f\|_{L^2} \tag{1.58}$$

for all $f \in L^2$. By combining (1.56) – (1.58), we obtain

$$\forall 0 < t \leq 1: \quad \|e^{Lt} f\|_{\mathcal{H}_r} \leq \beta t^{-r(\tau + \frac{1}{2})} \|f\|_{L^2}, \tag{1.59}$$

with $\beta := \tilde{c}^r$.

<u>Step 3 (compact resolvent)</u>: For $r := \frac{1}{\tau+1} > 0$, we can integrate (1.59) on $(0, 1)$. This yields

$$\| \int_0^1 e^{Lt} f \, \mathrm{d}t \|_{\mathcal{H}_r} \leq c \|f\|_{L^2}. \tag{1.60}$$

By a well-known result for semigroups (see e.g. [32], section II.1, Lemma 1.3 or [56], section 1.2, Theorem 2.4), for any $\lambda > 0$ it holds that

$$\forall f \in D(L) \; \forall t > 0: \quad \int_0^t e^{(L-\lambda)s} (L - \lambda) f \, \mathrm{d}s = e^{(L-\lambda)t} f - f.$$

Due to (1.57), $e^{(L-\lambda)t}$ decays exponentially and we conclude

$$\int_1^\infty e^{(L-\lambda)t} (\lambda - L) f \, \mathrm{d}t = e^{L-\lambda} f \tag{1.61}$$

for all $f \in D(L)$. Moreover (see e.g. [32], section II.1, Theorem 1.10 or [56], section 1.3, Theorem 3.1), the resolvent $R(\lambda, L) := (\lambda - L)^{-1}$ has the representation

$$R(\lambda, L) = \int_0^\infty e^{(L-\lambda)t} \, \mathrm{d}t = \int_0^1 e^{(L-\lambda)t} \, \mathrm{d}t + \int_1^\infty e^{L-\lambda)t} \, \mathrm{d}t.$$

We apply this representation to (1.60) and obtain

$$c\|f\|_{L^2} \geq \|[R(\lambda, L) - \int_1^\infty e^{(L-\lambda)t}]f \, \mathrm{d}t\|_{\mathcal{H}_r},$$

which yields

$$\|R(\lambda, L)f\|_{\mathcal{H}_r} \leq c\|f\|_{L^2} + \|\int_1^\infty e^{(L-\lambda)t} f \, \mathrm{d}t\|_{\mathcal{H}_r}. \tag{1.62}$$

For $g \in D(L)$, we replace f in (1.62) by $(\lambda - L)g$ and obtain, using (1.61),

$$\begin{aligned} \|g\|_{\mathcal{H}_r} &\leq c\|(\lambda - L)g\|_{L^2} + \|\int_1^\infty e^{(L-\lambda)t}(\lambda - L)g \, \mathrm{d}t\|_{\mathcal{H}_r} \\ &= c\|(\lambda - L)g\|_{L^2} + e^{-\lambda}\|e^L g\|_{\mathcal{H}_r}. \end{aligned}$$

Applying (1.59) with $t = 1$ to the last term yields

$$\|g\|_{\mathcal{H}_r} \leq c\|(\lambda - L)g\|_{L^2} + \beta e^{-\lambda}\|g\|_{L^2}.$$

Choosing $\lambda > \ln \beta$ allows to "absorb" the last term into the left-hand side. Due to the spectral representation of $\mathcal{H}_r$ in (1.55), the embedding $\mathcal{H}_r \hookrightarrow L^2$ is compact for $r > 0$. Hence, $R(\lambda, L)$ is compact for the chosen λ, and by the *first resolvent formula* then also for all λ in the resolvent set. $\square$

Remark: For $r = 1$, the compactness of the embedding can also be shown by the method in [38].

In the next lemma, we compute (generalised) eigenfunctions of L. Here we shall use the following notation for multi-indexes. Let $\alpha \in \mathbb{N}_0^d$ be a multi-index. We write $|\alpha| = \sum_{j=1}^d \alpha_j$, $\nabla^\alpha := \sum_{j=1}^d \partial_j^{\alpha_j}$. We also introduce the notation α_{l-} and

α_{l+}:

$$\begin{aligned}(\alpha_{l+})_j &:= \alpha_j \ (j \neq l), \quad (\alpha_{l+})_l := \alpha_l + 1,\\ (\alpha_{l-})_j &:= \alpha_j \ (j \neq l), \quad (\alpha_{l-})_l := \alpha_l - 1 \quad \text{ if } \alpha_l > 1,\\ &\alpha_{l-} := 0 \in \mathbb{N}_0^d \quad \text{ if } \alpha_l = 0.\end{aligned}$$

So α_{l-}, α_{l+} denote the multi-index that one gets by lowering or raising the l-th entry of α by 1. Analogously we define iterated index shifts like, e.g., $(\alpha_{l-})_{m-}$.

Lemma 1.34. *There is a bijection Φ between $\mathbb{N}_0^d$ and the (generalised) eigenfunctions $\varphi \in \mathcal{Q}$ of L. For $\alpha \in \mathbb{N}_0^d$, the polynomial part of $\Phi(\alpha)$ has degree $|\alpha|$, and the eigenvalue corresponding to $\Phi(\alpha)$ is*

$$\nu_\alpha := -\sum_{l=1}^d \alpha_l \lambda_l.$$

Proof: We make the following ansatz for the eigenfunctions of L:

$$\varphi(x) = q(x) f_\infty \in \mathcal{Q},$$

and obtain, using $D + R = CK$ (see section 1.2.1):

$$\begin{aligned}L\varphi &= \operatorname{div}(f_\infty (D+R)\nabla(\tfrac{\varphi}{f_\infty}))\\ &= \operatorname{div}(f_\infty (D+R)\nabla q) = f_\infty \operatorname{div}((D+R)\nabla q) - f_\infty (x^T K^{-1}(D+R)\nabla q)\\ &= f_\infty \left[\operatorname{div}(D\nabla q) - x^T K^{-1} C K \nabla q\right].\end{aligned}$$

This implies that we need to find a $q \in \mathcal{P}(\mathbb{R}^d)$ such that

$$L^{\mathcal{P}} q(x) := \nabla^T D \nabla q(x) - x^T K^{-1} C K \nabla q(x) = \nu q(x).$$

Since the eigenvalues of C (and thus of Q) may be complex, we shall consider the polynomial q in the space $\mathcal{P}(\mathbb{C}^d)$ in the sequel. As in Lemma 1.19, we shall now use the Jordan normal form J of $Q^T = K^{-1}CK$, with $A^{-1}JA = Q^T$ for some regular $A \in \mathbb{C}^{d\times d}$.
We introduce the (complex) coordinate transformation

$$\begin{aligned}y &:= (A^{-1})^T x, \text{ with } y \in \mathbb{C}^d,\\ p(y) &:= q(A^T y) = q(x) \in \mathcal{P}(\mathbb{C}^d).\end{aligned}$$

It follows that

$$\partial_{x_j} q = (\partial_{y_l} p)\frac{\partial y_l}{\partial x_j} = [(A^{-1})^T]_{lj}\partial_{y_l} p,$$

and thus

$$\nabla_x q = A^{-1}\nabla_y p.$$

From this we have

$$x^T Q^T \nabla_x q = y^T A Q^T A^{-1}\nabla_y p = y^T J \nabla_y p$$

and

$$\nabla_x^T D \nabla_x q = \nabla_y^T (A^{-1})^T D A^{-1} \nabla_y p.$$

So we obtain the following equation for the (transformed) eigenfunctions of $L^{\mathcal{P}}$:

$$\tilde{L}^{\mathcal{P}} p(y) := \nabla_y^T (A^{-1})^T D A^{-1} \nabla_y p(y) - y^T J \nabla_y p(y) = \nu p(y). \tag{1.63}$$

A basis of the polynomials (over $\mathbb{C}$) of degree n or lower is given by the monomials $\{y^\alpha | \alpha \in \mathbb{N}_0^d,\ |\alpha| \le n\}$. We order this basis by increasing degree, and in decreasing lexicographic order for monomials of the same degree. Next, we compute the matrix representation $M_{\mathcal{P}}$ of $\tilde{L}^{\mathcal{P}}$ with respect to this basis. Let e_l denote the l-th unit vector in $\mathbb{C}^d$, and I_{def} be the set of all $l \in \{1, \dots, d\}$ for which e_l is not an ordinary eigenvector of J. We compute

$$\begin{aligned}
\tilde{L}^{\mathcal{P}} y^\alpha &= [\nabla^T (A^{-1})^T D A^{-1} - y^T J] \sum_{l=1}^{d} (\alpha_l e_l y^{\alpha_{l-}}) \\
&= \sum_{l,m=1}^{d} \left([\alpha_m - \delta_{lm}]\alpha_l e_m^T (A^{-1})^T D A^{-1} e_l y^{(\alpha_{l-})_{m-}} \right) \\
&\quad - \sum_{l=1}^{d} (\alpha_l \lambda_l y^\alpha) - \sum_{l \in I_{def}} \alpha_l y^T e_{l-1} y^{\alpha_{l-}} \\
&= \sum_{l,m=1}^{d} \left(d_{lm}(\alpha) y^{(\alpha_{l-})_{m-}} \right) + \nu_\alpha y^\alpha - \sum_{l \in I_{def}} \alpha_l y^{(\alpha_{l-})_{(l-1)+}},
\end{aligned} \tag{1.64}$$

where $d_{lm}(\alpha) := [\alpha_m - \delta_{lm}]\alpha_l e_m^T (A^{-1})^T D A^{-1} e_l$. The first term of the r.h.s. has degree $\max(|\alpha| - 2, 0)$. The second and the third term both have degree $|\alpha|$, but the exponents of the third term come "earlier" in lexicographic order. Due to our ordering of the basis $\{y^\alpha |\ |\alpha| \le n\}$, this implies that $M_{\mathcal{P}}$ is an upper

triangular matrix. The entries on the diagonal are just the ν_α, which are hence the eigenvalues of $\tilde{L}^{\mathcal{P}}$ and hence of $L^{\mathcal{P}}$. It follows that, by transforming $M_{\mathcal{P}}$ into Jordan form, one can find a basis of the polynomials of degree n or lower consisting of (generalised) eigenfunctions of $\tilde{L}^{\mathcal{P}}$.
If $\alpha_l = 0$ for all $l \in I_{def}$, then (1.64) only contains terms of lower order besides of $\nu_\alpha y^\alpha$. Hence, y^α will be the leading term of an ordinary eigenfunction of $\tilde{L}^{\mathcal{P}}$.
To define Φ, we observe that for any $\beta \in \mathbb{N}_0^d$, the set $\{y^\alpha \big| \, |\alpha| \le |\beta|\} \bigcap \{y^\alpha \big| \, |\alpha| = |\beta|,\ \alpha \ge \beta \text{ lexicographically}\}$ is still the basis of an $\tilde{L}^{\mathcal{P}}$-invariant subspace U_β of $\mathcal{P}(\mathbb{C}^d)$ – due to the upper triangular form of $M_{\mathcal{P}}$. Further, if α follows β in the order introduced above for multiindices, then U_α has dimension 1 greater than U_β. Hence, U_α contains one additional (generalised) eigenfunction over U_β, which has to include the term y^α (else it would be in U_β). Setting $\Phi(\alpha)$ as this (generalised) eigenfunction completes the proof. □

An immediate consequence of Lemma 1.34 is:

Corollary 1.35. *There are only finitely many eigenfunctions $\varphi \in \mathcal{Q}$ of L to a given eigenvalue μ. The (generalised) eigenfunctions of L from Lemma 1.34 form a basis of $\mathcal{Q}$.*

Proof: Since L has compact resolvent and 0 is an eigenvalue, $\sigma(L) = \sigma_p(L)$. Moreover, the eigenvalues have no accumulation point, and all eigenspaces are finite dimensional. The (generalised) eigenfunctions of L form a basis of $\mathcal{Q}$ since Φ from Lemma 1.34 is a bijection. □

With Lemma 1.34 and Corollary 1.35, we have characterised the spectrum of $L|_{\mathcal{Q}}$. We will now show that this is the same as the spectrum of L in $L^2(\mathbb{R}^d, f_\infty^{-1})$. To do so, we introduce a change of coordinates. Let

$$\begin{aligned} y &:= K^{-\frac{1}{2}} x, \\ g_0(y) &:= f_\infty(K^{\frac{1}{2}} y) = c_K \exp(-\frac{|y|^2}{2}). \end{aligned}$$

Now let

$$\begin{aligned} g_\alpha(y) &:= \nabla^\alpha g_0(y), \quad \alpha \in \mathbb{N}_0^d, \\ V_m &:= \operatorname{span}\left\{g_\alpha \big| \, |\alpha| = m\right\} \subset \tilde{\mathcal{Q}} := \mathcal{P}(\mathbb{R}^d) g_0, \quad m \in \mathbb{N}_0 . \end{aligned}$$

Note that the g_α are in $\tilde{\mathcal{Q}}$, and the polynomial part of g_α has degree $|\alpha|$. From [39], [57] we know that $\{g_\alpha\}_{\alpha \in \mathbb{N}_0^d}$ forms an orthogonal basis of $L^2(\mathbb{R}^d, g_0^{-1})$. Hence, the subspaces V_m are also mutually orthogonal.

Acting on the transformed function $g(y) := f(K^{\frac{1}{2}}y) \in L^2(\mathbb{R}^d, g_0^{-1})$, L has the form

$$\begin{aligned}
\tilde{L}g &:= \operatorname{div}[(\tilde{D}+\tilde{R})(\nabla g + yg)], \\
\tilde{D} &:= K^{-\frac{1}{2}}DK^{-\frac{1}{2}}, \\
\tilde{R} &:= K^{-\frac{1}{2}}RK^{-\frac{1}{2}}.
\end{aligned}$$

Lemma 1.36. *For every $m \in \mathbb{N}_0$, V_m is invariant under both $\tilde{L}$ and its adjoint $\tilde{L}^\dagger$ (w.r.t. $L^2(\mathbb{R}^d, g_0^{-1})$).*

Proof: Note that the following properties of D, R, and C also hold for the transformed versions (with $\tilde{C} := K^{-\frac{1}{2}}CK^{-\frac{1}{2}}$):

$$2\tilde{D} = \tilde{C}K + K\tilde{C}^T, \quad \tilde{R}^T = -\tilde{R}.$$

The adjoint of $\tilde{L}$ has the form

$$\tilde{L}^\dagger g = \operatorname{div}[(\tilde{D}-\tilde{R})(\nabla g + yg)].$$

Now compute

$$\begin{aligned}
\partial_l g_\alpha(y) &= \nabla^\alpha \partial_l g_0(y) = -\nabla^\alpha(y_l g_0(y)) \\
&= -\alpha_l g_{\alpha_{l-}}(y) - y_l g_\alpha(y).
\end{aligned}$$

So we have, writing $h_\alpha := (\alpha_l g_{\alpha_{l-}}(y))_{l=1,\dots,d}$,

$$\nabla g_\alpha(y) = -h_\alpha(y) - y g_\alpha(y).$$

Inserting this into $\tilde{L}$ gives

$$\begin{aligned}
\tilde{L}g_\alpha &= \operatorname{div}[(\tilde{D}+\tilde{R})(-h_\alpha(y) - yg_\alpha(y) + yg_\alpha(y))] \\
&= -\operatorname{div}(\tilde{D}h_\alpha(y)) - \operatorname{div}(\tilde{R}h_\alpha(y)), \\
\tilde{L}^\dagger g_\alpha &= -\operatorname{div}(\tilde{D}h_\alpha(y)) + \operatorname{div}(\tilde{R}h_\alpha(y)).
\end{aligned}$$

Compute further

$$\begin{aligned}
\operatorname{div}(\tilde{D}h_\alpha) &= \sum_{j,l=1}^{d} \partial_j \left(\tilde{D}_{jl}\alpha_l g_{\alpha_{l-}}\right)(y) = \sum_{j,l=1}^{d} \alpha_l \tilde{D}_{jl} g_{(\alpha_{l-})_{j+}}(y), \\
\operatorname{div}(\tilde{R}h_\alpha) &= \sum_{j,l=1}^{d} \alpha_l \tilde{R}_{jl} g_{(\alpha_{l-})_{j+}}(y)\,.
\end{aligned}$$

Thus we get, using $R = \frac{1}{2}(CK - KC^T)$ and $D = \frac{1}{2}(CK + KC^T)$,

$$\begin{aligned}
\tilde{L} g_\alpha &= -\sum_{j,l=1}^{d} \alpha_l (\tilde{D} + \tilde{R})_{jl} g_{(\alpha_{l-})_{j+}}(y) \\
&= -\sum_{j,l=1}^{d} \alpha_l (K^{-\frac{1}{2}} C K^{\frac{1}{2}})_{jl} g_{(\alpha_{l-})_{j+}}(y), \\
\tilde{L}^\dagger g_\alpha &= -\sum_{j,l=1}^{d} \alpha_l (K^{\frac{1}{2}} C^T K^{-\frac{1}{2}})_{jl} g_{(\alpha_{l-})_{j+}}(y).
\end{aligned}$$

We see that $\tilde{L} g_\alpha$, $\tilde{L}^\dagger g_\alpha$ are linear combinations only of terms g_β, $\beta \in \mathbb{N}_0^d$, with $|\beta| = |\alpha|$. This completes the proof. $\square$

We now have the tools to prove Theorem 1.32:

Proof (of Theorem 1.32):

With Lemmas 1.33, 1.34 we have already established that

$$\sigma(L) = \sigma_p(L) \supset \Big\{ -\sum_{j=1}^{d} \alpha_j \lambda_j \Big| \alpha \in \mathbb{N}_0^d \Big\}. \tag{1.65}$$

All that remains to show is that there are no further eigenvalues in the point spectrum of L or, equivalently, of $\tilde{L}$. Assume there were an additional eigenvalue λ of $\tilde{L}$ with $\tilde{L} g = \lambda g$ for some $g \in L^2(\mathbb{R}^d, g_0^{-1})$. Then g has a unique L^2–decomposition in $\{V_m\}$:

$$g = \sum_{m=0}^{\infty} g_m\,, \quad \text{with } g_m \in V_m\,.$$

By the orthogonality and $\tilde{L}$–invariance of the V_m we have $\tilde{L} g_m = \lambda g_m \; \forall\, m \in \mathbb{N}_0$.

By Lemma 1.34, all eigenfunctions of $\tilde{L}$ in V_m satisfy

$$\nu = -\sum_{j=1}^{d} \alpha_j \lambda_j$$

for some $\alpha \in \mathbb{N}_0^d$ with $|\alpha| = m$. Hence, $\Re\{\nu\} \le -\mu m$ with $\mu = \min\limits_{j=1,\dots,d} \Re\{\lambda_j\} > 0$. This implies $g_m = 0$ for all $m > \tilde{m} := \frac{-\Re\{\lambda\}}{\mu}$. Thus, $g = q g_0$, with a polynomial q of deg $q \le \tilde{m}$. Hence, $g \in \tilde{\mathcal{Q}}$, and λ is already included in the r.h.s. of (1.65). $\square$

1.5 Examples

In this section, we study three 2-dimensional toy problems to illustrate our results and some of the mechanisms behind hypocoercive behaviour.
As first model, in Example 1 we discuss a degenerate Fokker-Planck equation with a confinement potential whose gradient lines are not aligned with the dissipative subspace. Hence, they cross this subspace at a non-zero angle, which yields (A.i). (A.ii) is provided by the confinement potential. In this case, C is symmetric and has a full complement of real eigenvectors and eigenvalues, and thus the spectrum of L is real.
In Example 2, we study as a second model the case of a defective C, where the dissipative subspace is a generalised eigenspace of C. Hence the degeneracy subspace of the dissipation is not invariant under C^T, and the model fulfils (A.i). (A.ii) again comes from a confinement potential. The spectrum of L remains real, but, as seen in Theorem 1.27, there is no precise sharp rate due to the defectiveness and generalised eigenfunctions.
As a last model, we discuss in Example 3 a kinetic equation with a skew-symmetric transport between the two variables, but diffusion only in the first variable. Here, (A.i) is due to a "mixing" of space variables by the transport terms, which is also responsible for "extending" the confinement by the potential (which only depends on x_1) to the whole space. Depending on the scaling of the transport term, this example exhibits the same behaviour as the first two or gives rise to two complex conjugate eigenvalues of C. This is the standard kinetic Fokker-Planck-model often used in literature (see e.g. [30], [10], [31]), with $x_1 = v$ and $x_2 = x$.

Example 1. Let L be defined as in (2) with

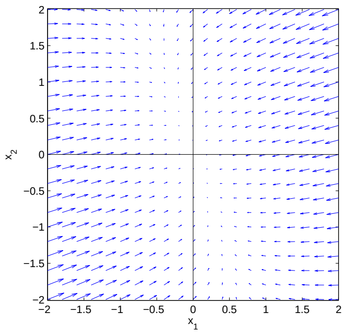
$$D := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad C := \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix}.$$

Then the unique normalised steady state from Theorem 1.12 is

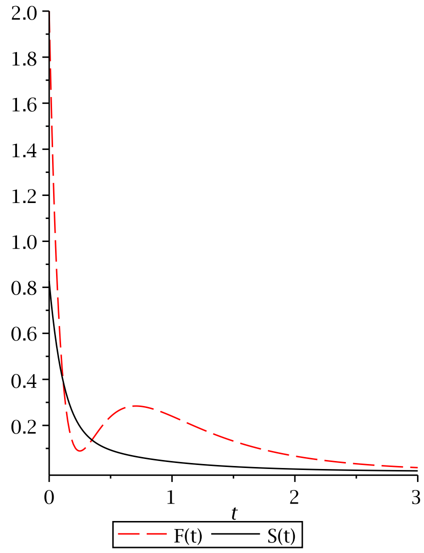
$$f_\infty = \frac{15}{2\pi\sqrt{3}} \exp(-\frac{5x_1^2 + 10x_1x_2 + 20x_2^2}{2}),$$

and the sharp decay rate from Theorem 1.27 is $\mu = \frac{5-\sqrt{13}}{2}$.

Figure 1.1: Example 1: Symmetric drift coefficients from a skew-aligned potential. Figure (b) shows that the correct choice of P is crucial.



(a) Field plot of the drift coefficient $x \mapsto Cx$.



(b) Plot of the modified entropy dissipation S and the functional F obtained by replacing P in S with Id.

Proof: To find the unique normalised steady state, we have to solve (1.12). So let

$$K = \begin{pmatrix} k_1 & k_2 \\ k_2 & k_3 \end{pmatrix}.$$

We obtain the equations

$$\begin{aligned} 2 &= 8k_1 + 2k_2, \\ 0 &= k_1 + 5k_2 + k_3, \\ 0 &= 2k_2 + 2k_3, \end{aligned}$$

which have the solution

$$K = \frac{1}{15} \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix}, \quad K^{-1} = 5 \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}.$$

This means the normalised steady state has the form

$$f_\infty = \frac{15}{2\pi\sqrt{3}} \exp(-\frac{5x_1^2 + 10x_1x_2 + 20x_2^2}{2}).$$

For the rate of decay, we compute the eigenvalues of C as $\lambda_\pm = \frac{5\pm\sqrt{13}}{2}$. From Theorem 1.27, we thus know that $\mu = \frac{5-\sqrt{13}}{2}$ is the sharp convergence rate. The eigenvectors of KC^TK^{-1} to $\lambda_\pm$ are

$$v_+ = \begin{pmatrix} 1 \\ -\frac{7-\sqrt{13}}{18} \end{pmatrix},$$
$$v_- = \begin{pmatrix} 1 \\ -\frac{7+\sqrt{13}}{18} \end{pmatrix}.$$

From this, we can compute P as

$$P = v_+v_+^T + v_-v_-^T = \begin{pmatrix} 2 & -\frac{7}{9} \\ -\frac{7}{9} & \frac{31}{81} \end{pmatrix}.$$

The behaviour of S for the initial condition

$$f_0 = (x_1 + x_2 + 1)f_\infty$$

(see also Lemma 1.30) is shown in Figure 1.1b. Also shown is the behaviour of the functional F one obtains by replacing D in I_ψ (see (1.18)) with Id. This is analogous to $P = \mathrm{Id}$ in the definition of S (1.19) and retains information on all derivatives. As can be seen from the graph, there is no hope of obtaining a decay estimate on F since the functional is not monotonous. This shows that one has to be careful in choosing P. □

Example 2. Let L be defined as in (2) with

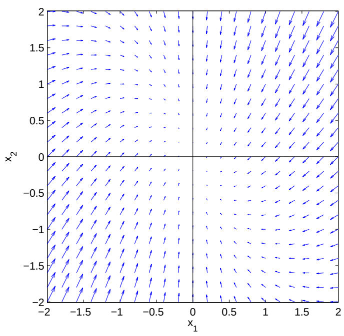
$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then the unique normalised steady state from Theorem 1.12 is

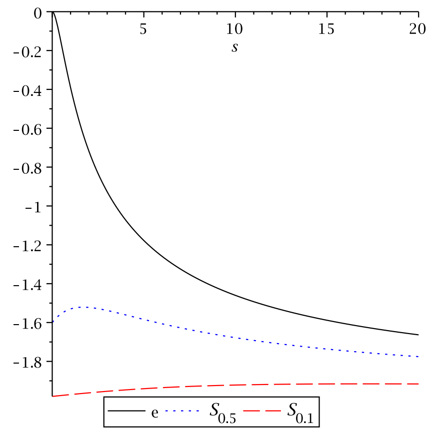
$$f_\infty = \frac{1}{\pi}\exp(-\frac{2x_1^2 + 4x_1x_2 + 4x_2^2}{2}),$$

and the decay rate from Theorem 1.27 is $\mu = 1 - \varepsilon$ for any $\varepsilon \in (0, 1)$.

Figure 1.2: Example 2: Defective matrix for the drift coefficient.



(a) Field plot of the drift coefficient $x \mapsto Cx$ for Example 2.



(b) Comparison of effective decay rates for e_ψ and S_ε, with $\varepsilon = 0.5$ and $\varepsilon = 0.1$.

Proof: As equations for

$$K = \begin{pmatrix} k_1 & k_2 \\ k_2 & k_3 \end{pmatrix}$$

we obtain from (1.12):

$$\begin{aligned} 2 &= 2k_1, \\ 0 &= k_1 + 2k_2, \\ 0 &= 2k_2 + 2k_3. \end{aligned}$$

This has the solution

$$K = \frac{1}{2}\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}, \quad K^{-1} = 2\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

so the normalised stationary state is

$$f_\infty = \frac{1}{\pi}\exp(-\frac{2x_1^2 + 4x_1x_2 + 4x_2^2}{2}).$$

C has the defective eigenvalue 1, so there is no exact sharp rate, and the best we can achieve is $1 - \varepsilon$ for $\varepsilon \in (0, 1)$, see Theorem 1.27. To obtain P, we compute the eigenvector v and generalised eigenvector h of $Q = KC^TK^{-1}$ as

$$v = \begin{pmatrix} -2 \\ 1 \end{pmatrix},$$

$$h = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

We then set, for $\varepsilon \in (0,1)$,

$$P_\varepsilon := vv^T + \varepsilon^2 hh^T.$$

One easily verifies that

$$QP_\varepsilon + P_\varepsilon Q^T \geq 2(1-\varepsilon)P_\varepsilon.$$

For the initial condition

$$f_0 = (x_2 + 1) f_\infty,$$

one can expect the defectiveness of the eigenvalue to show in the decay rate (compare Theorem 1.28 (iii)).
The effective decay rate for the entropy e on the interval $[0,s]$ can be computed as

$$\lambda_{eff}(s) := \frac{1}{s} \log \frac{e(s)}{e(0)}.$$

In Figure 1.2b, the effective decay rate for e is shown alongside the effective decay rate for two modified entropy dissipation functionals S corresponding to P_ε for $\varepsilon = 0.5$ and $\varepsilon = 0.1$. The estimated rate for S_ε is $2(1-\varepsilon)$. As can be seen, the estimated decay rate is not optimal (this could also be seen from the fact that the "remainder" M_n in the proof of Lemma 1.19 is positive definite, not positive semidefinite), as the depicted rates are better than the estimate. For small s the effective decay rate is better for modified entropy dissipation functionals, improving as $\varepsilon \to 0$. However, this comes at the price of increasing the constant c_ε in the estimate from Theorem 1.27.
$\square$

Example 3. Let $0 \neq \nu \in \mathbb{R}$ and L be defined as in (2) with

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -\nu \\ \nu & 0 \end{pmatrix}.$$

The unique normalised steady state from Theorem 1.12 is

$$f_\infty = \frac{1}{2\pi} \exp(-\frac{|x|^2}{2}),$$

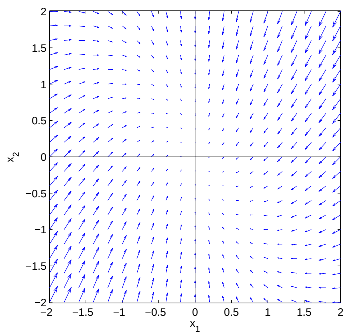
and the decay rate from Theorem (1.27) depends on ν in the following way:

- If $|\nu| < \frac{1}{2}$, then $\mu = \frac{1-\sqrt{1-4\nu^2}}{2}$.

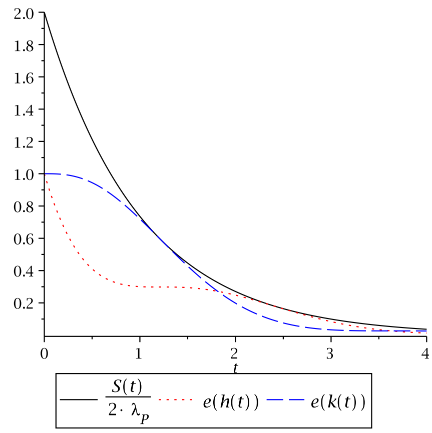
- If $|\nu| = \frac{1}{2}$, then the eigenvalue with smallest real part is defective, and we can obtain the rate $\mu = \frac{1}{2} - \varepsilon$ for $\varepsilon \in (0, \frac{1}{2})$.

- If $|\nu| > \frac{1}{2}$, then $\mu = \frac{1}{2}$.

Figure 1.3: Example 3: Rotation and a one-dimensional Fokker-Planck operator.



(a) Field plot of the drift coefficient $x \mapsto Cx$ for Example 3, with $\nu = 1$.



(b) Plot of the relative entropy for initial conditions k_0, h_0, as well as the bound from the log-Sobolev inequality (1.29)

Proof: As equations for

$$K = \begin{pmatrix} k_1 & k_2 \\ k_2 & k_3 \end{pmatrix}$$

we get from (1.12):

$$2 = 2k_1 - 2\nu k_2$$
$$0 = k_2 + \nu(k_1 - k_3)$$
$$0 = 2\nu k_2$$

The solution is $K = \mathrm{Id}$, and we obtain the steady state

$$f_\infty = \exp(-\frac{|x|^2}{2}).$$

The eigenvalues of $K^{-1}CK = C$ are

$$\lambda_+ = \frac{1+\sqrt{1-4\nu^2}}{2}, \quad \lambda_- = \frac{1-\sqrt{1-4\nu^2}}{2}.$$

If

$$|\nu| = \tfrac{1}{2},$$

both eigenvalues are $\frac{1}{2}$, but there is only one eigenvector

$$v = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

so the eigenvalue $\frac{1}{2}$ is defective. We thus get three cases:

- If $0 < |\nu| < \frac{1}{2}$, both eigenvalues are real; we are in the same situation as the first example, and the rate is $\mu = \frac{1-\sqrt{1-4\nu^2}}{2} < \frac{1}{2}$.

- If $|\nu| = \frac{1}{2}$, there is a single, real defect eigenvalue, so the situation is the same as in the second example, with the rate $\frac{1}{2} - \varepsilon$ for $\varepsilon \in (0, \frac{1}{2})$.

- If $|\nu| > \frac{1}{2}$, both eigenvalues are complex, and the from Theorem 1.27 is $\mu = \frac{1}{2}$.

We compute P for the case $\nu = 1$. In this case, $\lambda := \frac{1}{2} \pm i\frac{\sqrt{3}}{2}$, and $KC^TK^{-1} = C^T$ has the eigenvector

$$v := \begin{pmatrix} -\lambda \\ 1 \end{pmatrix}$$

to the eigenvalue λ, and $\overline{v}$ to $\overline{\lambda}$. Thus, we can set

$$P := vv^H + \overline{v}\overline{v}^H = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

In this case, the constant λ_P from (1.29) is 1. We compare the logarithmic entropy for the two initial conditions

$$h_0 := (x_1 + 1)f_\infty, \quad k_0 := (x_2 + 1)f_\infty.$$

k_0 is in the kernel of L_s, and $h_0 - f_\infty$ in its complement. Shown in Fig. 1.3b is the relative quadratic entropy for both cases, as well as the bound $\frac{S}{2\lambda_P}$ from (1.29) with the modified entropy dissipation S. In this case, the bound is sharp. However, this bound requires the initial condition f_0 to have finite modified entropy dissipation S, which in general will not be true. The general bound from Theorem 1.27 is strictly greater. □

1.6 Another proof of hypocoercivity

In this section, we establish exponential convergence towards the unique normalised steady state (see Theorem 1.12) in a weighted H^1-norm under condition (A). We follow a method established by Villani in [67], reformulating his Theorem 24 in a version tailored to our specific problem. As in [67], we are here not interested in a sharp decay rate, which we have already established in §1.3. Instead, the idea is to present another approach to the problem.

Before we start, we briefly elaborate on the strategy and idea behind the proof. The symmetric part of L can be written as A^*A. Due to the semidefiniteness and singularity of D, this is coercive on $\ker A = \ker A^*A$, which is strictly larger than $\ker L$. In his proof of hypoellipticity, Hörmander [42] uses iterated commutators $F_1 := [A, B]$, $F_2 := [[A, B], B]$, ..., where $B = -B^*$ is the antisymmetric part of L. Associate with A, F_j the vector fields $a(x)$, $f_j(x)$ of their coefficients (i.e, replacing ∂_j with the unit vectors e_j). The hypoellipticity then follows if $a(x)$, $f_1(x)$, $f_2(x)$, ... span the whole of $\mathbb{R}^d$ for all $x \in \mathbb{R}^d$. The idea is now to consider the operators A^*A, $F_1^*F_1$, $F_2^*F_2$, ..., all of which are coercive only on "too small" a subset of $\mathcal{H} = \{f \in L^2 | \nabla(\frac{f}{f_\infty}) \in (L^2(\mathbb{R}^d, f_\infty))^d\}$. However, if their sum is coercive on $\ker L^\perp$, they can then be used to construct a Lyapunov functional that yields decay in a $\mathcal{H}^1$-norm. As it turns out, this is the case exactly iff condition (A) is fulfilled.

We recall that, under condition (A), L decomposes on the weighted space L^2 as

$L_s + L_{as}$,

$$L_s f = \operatorname{div}(D\nabla(\tfrac{f}{f_\infty})f_\infty),$$
$$L_{as} f = x^T T \nabla f =: Bf,$$

where

$$T := \tfrac{1}{2}(C^T - K^{-1}CK) \in \mathbb{R}^{d\times d} \tag{1.66}$$

with $\mathrm{Tr}(T) = 0$ (see Theorem 1.16).
We work in the weighted H^1-space $\mathcal{H} = \{f \in L^2 | \nabla(\frac{f}{f_\infty}) \in (L^2(\mathbb{R}^d, f_\infty))^d\}$, $\|f\|^2_{\mathcal{H}} = \int\limits_{\mathbb{R}^d} |f|^2 f_\infty^{-1}\, \mathrm{d}x + \int\limits_{\mathbb{R}^d} |\nabla \frac{f}{f_\infty}|^2 f_\infty\, \mathrm{d}x$ from Lemma 1.33. Let $A : \mathcal{H} \subset L^2 \to (L^2)^d, \quad f \mapsto D^{\frac{1}{2}} \nabla(\frac{f}{f_\infty}) f_\infty$. Then we get for $g \in (H^1(\mathbb{R}^d, f_\infty^{-1}))^d$:

$$\langle Af, g\rangle = \int\limits_{\mathbb{R}^d} D^{\frac{1}{2}} \nabla(\tfrac{f}{f_\infty}) f_\infty g f_\infty^{-1}\, \mathrm{d}x$$
$$= -\int\limits_{\mathbb{R}^d} \operatorname{div}(D^{\frac{1}{2}} g) f f_\infty^{-1}\, \mathrm{d}x.$$

Hence, $A^* g = -\operatorname{div}(D^{\frac{1}{2}} g)$ and $L = -A^*A + B$, with the Operator $B := x^T T \nabla$ as introduced in (1.17). Note that $\|Af\| \leq \tilde{c}_1 \|f\|_{\mathcal{H}}$.

Now, we can state the main result of this section: The operator L is hypocoercive in the sense of [67] (see also Definition 0.1).

Theorem 1.37. *Let condition (A) hold. Then there exist $\lambda > 0$ and $C > 1$ such that all solutions of (2) with $f_0 \in \mathcal{H}$ satisfy*

$$\forall t \geq 0 : \|f(t) - f_\infty\|_{\mathcal{H}} \leq C \exp(-\lambda t) \|f_0 - f_\infty\|_{\mathcal{H}}.$$

Before we begin the proof of Theorem 1.37, we repeat the notations for commutators used in [67]. In this paper, the domain of all operators contains C_0^∞, and thus the commutators can always be defined on C_0^∞ and then be closed. For a proof of the density of C_0^∞ in L^2, see [50], Theorem 8.1.26.

For two operators $X : D(X) \subset L^2 \to L^2$, $Y : D(Y) \subset L^2 \to L^2$, the commutator $[X, Y]$ is defined as the closure of

$$C_0^\infty \subset L^2 \to L^2,$$
$$f \mapsto XYf - YXf.$$

If $X : D(X) \subset (L^2)^d \to L^2$, $\;g = (g_j)_{1\le j\le d} \mapsto \sum_{j=1}^{d} X_j g_j$ and $Y : D(Y) \subset L^2 \to L^2$, we define $[X,Y]$ as the closure of

$$\begin{aligned}&(C_0^\infty)^d \subset (L^2)^d \to L^2,\\ &g = (g_j)_{1\le j\le d} \mapsto \sum_{j=1}^{d} [X_j, Y] g_j \in L^2.\end{aligned}$$

Similarly, if $X : D(X) \subset L^2 \to (L^2)^d$, $g \mapsto (X_j g)_{1\le j\le d}$ and $Y : D(Y) \subset L^2 \to L^2$, we define $[X,Y]$ as the closure of

$$\begin{aligned}&C_0^\infty \subset L^2 \to (L^2)^d,\\ &\qquad g \mapsto ([X_j, Y]g)_{1\le j\le d} \in (L^2)^d.\end{aligned}$$

Finally, if $X : D(X) \subset L^2 \to (L^2)^d$, $g \mapsto (X_j g)_{1\le j\le d}$ and either $Y : D(Y) \subset L^2 \to (L^2)^d$, $g \mapsto (Y_j g)_{1\le j\le d}$ or $Y : D(Y) \subset (L^2)^d \to L^2$, $g = (g_j)_{1\le j\le d} \mapsto \sum_{j=1}^{d} Y_j g_j$, we define $[X,Y]$ as the closure of

$$\begin{aligned}&C_0^\infty \subset L^2 \to (L^2)^{d\times d}\\ &\qquad g \mapsto ([X_j, Y_k]g)_{1\le j,k\le d} \in (L^2)^{d\times d}.\end{aligned}$$

In the next lemma, we compute the iterated commutators for the Lyapunov functional.

Lemma 1.38. *Let $F_0 := A$, and for $j \in \mathbb{N}_0$ let $F_{j+1} := [F_j, B]$. Then*

$$\begin{aligned}&F_j : \mathcal{H} \subset L^2 \to (L^2)^d,\\ &\qquad f \mapsto D^{\frac{1}{2}} T^j \nabla((\frac{f}{f_\infty}) f_\infty.\end{aligned} \tag{1.67}$$

Proof: We proceed by induction. For $j = 0$, $F_0 = A$ has the postulated form. So let $j \ge 0$, and $M := D^{\frac{1}{2}} T^j$. Then

$$\begin{aligned}F_j f &= M\nabla(\tfrac{f}{f_\infty}) f_\infty = \left(\sum_{p=1}^{d} M_{lp}(f_{,p} + (K^{-1}x)_p f)\right)_{l=1,\dots,d},\\ Bf &= x^T T \nabla f = \sum_{r,s=1}^{d} x_r T_{rs} f_{,s}.\end{aligned}$$

This implies

$$\begin{aligned}
B(F_j)_l f &= \sum_{r,s,p=1}^{d} x_r T_{rs} M_{lp} (f_{,p} + (K^{-1}x)_p f)_{,s} \\
&= \sum_{r,s,p=1}^{d} x_r T_{rs} M_{lp} (f_{,ps} + K^{-1}_{ps} f + (K^{-1}x)_p f_{,s}), \\
(F_j)_l B f &= \sum_{r,s,p=1}^{d} M_{lp} T_{rs} \left((x_r f_{,s})_{,p} + (K^{-1}x)_p x_r f_{,s} \right) \\
&= \sum_{r,s,p=1}^{d} M_{lp} T_{rs} \left(\delta_{rp} f_{,s} + x_r f_{,sp} + (K^{-1}x)_p x_r f_{,s} \right),
\end{aligned}$$

and thus

$$\begin{aligned}
[(F_j)_l, B] f &= \sum_{r,s,p=1}^{d} M_{lp} T_{rs} \left(\delta_{rp} f_{,s} - x_r K^{-1}_{ps} f \right), \\
F_{j+1} f &= MT\nabla f - MK^{-1}T^T x f = M(T\nabla f - K^{-1}T^T K(K^{-1}x) f) \\
&= MT(\nabla f + K^{-1}xf) = MT\nabla(\tfrac{f}{f_\infty}) f_\infty,
\end{aligned}$$

where we have used the form of T from Theorem 1.16. □

Corollary 1.39. *For $j \in \mathbb{N}_0$ it holds:*

$$F_j^* = -\operatorname{div}((T^T)^j D^{\frac{1}{2}} \cdot), \tag{1.68}$$

$$F_j f_\infty = 0, \tag{1.69}$$

$$\forall f \in \mathcal{H} : \|F_j f\|^2 \le c_j \|f\|^2_{\mathcal{H}} \textit{ with } c_j := \|D\| \|T\|^{2j} > 0. \tag{1.70}$$

Proof: Using (1.67), an integration by parts immediately yields (1.68). (1.69) follows from $Af_\infty = Bf_\infty = 0$, and (1.70) from Lemma 1.38. □

As mentioned, the proof of Theorem 1.37 relies on analysing a Lyapunov functional that incorporates the iterated commutators F_j. The idea is to "complete" the degenerate diffusion term A^*A in L by adding more diffusion operators of the form $F_j^* F_j$. The next lemma establishes that a finite sum of these operators will always be sufficient for the diffusion to act on the whole of $\mathbb{R}^d$:

Lemma 1.40. *Let T be defined as above. Let condition (A) be fulfilled. Then there exists $\kappa > 0$ such that*

$$\sum_{j=0}^{d-k} (T^T)^j D T^j \geq \kappa \, \mathrm{Id},$$

where $k = \operatorname{rank} D$.

Proof: This is a direct consequence of Lemma 1.3, replacing C^T by $T = \frac{1}{2}(C^T - K^{-1}CK) = C^T - K^{-1}D$ and noting that $Tv = C^T v$ for $v \in \ker D$. $\square$

Corollary 1.41. *It holds that $\sum\limits_{j=0}^{d-k} \|F_j f\|^2 \geq \gamma \|f - f_\infty\|_{\mathcal{H}}^2$ for some $\gamma > 0$.*

Proof: Lemma 1.40 gives

$$\begin{aligned}
\sum_{j=0}^{d-k} \|F_j f\|^2 &= \sum_{j=0}^{d-k} \langle f, F_j^* F_j f \rangle \\
&= -\langle f, \operatorname{div}(\sum_{j=0}^{d-k} (T^T)^j D T^j \nabla (\frac{f}{f_\infty}) f_\infty) \rangle \\
&= \langle \nabla (\frac{f}{f_\infty}) f_\infty, \sum_{j=0}^{d-k} (T^T)^j D T^j \nabla (\frac{f}{f_\infty}) f_\infty \rangle \\
&\geq \gamma' \int\limits_{\mathbb{R}^d} |\nabla \frac{f}{f_\infty}|^2 f_\infty \, \mathrm{d}x \\
&\geq \frac{\gamma' \lambda_1}{1 + \lambda_1} \left(\int\limits_{\mathbb{R}^d} |\nabla \frac{f}{f_\infty}|^2 f_\infty \, \mathrm{d}x + \|f - f_\infty\|^2 \right),
\end{aligned}$$

where we used Lemma 1.20. $\square$

Finally, we need an estimate on two more commutators that will appear in our Lyapunov functional:

Lemma 1.42. *The operators F_j defined in (1.67) satisfy*

$$[F_j, A] = 0, \tag{1.71}$$

$$\|[F_j, A^*] f\| \leq \alpha_j \|f\|, \tag{1.72}$$

for some $\alpha_j > 0$.

Proof: Remember our notational convention: $[F_j, A]$ is the closure of an operator from C_0^∞ to $(L^2)^{d\times d}$, as is $[F_j, A^*]$.
Now let $j \in \mathbb{N}_0$ be fixed. Since $F_j f = M\nabla(\frac{f}{f_\infty})f_\infty$ for some matrix M, we get for the components of our operators, using $D^{\frac{1}{2}} = D$,

$$\begin{aligned}
(F_j)_l f &= \sum_{p=1}^{d} M_{lp} f_{,p} + (MK^{-1})_{lp} x_p f, \\
A_r f &= \sum_{s=1}^{d} D_{rs} f_{,s} + (DK^{-1})_{rs} x_s f, \\
A_r^* f &= -\sum_{s=1}^{d} D_{rs} f_{,s}.
\end{aligned}$$

Denote by m_l the l-th row of M, by d_r the r-th row of D. Then we have for $f \in C_0^\infty$:

$$(F_j)_l A_r = m_l \cdot \nabla(f_\infty^{-1} d_r \cdot \nabla \frac{f}{f_\infty}) f_\infty = (m_l)^T [\tfrac{\partial^2}{\partial x^2} \frac{f}{f_\infty}] d_r f_\infty.$$

Since the matrix $\frac{\partial^2}{\partial x^2} \frac{f}{f_\infty}$ is symmetric, it follows that $[(F_j)_l, A_r] = 0$. Similarly, we have

$$\begin{aligned}
[(F_j)_l, A_r^*] f &= -m_l \cdot \nabla(\frac{d_r \cdot \nabla f}{f_\infty}) f_\infty + d_r \cdot \nabla(m_l \cdot \nabla(\frac{f}{f_\infty}) f_\infty \\
&= -\mathrm{Tr}(d_r^T \otimes m_l \cdot \nabla \otimes (\frac{\nabla f}{f_\infty}) f_\infty + \mathrm{Tr}(m_l^T \otimes d_r \cdot \nabla \otimes (\nabla(\frac{f}{f_\infty}) f_\infty)) \\
&= -\mathrm{Tr}(d_r^T \otimes m_l \cdot \nabla \otimes \frac{\nabla f_\infty}{f_\infty}) f
\end{aligned}$$

which proves the inequality

$$\|[F_j, A^*] f\| \leq c(M, K, D)\|f\|.$$

$\square$

With these results, we can prove convergence to equilibrium under condition (A):

Proof: (of Theorem 1.37)

We reiterate the proof of Theorems 18 and 24 in [67] tailored to (2). Remember that $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ refer to the norm and scalar product on the weighted L^2-space. Let $N := d-k$. For a solution f to (2) let

$$\mathcal{L}(f) := c\|f-f_\infty\|^2 + \sum_{j=0}^{N} a_j\|F_j f\|^2 - 2\sum_{j=0}^{N-1} b_j\langle F_j f, F_{j+1} f\rangle,$$

where the positive constants a_j, b_j, c are yet to be determined. We will show that $\mathcal{L}$ is a Lyapunov functional. First, we shall need

$$\mathcal{L}(f) \geq \tilde{c}_2\|f-f_\infty\|_{\mathcal{H}}^2, \tag{1.73}$$
$$\mathcal{L}(f) \leq \tilde{c}_3\|f-f_\infty\|_{\mathcal{H}}^2 \tag{1.74}$$

with some $\tilde{c}_2, \tilde{c}_3 > 0$ still to be chosen. For (1.73), we can choose

$$a_j a_{j+1} > 4b_j^2 \tag{1.75}$$

and obtain:

$$\begin{aligned}
&\sum_{j=0}^{N} a_j\|F_j f\|^2 - 2\sum_{j=0}^{N-1} b_j\langle F_j f, F_{j+1} f\rangle \\
&= \frac{a_0}{2}\|F_0 f\|^2 + \frac{a_N}{2}\|F_N f\|^2 + \sum_{j=0}^{N-1} \frac{a_j}{2}\|F_j f\|^2 - 2b_j\langle F_j f, F_{j+1} f\rangle + \frac{a_{j+1}}{2}\|F_{j+1} f\|^2 \\
&\geq \tilde{c}_4\sum_{j=0}^{d-k}\|F_j f\|^2 \geq \tilde{c}_2\|f-f_\infty\|_{\mathcal{H}}^2,
\end{aligned}$$

where we have used Corollary 1.41 for the last inequality.

With (1.69) we get

$$\begin{aligned}
\mathcal{L}(f) &= c\|f-f_\infty\|^2 + \sum_{j=0}^{N} a_j\|F_j(f-f_\infty)\|^2 - 2\sum_{j=0}^{N-1} b_j\langle F_j(f-f_\infty), F_{j+1}(f-f_\infty)\rangle \\
&\leq c\|f-f_\infty\|^2 + \tilde{c}_5\sum_{j=0}^{N}\|F_j(f-f_\infty)\|^2 \leq \tilde{c}_3\|f-f_\infty\|_{\mathcal{H}}^2,
\end{aligned}$$

where we have used (1.70) for the last inequality.

Our aim is now to prove that $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}(f) \leq -\sum_{j=0}^{N}\gamma_j\|F_j f\|^2$ for some $\gamma_j > 0$. This will give the desired result with Corollary 1.41, since $\mathcal{L}$ is equivalent to

$\|\cdot\|_{\mathcal{H}}^2$. So compute, using $\int_{\mathbb{R}^d} f_t \, \mathrm{d}x = 0$ and (1.71):

$$
\begin{aligned}
&\tfrac{\mathrm{d}}{\mathrm{dt}}\|f - f_\infty\|^2 = 2\langle f, f_t\rangle = -2\|Af\|^2,\\
&\tfrac{\mathrm{d}}{\mathrm{dt}}\|F_j f\|^2 = 2\langle F_j f, F_j f_t\rangle = -2\langle F_j f, F_j A^* A f\rangle + 2\langle F_j f, F_j B f\rangle\\
&= -2\langle F_j f, [F_j, A^*]Af\rangle - 2\langle AF_j f, F_j Af\rangle + 2\langle F_j f, [F_j, B]f\rangle + 2\langle F_j f, BF_j f\rangle\\
&= -2\langle F_j f, [F_j, A^*]Af\rangle - 2\|AF_j f\|^2 + 2\langle F_j f, F_{j+1} f\rangle,\\
&\tfrac{\mathrm{d}}{\mathrm{dt}}\langle F_j f, F_{j+1} f\rangle\\
&= \langle F_j f_t, F_{j+1} f\rangle + \langle F_j f, F_{j+1} f_t\rangle\\
&= -\langle F_j A^* A f, F_{j+1} f\rangle + \langle F_j B f, F_{j+1} f\rangle - \langle F_j f, F_{j+1} A^* A f\rangle + \langle F_j f, F_{j+1} B f\rangle\\
&= -\langle [F_j, A^*]Af, F_{j+1} f\rangle - \langle F_j Af, AF_{j+1} f\rangle + \langle F_j B f, F_{j+1} f\rangle - \langle F_j f, [F_{j+1}, A^*]Af\rangle\\
&\quad - \langle AF_j f, F_{j+1} Af\rangle + \langle F_j f, [F_{j+1}, B]f\rangle + \langle F_j f, BF_{j+1} f\rangle\\
&= -\langle [F_j, A^*]Af, F_{j+1} f\rangle - 2\langle AF_j f, AF_{j+1} f\rangle - \langle F_j f, [F_{j+1}, A^*]Af\rangle\\
&\quad + \|F_{j+1} f\|^2 + \langle F_j f, F_{j+2} f\rangle.
\end{aligned}
$$

Using this we get

$$
\begin{aligned}
&\tfrac{1}{2}\tfrac{\mathrm{d}}{\mathrm{dt}}\mathcal{L}(f)\\
&= -c\|Af\|^2 - \sum_{j=0}^{N} a_j \left(\|AF_j f\|^2 - \langle F_j f, F_{j+1} f\rangle - \langle F_j f, [F_j, A^*]Af\rangle\right) \qquad (1.76)\\
&\quad + \sum_{j=0}^{N-1} b_j \Big(\langle [F_j, A^*]Af, F_{j+1} f\rangle + 2\langle AF_j f, AF_{j+1} f\rangle + \langle F_j f, [F_{j+1}, A^*]Af\rangle\Big)\\
&\quad - \sum_{j=0}^{N-1} b_j \Big(\|F_{j+1} f\|^2 - \langle F_j f, F_{j+2} f\rangle\Big). \qquad (1.77)
\end{aligned}
$$

Note that

$$
0 \geq -\sum_{j=0}^{N} a_j \|AF_j f\|^2 + \sum_{j=0}^{N-1} 2b_j \langle AF_j f, AF_{j+1} f\rangle
$$

due to (1.75). We further estimate, using (1.72),

$$
\begin{aligned}
|\langle F_j f, F_{j+1} f\rangle| &\leq \tfrac{a_j}{2}\|F_j f\|^2 + \tfrac{1}{2a_j}\|F_{j+1} f\|^2,\\
|\langle F_j f, [F_j, A^*]Af\rangle| &\leq \tfrac{\alpha_j a_j}{2}\|Af\|^2 + \tfrac{1}{2a_j}\|F_j f\|^2,\\
|\langle [F_j, A^*]Af, F_{j+1} f\rangle| &\leq \tfrac{\alpha_j b_j}{2}\|Af\|^2 + \tfrac{1}{2b_j}\|F_{j+1} f\|^2,\\
|\langle F_j f, [F_{j+1}, A^*]Af\rangle| &\leq \tfrac{\alpha_{j+1} b_j}{2}\|Af\|^2 + \tfrac{1}{2b_j}\|F_j f\|^2,\\
|\langle F_j f, F_{j+2} f\rangle| &\leq \tfrac{b_j}{2}\|F_j f\|^2 + \tfrac{1}{2b_j}\|F_{j+2} f\|^2.
\end{aligned}
$$

Inserting these estimates into (1.76) yields

$$
\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{dt}}\mathcal{L}(f) \leq &-2c\|Af\|^2 + \sum_{j=0}^{N}\left(a_j^2\|F_jf\|^2 + \|F_{j+1}f\|^2 + \alpha_j a_j^2\|Af\|^2 + \|F_jf\|^2\right) \\
&+ \sum_{j=0}^{N-1}\Big(\alpha_j b_j^2\|Af\|^2 + \|F_{j+1}f\|^2 + \alpha_{j+1}b_j^2\|Af\|^2 + \|F_jf\|^2\Big) \\
&+ \sum_{j=0}^{N-1}\Big(b_j^2\|F_jf\|^2 + \|F_{j+2}f\|^2\Big) \\
&- \sum_{j=1}^{N} 2b_{j-1}\|F_jf\|^2.
\end{aligned} \tag{1.78}
$$

Note that there are two terms containing $\|F_{N+1}\|$. With $F_{N+1}f_\infty = 0$, (1.70), and Corollary 1.41 we get

$$
\|F_{N+1}f\|^2 \leq c_{N+1}\|f - f_\infty\|_{\mathcal{H}}^2 \leq \beta\sum_{j=0}^{N}\|F_jf\|^2
$$

with some $\beta > 0$. Now we analyse the coefficients of $\|F_jf\|^2$, $0 \leq j \leq N$ (recall $F_0 = A$): on the right hand side of (1.78):

$$
\|Af\|^2 : \quad -2c + a_0^2 + \sum_{j=0}^{N}\alpha_j a_j^2 + \sum_{j=0}^{N-1}\left(\alpha_j b_j^2 + \alpha_{j+1}b_j^2\right) + b_0^2 + 2 + 2\beta, \tag{1.79}
$$

$$
\|F_1f\|^2 : \quad -2b_0 + a_1^2 + b_1^2 + 4 + 2\beta, \tag{1.80}
$$

$$
\|F_jf\|^2, 2 \leq j < N : \quad -2b_{j-1} + a_j^2 + b_j^2 + 5 + 2\beta, \tag{1.81}
$$

$$
\|F_Nf\|^2 : \quad -2b_{N-1} + a_N^2 + 4 + 2\beta. \tag{1.82}
$$

We can choose $a_N := 1, b_{N-1} := \frac{6+2\beta}{2}$, so the term (1.82) is -1. Then, we successively set $a_j := \frac{4b_j^2+1}{a_{j+1}}$ to fulfil (1.75) and $b_{j-1} := \frac{6+2\beta+a_j^2+b_j^2}{2}$, $b_0 := \frac{5+2\beta+a_1^2+b_1^2}{2}$, so the terms in (1.81), (1.80) are -1. Finally, we can choose $c > 0$ such that the term (1.79) is also -1 and obtain

$$
\tfrac{\mathrm{d}}{\mathrm{dt}}\mathcal{L}(f) \leq -\sum_{j=0}^{N}\|F_jf\|^2.
$$

Using (1.41) and (1.74), this directly yields

$$
\tfrac{\mathrm{d}}{\mathrm{dt}}\mathcal{L}(f) \leq -\kappa\|f - f_\infty\|_{\mathcal{H}}^2 \leq -\tfrac{\kappa}{\tilde{c}_3}\mathcal{L}(f).
$$

Using Gronwall's inequality we obtain the exponential decay of $\mathcal{L}(f)$, and thus with (1.73), (1.74),

$$\|f(t) - f_\infty\|_{\mathcal{H}}^2 \leq \tfrac{1}{\tilde{c}_2}\mathcal{L}(f(t)) \leq \tfrac{1}{\tilde{c}_2}\mathcal{L}(f_0)\exp(-\tfrac{\kappa}{\tilde{c}_3}t) \leq \tfrac{\tilde{c}_3}{\tilde{c}_2}\|f_0 - f_\infty\|_{\mathcal{H}}^2 \exp(-\tfrac{\kappa}{\tilde{c}_3}t)$$

for all $t > 0$. $\square$

Remark: We did not keep track of the decay rate in this theorem - it is far from optimal, as it is also in the general theorem given in [67]. In his book, Villani remarks that in some cases, it is possible to get a better rate (that is only "wrong" by one order of magnitude) by tailoring more specifically to the particular form of the operators A, B. In this thesis, we get a sharp rate from the modified entropy method (§1.3).

1.7 Extension to Nonlinear Drift terms

In this section, we investigate how the results of §1.1-§1.3 extend to nonlinear drift coefficients F. We do not reproduce the results of the previous sections; instead, the aim is to show some of the difficulties that arise and present an idea for a relatively strict, but simple extension of condition (A) (see Definition 1.1). We consider the equation

$$\begin{aligned} f_t = Lf &:= \operatorname{div}(D\nabla f + Ff), \qquad (1.83)\\ f(t=0) &= f_0. \end{aligned}$$

We assume $\int\limits_{\mathbb{R}^d} f_0 \, \mathrm{d}x = 1$ and analyse solutions in $L^1_+(\mathbb{R}^d)$. We still assume

$$D = \operatorname{diag}(\underbrace{1,\dots,1}_{k},\underbrace{0,\dots,0}_{d-k}),$$

but we no longer assume that F is linear in the space variable x.

1.7.1 Existence of solutions and stationary states

To guarantee existence of stationary state and a contraction semigroup e^{tL}, we make some assumptions on F. They take the following form:

Definition 1.43. The operator L from (1.83) fulfils condition (B) iff F is smooth and

(i) $\frac{\partial F}{\partial x}$ is uniformly positively stable, i.e. there is a $\mu > 0$ such that for all $x \in \mathbb{R}^d$, the spectrum $\sigma\left(\frac{\partial F}{\partial x}(x)\right)$ is contained in $[\mu, \infty) \times i\mathbb{R}$.

(ii) There exists a matrix $R = -R^T \in \mathbb{R}^{d\times d}$ constant in x such that $D + R$ is invertible and

$$(D+R)^{-1}\frac{\partial F}{\partial x} \tag{1.84}$$

is symmetric for all $x \in \mathbb{R}^d$.

These conditions are inspired by (1.1): We still want a confinement potential (i), and (ii) ensures that there is an easily computable stationary state for 1.83. In contrast to condition (A), condition (B) is not equivalent to existence of stationary states and exponential decay for e^{tL}. In particular, the assumption that the matrix R in (1.84) should be constant is very strict. In §1.7.3, we compute two examples for coefficients F that fulfil condition (B).

Proposition 1.44. *Let L fulfil condition (B). Then the following holds:*

(i) There is a smooth solution V to

$$(D+R)\nabla V = F, \tag{1.85}$$

which is unique up to a constant. Further,

$$f_\infty := c_V e^{-V} \in L^1_+(\mathbb{R}^d) \cap L^2(\mathbb{R}^d),$$

where $c_V > 0$ is such that $\int\limits_{\mathbb{R}^d} f_\infty \,\mathrm{d}x = 1$.

(ii) On the Hilbert space

$$\mathcal{H} := \{f \in L^2(\mathbb{R}^d, f_\infty^{-1}) | \nabla \frac{f}{f_\infty} \in L^2(\mathbb{R}^d, f_\infty)\},$$

the kernel of L is spanned by f_∞.

(iii) $Lf = \operatorname{div}(f_\infty[D+R]\nabla \frac{f}{f_\infty})$.

Proof:

(i) From (B.ii), it follows that

$$(D+R)^{-1}F$$

is a gradient field. Differentiating (1.85) yields

$$\frac{\partial^2 V}{\partial x^2} = (D+R)^{-1}\frac{\partial F}{\partial x} = \frac{\partial F}{\partial x}^T (D-R)^{-1}. \tag{1.86}$$

This implies that $\frac{\partial^2 V}{\partial x^2}$ is regular and smooth, since both $\frac{\partial F}{\partial x}$ and $D+R$ are. Furthermore, $\frac{\partial^2 V}{\partial x^2}$ is a (pointwise) solution to the continuous Lyapunov equation

$$(D+R)\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial x^2}(D+R)^T = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial x}^T.$$

Since $D+R$ is regular and $v^T(D+R)v = v^T Dv \geq 0$ for all $v \in \mathbb{R}^d$, $D+R$ is positively stable. Since $\frac{\partial F}{\partial x}$ is positively stable for any $x \in \mathbb{R}^d$, the right hand side is positive semidefinite; this means that $\frac{\partial^2 V}{\partial x^2}$ is positive semidefinite, too. With the regularity of V, it follows that $\frac{\partial^2 V}{\partial x^2}$ is positive definite for any $x \in \mathbb{R}^d$ (see [44] Theorem 2.2.3, [61] Theorem 2.2). In fact, since $\frac{\partial F}{\partial x}$ is uniformly positively stable, $\frac{\partial^2 V}{\partial x^2}$ is uniformly positive definite. It then follows that V grows at least quadratically for $|x| \to \infty$ and thus $\exp(-V) \in L^1_+(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.
(ii) Inserting f_∞ into L yields

$$\operatorname{div}(f_\infty(-D\nabla V + F)) = f_\infty\left[(\nabla V)\cdot(D\nabla V - F) - \operatorname{div}(D\nabla V - F)\right].$$

Applying (1.85), we obtain

$$Lf_\infty = -f_\infty\left[(\nabla V)\cdot(R\nabla V) - \operatorname{div}(R\nabla V)\right] = 0.$$

So f_∞ is in the kernel of L. Compute

$$\operatorname{div}\left(f_\infty(D+R)\nabla\frac{f}{f_\infty}\right) = \operatorname{div}\left((D+R)\nabla f + (D+R)\nabla V f\right) = \operatorname{div}(D\nabla f + Ff).$$

Thus we can write

$$Lf = \operatorname{div}\left(f_\infty(D+R)\nabla\frac{f}{f_\infty}\right).$$

This proves (iii). Now assume $Lf = 0$. It follows that for all $g \in \mathcal{H}$,

$$0 = \langle Lf, g\rangle = -\int_{\mathbb{R}^d} [(D+R)\nabla\frac{f}{f_\infty}]\cdot[\nabla\frac{g}{f_\infty}]f_\infty \,\mathrm{d}x.$$

Since $D+R$ is regular, this implies $\nabla\frac{f}{f_\infty} = 0$ and thus $f = \alpha f_\infty$. □

Remark: The matrix R from condition (B) is unique and can be computed explicitly: Multiplying (1.86) by $D+R$ from left and $D-R$ from right yields

$$\frac{\partial F}{\partial x}D - \frac{\partial F}{\partial x}R = D\frac{\partial F}{\partial x}^T + R\frac{\partial F}{\partial x}^T,$$

which can be rearranged into the continuous Lyapunov equation

$$\frac{\partial F}{\partial x}R + R\frac{\partial F}{\partial x}^T = \frac{\partial F}{\partial x}D - D\frac{\partial F}{\partial x}^T \tag{1.87}$$

for the unknown matrix R. Point (ii) of condition (B) implies that (1.87) has a unique solution, since $\frac{\partial F}{\partial x}$ is positively stable for any $x \in \mathbb{R}^d$. This solution can be written as

$$R = \int_{\tau=0}^{\infty} \exp(-\tau\frac{\partial F}{\partial x})(\frac{\partial F}{\partial x}D - D\frac{\partial F}{\partial x}^T)\exp(-\tau\frac{\partial F}{\partial x}^T)\,\mathrm{d}\tau,$$

which immediately confirms that R will be skew-symmetric. The restriction of condition (B) is that R should be independent of x; examples are given at the end of this section.

We also remark that a split such as in Proposition 1.44.(iii) is to be expected for (1.83) if there is a unique normalised stationary state:

Lemma 1.45. *Let f_∞ be the unique normalised steady state of L from (1.83), with*

$$\forall x \in \mathbb{R}^d : f_\infty(x) > 0.$$

Then there exists $R = R(x) \in \mathbb{R}^{d\times d}$ such that $R(x) = -R(x)^T$ and

$$Lf = \mathrm{div}[f_\infty[D + R(x)]\nabla\frac{f}{f_\infty}] \tag{1.88}$$

in $L^2(\mathbb{R}^d, f_\infty^{-1})$.

Proof: Since f_∞ is strictly positive, we write $f_\infty(x) = \exp(-A(x))$. We compute, writing $\tilde{R}(x) := f_\infty R(x)$,

$$\begin{aligned}\mathrm{div}[f_\infty[D + R(x)]\nabla\frac{f}{f_\infty}] &= \mathrm{div}[D(\nabla f + f\nabla A)] + \mathrm{div}[\tilde{R}(x)\nabla\frac{f}{f_\infty}]\\ &= \mathrm{div}[D(\nabla f + f\nabla A)] + (\mathrm{div}\,\tilde{R}(x))\cdot\nabla\frac{f}{f_\infty}],\end{aligned} \tag{1.89}$$

where we have used $\mathrm{Tr}(\tilde{R}(x)\frac{\partial^2}{\partial x^2}\frac{f}{f_\infty}) = 0$ due to the antisymmetry of $\tilde{R}$. We know that

$$Lf_\infty = \operatorname{div}[(F - D\nabla A)f_\infty] = 0. \tag{1.90}$$

This implies

$$\operatorname{div}[(F - D\nabla A)f] = \operatorname{div}[(F - D\nabla A)f_\infty \frac{f}{f_\infty}) = (\nabla \frac{f}{f_\infty}) \cdot (F - D\nabla A)f_\infty. \tag{1.91}$$

Writing

$$Lf = \operatorname{div}[D(\nabla f + f\nabla A) + (F - D\nabla A)f] = \operatorname{div}[f_\infty D\nabla \frac{f}{f_\infty}] + \operatorname{div}[(F - D\nabla A)f],$$

we conclude that (1.88) holds if

$$\operatorname{div}[(F - D\nabla A)f] = \operatorname{div}[f_\infty R\nabla \frac{f}{f_\infty}].$$

With (1.89), (1.91), this can be written as

$$(\nabla \frac{f}{f_\infty}) \cdot (F - D\nabla A)f_\infty = (\operatorname{div} \tilde{R}(x)) \cdot \nabla \frac{f}{f_\infty},$$

which is equivalent to

$$\operatorname{div}(\tilde{R}(x)) = (F - D\nabla A)f_\infty.$$

This can always be solved: The right-hand side has divergence 0 due to (1.90), and for the left-hand side we obtain

$$\operatorname{div}(\operatorname{div}(\tilde{R}(x))) = \sum_{i,j=1}^{d} \partial_i \partial_j \tilde{r}_{ij} = 0$$

due to the antisymmetry of $\tilde{R}$. □

Proposition 1.46. *Assume condition (B). Then L generates a contraction semigroup on L^2.*

Proof: We compute

$$\langle Lf, f\rangle = \int_{\mathbb{R}^d} \operatorname{div}\left(f_\infty (D+R)\nabla \frac{f}{f_\infty}\right) \frac{f}{f_\infty}\, \mathrm{d}x = -\int_{\mathbb{R}^d} (\nabla \frac{f}{f_\infty})^T D\nabla \frac{f}{f_\infty} f_\infty \, \mathrm{d}x \le 0.$$

Thus, L is a dissipative operator on L^2. The adjoint L^* of L is easily computed as

$$L^* f = \operatorname{div}\left(f_\infty (D - R) \nabla \frac{f}{f_\infty} \right).$$

It then follows that L^* is dissipative, as well, and thus L generates a contraction semigroup (see e.g. Corollary 3.17 in §II of [32]). □

Corollary 1.47. *Let $0 \le f_0 \in L^2$, $\|f_0\|_{L^1(\mathbb{R}^d)} = 1$. Then the solution f of (1.83) is in L^1 and fulfils*

$$\forall t \ge 0 : f(t) \ge 0, \quad \|f(t)\|_{L^1(\mathbb{R}^d)} = 1.$$

Proof: $f(t) \ge 0$ holds due to the weak maximum principle for degenerate parabolic equations.
We compute

$$\begin{aligned} \|f(t)\|_{L^1(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} f(t) \,\mathrm{d}x = \int_{\mathbb{R}^d} f(t) f_\infty^{-\frac{1}{2}} f_\infty^{\frac{1}{2}} \,\mathrm{d}x \\ &\le \left(\int_{\mathbb{R}^d} |f(t)|^2 f_\infty^{-1} \,\mathrm{d}x \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d} f_\infty \,\mathrm{d}x \right)^{\frac{1}{2}} = \|f(t)\|_{L^2}. \end{aligned}$$

This shows $L^2 \hookrightarrow L^1$, so $f(t) \in L^1$ for all $t \ge 0$. The norm preservation then follows from the divergence form of L. □

1.7.2 Entropy method

In the case of a linear drift coefficient F, we considered the modified entropy functional

$$S(f) := \int_{\mathbb{R}^d} \psi''\left(\frac{f}{f_\infty}\right) u^T P u f_\infty \,\mathrm{d}x \tag{1.92}$$

where $0 < P \in \mathbb{R}^d$ is carefully chosen such that inequality (1.25) holds:

$$QP + PQ^T \ge \mu P.$$

There, Q is computed from the stationary state as $Q := (D - R)\frac{\partial^2 V}{\partial x^2}$. In the case of non-linear drift, the computations leading to Proposition 1.25 can all be repeated (under regularity assumptions on the solution). However, the potential V will not be quadratic, and thus Q is no longer a constant matrix. The linear method can still be used for a perturbation result:

Proposition 1.48. *Let f be a solution to (1.83) under condition (B). Let*

$$\frac{\partial F}{\partial x} = C_0 + C_r(x), \tag{1.93}$$

where $C_0 \in \mathbb{R}^d$ is positively stable with $\mu := \min\{\Re\lambda_j | \lambda_j \in \sigma(C_0)\} > 0$. Further, assume that for the matrix P from Lemma 1.19 corresponding to C_0, it holds that

$$\forall x \in \mathbb{R}^d : Q_r(x)P + PQ_r(x)^T \geq -\nu P \tag{1.94}$$

for some $\nu \in [0, \mu)$. Then the modified entropy production corresponding to P (see (1.24)) fulfils

$$\forall t \geq 0 : \quad S(t) \leq S(0) \exp((\nu - \mu)t).$$

Proof: From (1.93), (1.86) we obtain

$$\begin{aligned} Q &= (D-R)\frac{\partial^2 V}{\partial x^2} = (D-R)\frac{\partial F}{\partial x}^T (D-R)^{-1} \\ &= \underbrace{(D-R)C_0(D-R)^{-1}}_{=:Q_0} + \underbrace{(D-R)C_r(x)(D-R)^{-1}}_{:=Q_r(x)}. \end{aligned}$$

Hence, Q_0 is similar to C_0 and therefore also positively stable. We can apply Lemma 1.19 and obtain a symmetric $P > 0$ such that

$$Q_0 P + PQ_0^T \geq \mu P.$$

Since we assumed

$$Q_r(x)P + PQ_r(x)^T \geq -\nu P,$$

it follows that

$$QP + PQ^T \geq (\mu - \nu)P,$$

and decay of S with rate $\mu - \nu$ follows as in Proposition 1.25. $\square$

The rate obtained by the perturbation result is not going to be sharp. Also, it only covers cases where the growth of $\frac{\partial F}{\partial x}$ – and thus of V – is at most quadratic. As we will see in the examples of the next subsection, for faster-growing potentials it is in general impossible to obtain inequalities using a constant matrix

P. The next logical step would be to use a non-constant matrix P in S. In the following Proposition, we compute the time-derivative of the resulting modified entropy dissipation. While this result leads to a decay estimate in the form $\mathcal{A}(P) \geq \mu P$, it is only of theoretical value without an example that this estimate can actually work.

Proposition 1.49. *Let condition (B) hold. Let f be a solution to (1.83) such that $f(t) > 0$ for all $t \geq 0$, and let*

$$P : \mathbb{R}^d \to \mathbb{R}^{d\times d}$$

be a smooth function such that $P(\cdot) \geq p_0 \,\mathrm{Id}$ uniformly. Then the modified entropy dissipation

$$S(t) := \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u^T P u f_\infty \,\mathrm{d}x$$

fulfils

$$-\tfrac{d}{dt} S(f(t)|f_\infty) \geq -\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) u^T \mathcal{A}(P) u f_\infty \,\mathrm{d}x,$$

where

$$\mathcal{A}(P) = (D-R)\frac{\partial^2 V}{\partial x^2} P + P \frac{\partial^2 V}{\partial x^2}(D+R) - [\nabla V (D-R)\nabla] P + [\nabla^T (D-R)\nabla] P.$$

Proof: Define $u := \nabla \frac{f}{f_\infty}$, then we have

$$\begin{aligned} Lf &= \operatorname{div}((D+R)u f_\infty) = [(D_{lk} + R_{lk}) u_k f_\infty]_{,l} \\ &= D_{lk} u_{k,l} f_\infty - (D_{lk} + R_{lk}) u_k V_{,l} f_\infty, \end{aligned}$$

and thus, using $u_{k,lj} = u_{j,lk}$,

$$\begin{aligned} u_{j,t} &= (\frac{f_t}{f_\infty})_{,j} = (\frac{Lf}{f_\infty})_{,j} \\ &= \left[D_{lk} u_{k,l} - (D_{lk} + R_{lk}) u_k V_{,l}\right]_{,j} \\ &= D_{lk} u_{j,lk} - (D_{lk} + R_{lk}) u_{k,j} V_{,l} - (D_{lk} + R_{lk}) u_k V_{,lj}. \end{aligned}$$

We compute

$$Z_\psi(f(t)) := \tfrac{\mathrm{d}}{\mathrm{dt}} S_\psi(f(t))$$

$$= \underbrace{2\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty})(u_t)^T P u f_\infty \, \mathrm{d}x}_{=:(I)} + \underbrace{\int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) u^T P u f_t \, \mathrm{d}x}_{=:(II)},$$

where we have used the symmetry of P. We have

$$\begin{aligned}
(I) &= 2\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty})(u_t)^T P u f_\infty \, \mathrm{d}x \\
&= 2\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) D_{lk} u_{j,lk} P_{jr} u_r f_\infty \, \mathrm{d}x - 2\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})(D_{lk}+R_{lk}) u_{k,j} V_{,l} P_{jr} u_r f_\infty \, \mathrm{d}x \\
&\quad - 2\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})(D_{lk}+R_{lk}) u_k V_{,lj} P_{jr} u_r f_\infty \, \mathrm{d}x.
\end{aligned}$$

For the first term, compute

$$\begin{aligned}
&2\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) D_{lk} u_{j,kl} P_{jr} u_r f_\infty \, \mathrm{d}x \\
&= -2\int\limits_{\mathbb{R}^d} D_{lk} u_{j,k} (\psi''(\tfrac{f}{f_\infty}) P_{jr} u_r f_\infty)_{,l} \, \mathrm{d}x \\
&= -2\int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} u_l P_{jr} u_r f_\infty \, \mathrm{d}x - 2\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr} u_{r,l} f_\infty \, \mathrm{d}x \\
&\quad + 2\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} V_{,l} P_{jr} u_r f_\infty \, \mathrm{d}x - 2\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr,l} u_r f_\infty \, \mathrm{d}x.
\end{aligned}$$

This implies, again using $u_{j,k} = u_{k,j}$

$$\begin{aligned}
(I) &= -2\int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} u_l P_{jr} u_r f_\infty \, \mathrm{d}x - 2\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr} u_{r,l} f_\infty \, \mathrm{d}x \\
&\quad + 2\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} V_{,l} P_{jr} u_r f_\infty \, \mathrm{d}x - 2\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr,l} u_r f_\infty \, \mathrm{d}x \\
&\quad - 2\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})(D_{lk}+R_{lk}) u_{k,j} V_{,l} P_{jr} u_r f_\infty \, \mathrm{d}x \\
&\quad - 2\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})(D_{lk}+R_{lk}) u_k V_{,lj} P_{jr} u_r f_\infty \, \mathrm{d}x \\
&= -2\int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} u_l P_{jr} u_r f_\infty \, \mathrm{d}x - 2\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr} u_{r,l} f_\infty \, \mathrm{d}x
\end{aligned}$$

$$- 2\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})D_{lk}u_{j,k}P_{jr,l}u_r f_\infty \,\mathrm{d}x - 2\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})R_{lk}u_{k,j}V_{,l}P_{jr}u_r f_\infty \,\mathrm{d}x$$
$$- 2\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})(D_{lk}+R_{lk})u_k V_{,lj}P_{jr}u_r f_\infty \,\mathrm{d}x.$$

Next, we investigate the term

$$\begin{aligned}
&-\int_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty})V_{,l}R_{lk}u_{j,k}P_{jr}u_r f_\infty \,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} u_j(R_{lk}P_{jr}V_{,l}u_r f_\infty \psi''(\tfrac{f}{f_\infty}))_{,k} \,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})u_j P_{jr,k}V_{,l}u_r f_\infty \,\mathrm{d}x + \int_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty})R_{lk}V_{,lk}u_j P_{jr}u_r f_\infty \,\mathrm{d}x \\
&+ \int_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty})V_{,l}R_{lk}u_{r,k}P_{jr}u_j f_\infty \,\mathrm{d}x - \int_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty})V_{,l}R_{lk}V_{,k}u_j P_{jr}u_r f_\infty \,\mathrm{d}x \\
&+ \int_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty})V_{,l}R_{lk}u_k u_j P_{jr}u_r f_\infty \,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})u_j P_{jr,k}R_{lk}V_{,l}u_r f_\infty \,\mathrm{d}x + \int_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty})V_{,l}R_{lk}u_{r,k}P_{jr}u_j f_\infty \,\mathrm{d}x \\
&+ \int_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty})V_{,l}R_{lk}u_k u_j P_{jr}u_r f_\infty \,\mathrm{d}x.
\end{aligned}$$

Here we have used the skew-symmetry of R to conclude $R_{lk}V_{,lk} = V_{,l}R_{lk}V_{,k} = 0$. We obtain

$$\begin{aligned}
&- 2\int_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty})V_{,l}R_{lk}u_{j,k}P_{jr}u_r f_\infty \,\mathrm{d}x \\
&= \int_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty})V_{,l}R_{lk}u_k u_j P_{jr}u_r f_\infty \,\mathrm{d}x + \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})u_j P_{jr,k}R_{lk}V_{,l}u_r f_\infty \,\mathrm{d}x.
\end{aligned}$$

So we arrive at

$$\begin{aligned}
(I) = &-2\int_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty})D_{lk}u_{j,k}u_l P_{jr}u_r f_\infty \,\mathrm{d}x - 2\int_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty})D_{lk}u_{j,k}P_{jr}u_{r,l}f_\infty \,\mathrm{d}x \\
&- 2\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})D_{lk}u_{j,k}P_{jr,l}u_r f_\infty \,\mathrm{d}x \\
&- 2\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty})(D_{lk}+R_{lk})u_k V_{,lj}P_{jr}u_r f_\infty \,\mathrm{d}x
\end{aligned}$$

$$+\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})V_{,l}R_{lk}u_ku_jP_{jr}u_rf_\infty\ \mathrm{d}x+\int_{\mathbb{R}^d}\psi''(\frac{f}{f_\infty})u_jP_{jr,k}R_{lk}V_{,l}u_rf_\infty\ \mathrm{d}x. \tag{1.95}$$

Next, we compute

$$\begin{aligned}(II)&=\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_rP_{rj}u_jLf\ \mathrm{d}x\\&=\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_rP_{rj}u_j(D_{lk}+R_{lk})u_{k,l}f_\infty\ \mathrm{d}x\\&\quad-\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_rP_{rj}u_jV_{,l}(D_{lk}+R_{lk})u_kf_\infty\ \mathrm{d}x.\end{aligned}$$

Take a closer look at

$$\begin{aligned}&\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_rP_{rj}u_j(D_{lk}+R_{lk})u_{k,l}f_\infty\ \mathrm{d}x\\&=\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_rP_{rj}u_jD_{lk}u_{k,l}f_\infty\ \mathrm{d}x\\&=-\int_{\mathbb{R}^d}u_kD_{lk}(P_{rj}u_ru_jf_\infty\psi'''(\tfrac{f}{f_\infty}))_{,l}\ \mathrm{d}x\\&=-\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_kD_{lk}u_{r,l}P_{rj}u_jf_\infty\ \mathrm{d}x-\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_kD_{lk}u_{j,l}P_{rj}u_rf_\infty\ \mathrm{d}x\\&\quad-\int_{\mathbb{R}^d}\psi^{IV}(\tfrac{f}{f_\infty})u_kD_{lk}u_lu_rP_{rj}u_jf_\infty\ \mathrm{d}x+\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_kD_{lk}V_{,l}u_rP_{rj}u_jf_\infty\ \mathrm{d}x\\&\quad-\int_{\mathbb{R}^d}\psi'''(\frac{f}{f_\infty})u_kD_{lk}P_{rj,l}u_ru_jf_\infty\ \mathrm{d}x,\end{aligned}$$

and it follows that

$$\begin{aligned}(II)=&-\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_kD_{lk}u_{r,l}P_{rj}u_jf_\infty\ \mathrm{d}x-\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_kD_{lk}u_{j,l}P_{rj}u_rf_\infty\ \mathrm{d}x\\&-\int_{\mathbb{R}^d}\psi^{IV}(\tfrac{f}{f_\infty})u_kD_{lk}u_lu_rP_{rj}u_jf_\infty\ \mathrm{d}x+\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_kD_{lk}V_{,l}u_rP_{rj}u_jf_\infty\ \mathrm{d}x\\&-\int_{\mathbb{R}^d}\psi'''(\frac{f}{f_\infty})u_kD_{lk}P_{rj,l}u_ru_jf_\infty\ \mathrm{d}x\\&-\int_{\mathbb{R}^d}\psi'''(\tfrac{f}{f_\infty})u_rP_{rj}u_jV_{,l}(D_{lk}+R_{lk})u_kf_\infty\ \mathrm{d}x\end{aligned}$$

$$= -2 \int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) u_k D_{lk} u_{r,l} P_{rj} u_j f_\infty \, \mathrm{d}x - \int\limits_{\mathbb{R}^d} \psi^{IV}(\tfrac{f}{f_\infty}) u_k D_{lk} u_l u_r P_{rj} u_j f_\infty \, \mathrm{d}x$$
$$- \int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) u_r P_{rj} u_j V_{,l} R_{lk} u_k f_\infty \, \mathrm{d}x - \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_k D_{lk} P_{rj,l} u_r u_j f_\infty \, \mathrm{d}x. \tag{1.96}$$

Combining (1.95) and (1.96), we obtain

$$\begin{aligned}
Z_\psi(f(t)) &= -2 \int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} u_l P_{jr} u_r f_\infty \, \mathrm{d}x \\
&- 2 \int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr} u_{r,l} f_\infty \, \mathrm{d}x - 2 \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr,l} u_r f_\infty \, \mathrm{d}x \\
&- 2 \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) (D_{lk} + R_{lk}) u_k V_{,lj} P_{jr} u_r f_\infty \, \mathrm{d}x \\
&+ \int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) V_{,l} R_{lk} u_k u_j P_{jr} u_r f_\infty \, \mathrm{d}x + \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_j P_{jr,k} R_{lk} V_{,l} u_r f_\infty \, \mathrm{d}x \\
&- 2 \int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) u_k D_{lk} u_{r,l} P_{rj} u_j f_\infty \, \mathrm{d}x - \int\limits_{\mathbb{R}^d} \psi^{IV}(\tfrac{f}{f_\infty}) u_k D_{lk} u_l u_r P_{rj} u_j f_\infty \, \mathrm{d}x \\
&- \int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) u_r P_{rj} u_j V_{,l} R_{lk} u_k f_\infty \, \mathrm{d}x - \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_k D_{lk} P_{rj,l} u_r u_j f_\infty \, \mathrm{d}x \\
&= -4 \int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) u_l D_{lk} u_{j,k} P_{jr} u_r f_\infty \, \mathrm{d}x - 2 \int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr} u_{r,l} f_\infty \, \mathrm{d}x \\
&- 2 \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr,l} u_r f_\infty \, \mathrm{d}x - 2 \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) (D_{lk} + R_{lk}) u_k V_{,lj} P_{jr} u_r f_\infty \, \mathrm{d}x \\
&+ \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_j P_{jr,k} R_{lk} V_{,l} u_r f_\infty \, \mathrm{d}x - \int\limits_{\mathbb{R}^d} \psi^{IV}(\tfrac{f}{f_\infty}) u_k D_{lk} u_l u_r P_{rj} u_j f_\infty \, \mathrm{d}x \\
&- \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_k D_{lk} P_{rj,l} u_r u_j f_\infty \, \mathrm{d}x.
\end{aligned}$$

We compute

$$\begin{aligned}
&\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr,l} u_r f_\infty \, \mathrm{d}x \\
&= - \int\limits_{\mathbb{R}^d} \left(\psi''(\frac{f}{f_\infty}) P_{jr,l} u_r f_\infty \right)_{,k} D_{lk} u_j \, \mathrm{d}x \\
&= - \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) P_{jr,l} u_k D_{lk} u_r u_j f_\infty \, \mathrm{d}x - \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) P_{jr,lk} u_r D_{lk} u_j f_\infty \, \mathrm{d}x
\end{aligned}$$

$$-\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) P_{jr,l} u_{r,k} D_{lk} u_j f_\infty \,\mathrm{d}x + \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) P_{jr,l} u_r V_{,k} f_\infty D_{lk} u_j \,\mathrm{d}x.$$

Since $P = P^T$, this implies

$$\begin{aligned}
&-2\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr,l} u_r f_\infty \,\mathrm{d}x \\
&= \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_r D_{lk} P_{jr,lk} u_j f_\infty \,\mathrm{d}x + \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) P_{jr,l} u_k D_{lk} u_r u_j f_\infty \,\mathrm{d}x \\
&\quad - \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) P_{jr,l} u_r V_{,k} f_\infty D_{lk} u_j \,\mathrm{d}x,
\end{aligned}$$

and thus

$$\begin{aligned}
Z_\psi(f(t)) &= -4\int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) u_l D_{lk} u_{j,k} P_{jr} u_r f_\infty \,\mathrm{d}x \\
&- 2\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr} u_{r,l} f_\infty \,\mathrm{d}x \\
&+ \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_r D_{lk} P_{jr,lk} u_j f_\infty \,\mathrm{d}x + \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) P_{jr,l} u_k D_{lk} u_r u_j f_\infty \,\mathrm{d}x \\
&- \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) P_{jr,l} u_r V_{,k} f_\infty D_{lk} u_j \,\mathrm{d}x \\
&- 2\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) (D_{lk} + R_{lk}) u_k V_{,lj} P_{jr} u_r f_\infty \,\mathrm{d}x \\
&+ \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_j P_{jr,k} R_{lk} V_{,l} u_r f_\infty \,\mathrm{d}x - \int\limits_{\mathbb{R}^d} \psi^{IV}(\tfrac{f}{f_\infty}) u_k D_{lk} u_l u_r P_{rj} u_j f_\infty \,\mathrm{d}x \\
&- \int\limits_{\mathbb{R}^d} \psi'''(\frac{f}{f_\infty}) u_k D_{lk} P_{rj,l} u_r u_j f_\infty \,\mathrm{d}x \\
= &-4\int\limits_{\mathbb{R}^d} \psi'''(\tfrac{f}{f_\infty}) u_l D_{lk} u_{j,k} P_{jr} u_r f_\infty \,\mathrm{d}x - 2\int\limits_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) D_{lk} u_{j,k} P_{jr} u_{r,l} f_\infty \,\mathrm{d}x \\
&+ \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_r D_{lk} P_{jr,lk} u_j f_\infty \,\mathrm{d}x \\
&- \int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) P_{jr,k} u_r V_{,l} f_\infty (D_{lk} - R_{lk}) u_j \,\mathrm{d}x \\
&- 2\int\limits_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) (D_{lk} + R_{lk}) u_k V_{,lj} P_{jr} u_r f_\infty \,\mathrm{d}x
\end{aligned}$$

$$
\begin{aligned}
&- \int_{\mathbb{R}^d} \psi^{IV}(\tfrac{f}{f_\infty}) u_k D_{lk} u_l u_r P_{rj} u_j f_\infty \, \mathrm{d}x \\
&= -2 \int_{\mathbb{R}^d} \mathrm{Tr}(XY) f_\infty \, \mathrm{d}x - \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) P_{jr,k} u_r V_{,l} f_\infty (D_{lk} - R_{lk}) u_j \, \mathrm{d}x \\
&+ \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_r D_{lk} P_{jr,lk} u_j f_\infty \, \mathrm{d}x \\
&- 2 \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) (D_{lk} + R_{lk}) u_k V_{,lj} P_{jr} u_r f_\infty \, \mathrm{d}x.
\end{aligned}
$$

Here, the matrices X, Y are given as (cf. Lemma 2.13 in [6])

$$
X = \begin{pmatrix} \psi''(\frac{f}{f_\infty}) & \psi'''(\frac{f}{f_\infty}) \\ \psi'''(\frac{f}{f_\infty}) & \frac{1}{2}\psi^{IV}(\frac{f}{f_\infty}) \end{pmatrix}, \quad Y = \begin{pmatrix} \mathrm{Tr}(D\frac{\partial u}{\partial x} P \frac{\partial u}{\partial x}) & u^T D \frac{\partial u}{\partial x} P u \\ u^T D \frac{\partial u}{\partial x} P u & (u^T P u)(u^T D u) \end{pmatrix}.
$$

Due to the assumptions on ψ (cf. Definition 1.17), $X \geq 0$. To see $Y \geq 0$, we use the Cauchy-Schwarz inequality for the Hilbert-Schmidt norm and the symmetry of D, P to obtain

$$
\begin{aligned}
(u^T D \frac{\partial u}{\partial x} P u)^2 &= \mathrm{Tr}\left(\sqrt{P} u u^T \sqrt{D} \cdot \sqrt{D} \frac{\partial u}{\partial x} \sqrt{P} \right)^2 \\
&\leq \mathrm{Tr}\left(\sqrt{P} u u^T \sqrt{D} \sqrt{D} u u^T \sqrt{P} \right) \mathrm{Tr}\left(\sqrt{D} \frac{\partial u}{\partial x} \sqrt{P} \sqrt{P} \frac{\partial u}{\partial x} \sqrt{D} \right) \\
&= [u^T D u][u^T P u] \mathrm{Tr}\left(D \frac{\partial u}{\partial x} P \frac{\partial u}{\partial x} \right).
\end{aligned}
$$

This implies $\mathrm{Tr}(XY) \geq 0$, and thus

$$
\begin{aligned}
Z_\psi(f(t)) \leq &\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u_r D_{lk} P_{jr,lk} u_j f_\infty \, \mathrm{d}x \\
&- \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) P_{jr,k} u_r V_{,l} f_\infty (D_{lk} - R_{lk}) u_j \, \mathrm{d}x \\
&- 2 \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) (D_{lk} + R_{lk}) u_k V_{,lj} P_{jr} u_r f_\infty \, \mathrm{d}x \\
= &- \int_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) u^T (D - R) \frac{\partial^2 V}{\partial x^2} P u f_\infty \, \mathrm{d}x \\
&- \int_{\mathbb{R}^d} \psi''(\tfrac{f}{f_\infty}) u^T P \frac{\partial^2 V}{\partial x^2} (D + R) u f_\infty \, \mathrm{d}x \\
&- \int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u^T ([\nabla V (D - R) \nabla] P) u f_\infty \, \mathrm{d}x
\end{aligned} \tag{1.97}
$$

$$+\int_{\mathbb{R}^d} \psi''(\frac{f}{f_\infty}) u^T([\nabla^T(D-R)\nabla]P)uf_\infty \, \mathrm{d}x.$$

$\square$

1.7.3 Examples

As a last part of this subsection, we give two simple prototype examples that fulfil condition (B).

Definition 1.50. Let

$$v, w, a, b \in C^\infty(\mathbb{R}).$$

We define the drift coefficients

$$F_1(x) := \begin{pmatrix} v'(x_1) + w'(x_2) \\ -v'(x_1) \end{pmatrix}, \tag{1.98}$$

$$F_2(x) := \begin{pmatrix} \beta a'(\beta x_1 + x_2) + b'(x_1 + \beta x_2) \\ a'(\beta x_1 + x_2) + \beta b'(x_1 + \beta x_2) \end{pmatrix}. \tag{1.99}$$

Further, let

$$D := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then we define the two Fokker-Planck type operators

$$L_j := \mathrm{div}(D\nabla f + F_j f), \quad j = 1, 2. \tag{1.100}$$

First, we investigate condition (B) for both cases:

Lemma 1.51. *Condition (B) holds for L_1, L_2 iff a, b, v, w are strictly convex.*

Proof: We compute, leaving out the arguments for sake of readability

$$\frac{\partial F_1}{\partial x} = \begin{pmatrix} v'' & w'' \\ -v'' & 0 \end{pmatrix},$$

$$\frac{\partial F_2}{\partial x} = \begin{pmatrix} \beta^2 a'' + b'' & \beta(a'' + b'') \\ \beta(a'' + b'') & a'' + \beta^2 b'' \end{pmatrix}.$$

It follows that, at any given $x \in \mathbb{R}^2$, the eigenvalues of $\frac{\partial F_1}{\partial x}$ are

$$\lambda_{1,1/2} = \frac{1}{2}\left(v'' \pm \sqrt{v''(v'' - 4w'')}\right).$$

For positive stability, it is thus necessary that $v'' > 0$ and $w'' > 0$ for all x, which means that v, w are convex. Since $\frac{\partial F_2}{\partial x}$ is symmetric, positive stability simply means that the matrix is positive definite. That requires by the principal minor criterion

$$0 < \beta^2 a'' + b'', \quad 0 < (1-\beta^2)^2 a'' b''.$$

Again we conclude $a'', b'' > 0$, as well as $\beta^2 \neq 1$. If these conditions are fulfilled, it follows immediately that both matrices don't leave the subspace $\{0\} \times \mathbb{R}$ invariant.
It remains to find the skew-symmetric, constant matrix R that fulfils (1.84). The subspace of skew-symmetric matrices in $\mathbb{R}^{2\times 2}$ is one-dimensional, and thus

$$R = \alpha \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

for some $\alpha \in \mathbb{R}$. It follows that $D + R$ is invertible iff $\alpha \neq 0$, and

$$(D+R)^{-1} = \frac{1}{|\alpha|^2} \begin{pmatrix} 0 & \alpha \\ -\alpha & 1 \end{pmatrix}.$$

Using this, we compute

$$(D+R)^{-1}\frac{\partial F_1}{\partial x} = \frac{1}{|\alpha|^2} \begin{pmatrix} -\alpha v'' & 0 \\ -\alpha v'' - v'' & -\alpha w'' \end{pmatrix},$$
$$(D+R)^{-1}\frac{\partial F_2}{\partial x} = \frac{1}{|\alpha|^2} \begin{pmatrix} \alpha\beta(a''+b'') & \alpha(a''+\beta^2 b'') \\ \beta(a''+b'') - \alpha(\beta^2 a'' + b'') & a'' + \beta^2 b'' - \alpha\beta(a''+b'') \end{pmatrix}.$$

So $(D+R)^{-1}\frac{\partial F_1}{\partial x}$ is symmetric iff $\alpha = \alpha_1 := -1$. In the second case, we get the condition

$$\alpha(a'' + \beta^2 b'') = \beta(a''+b'') - \alpha(\beta^2 a'' + b''),$$

which is equivalent to

$$(\alpha - \beta + \alpha\beta^2)(a'' + b'') = 0.$$

We thus need

$$\alpha = \frac{\beta}{1+\beta^2},$$

since $a'', b'' > 0$. This excludes $\beta = 0$ (else $\alpha = 0$ and $D+R$ is not invertible). $\square$

We can now compute the kernel of L_1, L_2 as in Proposition 1.44.

Lemma 1.52. *Let*

$$f_1(x) := \exp(-\frac{v(x_1)+w(x_2)}{2}), \quad f_2(x) := \exp(-\frac{a+\beta^2 b}{1+\beta^2}). \tag{1.101}$$

Then f_1 spans the kernel of L_1, and f_2 the kernel of L_2.

Proof: In the proof of Lemma 1.51, we have already computed the skew-symmetric matrices

$$R_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$
$$R_2 = \frac{\beta}{1+\beta^2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

for which $(D+R_j)^{-1}\frac{\partial F_j}{\partial x}$ is symmetric. From there, one obtains the equations

$$\nabla V_1 = (D+R_1)^{-1}F_1 = \begin{pmatrix} v'(x_1) \\ w'(x_2) \end{pmatrix},$$
$$\nabla V_2 = (D+R_2)^{-1}F_2 = \frac{1}{1+\beta^2}\begin{pmatrix} \beta a'(\beta x_1 + x_2) + \beta^2 b'(x_1+\beta x_2) \\ a'(\beta x_1+x_2) + \beta^3 b'(x_1+\beta x_2) \end{pmatrix}.$$

This implies

$$V_1 = v(x_1) + w(x_2) + c_1,$$
$$V_2 = \frac{a(\beta x_1 + x_2) + \beta^2 b(x_1 + \beta x_2)}{1+\beta^2} + c_2.$$

with constants c_1, c_2 determined by the normalisation. $\square$

A straight-forward computation now yields the matrices

$$Q_j := (D-R_j)\frac{\partial F_j}{\partial x}^T (D-R_j)^{-1}$$

that appear in §1.3.1, (1.25).

Lemma 1.53. *It is*

$$Q_1 = \begin{pmatrix} v'' & -w'' \\ v'' & 0 \end{pmatrix}, \tag{1.102}$$

$$Q_2 = \begin{pmatrix} \beta^2(a''+2b'')+2a''+b'' & \frac{1}{\beta}[2\beta^4 b''+\beta^2(a''+b'')+2a''] \\ -\beta(a''+b'') & -a''-\beta^2 b'' \end{pmatrix}. \tag{1.103}$$

As a last result, we show that a constant matrix P for the modified entropy dissipation will in general not suffice:

Corollary 1.54. *Let v, w grow faster than quadratic as $|x| \to \infty$. Then for any positive definite, symmetric $P \in \mathbb{R}^{2\times 2}$, there is $x_P \in \mathbb{R}^d$ such that*

$$Q_1(x_P)P + PQ_1^T(x_P)$$

is indefinite.

Proof: We set

$$P := \begin{pmatrix} p_1 & p_n \\ p_n & p_2 \end{pmatrix} \in \mathbb{C}^{d\times d}$$

constant, where $p_1, p_2 > 0$, $p_n^2 < p_1 p_2$. This yields

$$Q_1 P + PQ_1^T = \begin{pmatrix} 2p_1 v''(x_1) - 2p_n w''(x_2) & (p_1+p_n)v''(x_1) - p_2 w''(x_2) \\ (p_1+p_n)v''(x_1) - p_2 w''(x_2) & 2p_n v''(x_1) \end{pmatrix}.$$

For this matrix to be positive definite, we need $2p_n v''(x_1) > 0$ and thus $p_n > 0$. But then it follows that, since $v''(x_1)$ and $w''(x_2)$ are not bounded, the first entry will not have a sign independent of x. $\square$

1.8 Conclusion, open questions

For linear Fokker-Planck type equations (2),

$$\partial_t f = Lf = \operatorname{div}(D\nabla f + Cxf),$$

we established a characterisation of the hypoellipticity and hypocoercivity of L in terms of C and D (see condition (A), Definition 1.1). Condition (A) also turned out to be equivalent to L possessing a unique normalised ground state

$$f_\infty = c_K \exp(-\frac{x^T K^{-1} x}{2}),$$

where the covariance matrix K is computed from

$$2D = CK + KC^T.$$

We have extended the entropy method to deal with a singular D. As seen in Theorem 1.27, one can still obtain sharp decay rates in general admissible entropies, at the price of a constant $c > 1$ on the right hand side of the estimate

$$e_\psi(f(t)|f_\infty) \le c e_\psi(f_0|f_\infty) \exp(-2\mu t).$$

Our method does not guarantee that c is optimal. In fact, it is almost certainly not (the optimal c in §1.5, Example 3 is only valid for ψ-compatible initial data). So there is space for improvement with the constant c in Theorem 1.27.
The Bakry-Émery-analogon for the degenerate parabolic case turned out to be the matrix inequality (1.25):

$$QP + PQ^T \ge 2\mu P,$$

where $Q = KC^TK^{-1}$ is computed from the stationary state and the drift coefficient C.
As seen in §1.7, this structure does not easily translate to the case where the drift coefficients are not linear, and further research in this direction is required to extend it. A first step would be to consider the commonly found kinetic case $\mathbb{R}^d = \mathbb{R}_x^{\frac{d}{2}} \times \mathbb{R}_v^{\frac{d}{2}}$, $d \in 2\mathbb{N}$, and

$$f_\infty(x, v) = c_\infty \exp(-\frac{|v|^2}{2} - V(x)),$$

where V is not quadratic in x and fulfils the condition

$$\|\frac{\partial^2 V}{\partial x^2}\| \le c(1 + \|\nabla V\|)$$

for some $c > 0$ (see [10], [67]). Here, as seen in Corollary 1.54, one cannot use a constant matrix P for the modified entropy dissipation S ((1.24), Lemma 1.19).

We remark that it is entirely possible to apply these results to the case $k = d$. In this case, there is already an established decay rate from the entropy method for symmetric operators (see e.g. §2.4 of [6]; [3]). As it turns out, the rate computed in this thesis is different and in general better. In fact, the two rates describe different phenomena: The rate λ_1 from the symmetric method gives a "local" decay rate, applicable at any point t and with the estimate

$$e(t) \le e(0) \exp(-2\lambda_1 t).$$

In contrast, the rate μ established in this chapter is a "global" rate, best understood as the average rate of decay around which the actual rate oscillates. It holds true for estimates

$$e(t) \leq c \exp(-2\mu t),$$

where $c > e(0)$. Both rates help characterise the behaviour of solutions. In the case $k < d$, the "local" rate λ_1 is zero and thus not generally considered (of course this in itself is an interesting behaviour, see also Corollary 1.31). This difference between "local" and "global" rate is discussed in more depth in the paper version [33].

Chapter 2

Discrete open quantum systems

2.1 Outline

This chapter presents some work of the author on an entropy method for finite dimensional open quantum systems in Lindblad form. The equations are given as

$$\rho_t = D\rho := i[H,\rho] + \sum_k [L_k\rho, L_k^\dagger] + [L_k, \rho L_k^\dagger], \tag{2.1}$$
$$\rho(t=0) = \rho_0 \in \mathbb{C}^{d\times d}.$$

Here $[\cdot,\cdot]$ denotes the commutator of two matrices. The term

$$\sum_k [L_k\rho, L_k^\dagger] + [L_k, \rho L_k^\dagger]$$

is called the Lindblad part of D. We look for solutions in the space of *density matrices*, which in finite dimension are matrices $\rho \in \mathbb{C}^{d\times d}$ which are hermitian, positive semidefinite and have trace 1. The operator D from (2.1) generates a semigroup of completely positive, trace-preserving maps called a *quantum dynamical semigroup*. In fact, it is the general form for a generator of quantum dynamical semigroups, as has been established in [49] (for bounded generators), [35] (for finite dimensional systems).

The aim of the presented research was to establish a-priori estimates on the rate of convergence to the equilibrium via study of the relative entropy. As a first step, we show that equation (2.1) can be rewritten to closely resemble the Fokker–Planck equation from Chapter 1, both for $k = d$ and $k < d$. However, while this approach looks quite promising, it fails at two major obstacles: First, the non-commutativity of the underlying space $\mathbb{C}^{d\times d}$, and second, the fact that D has no simple decomposition such as L in §1.2.2. As such, this chapter gives an overview of the idea and a collection of some results, as well as a discussion of the problems facing an entropy method for open quantum systems.

A recent review on the origins and motivation of (2.1) can be found in [1], while [58] provides a good compact introduction. For a more in-depth discussion of open quantum systems, we refer to the books by Breuer and Petruccione [18] as well as Attal, Joye and Pillet [7]- [9]. A general discussion of quantum dynamical semigroups can be found in the book by Alicki and Lendi [2].

General interest in systems of the form (2.1) has resurged with the emergence of quantum engineering; one of the main problems with this idea is decoherence - that is, a move to a diagonal density matrix due to observation of the system by its environment. This phenomenon is unique to open systems. Mathematically, it is due to the influence of the Lindblad term in (2.1). For a good overview, we refer to [69]. Decoherence need not, however, be detrimental in nature - see e.g. [66], where it is shown that such behaviour can also drive the system in a desired direction.

A comprehensive study regarding stationary states and long term behaviour of (2.1) has been given in [11], [12]. The relative entropy of open quantum systems has been studied since their emergence, notably in [48], [47], [63] and [64]. Back then, the entropy method did not exist, and the results focus on the convexity of the entropy, not on Gronwall inequalities between first and second time derivative of the relative entropy of solutions with respect to the stationary state. In the classic case, there is a well-understood connection between convex Sobolev inequalities, hypercontractivity and the decay of relative entropies (see the Introduction). In the non-commutative case, hypercontractivity and convex Sobolev inequalities have been studied before (see for example [54]). However, there has only very recently been progress towards a connection of logarithmic Sobolev inequalities and decay of relative entropies, see [22] and [19].

2.2 Existence of solutions and stationary states

In this section, we discuss the existence of solutions and stationary states. Since the operator D generates a quantum dynamical semigroup, we always have a smooth solution which has some fundamental properties:

Lemma 2.1. *Let $0 \leq \rho_0 \in \mathbb{C}^{d\times d}$ be hermitian with $\mathrm{Tr}(\rho_0) = 1$. Then there exists a unique solution $\rho \in C([0,\infty), \mathbb{C}^{d\times d}) \cap C^\infty((0,\infty), \mathbb{C}^{d\times d})$ to (2.1) with the properties*

(i) $\rho(t) \geq 0$ *is hermitian.*

(ii) $\mathrm{Tr}(\rho(t)) = 1, t \geq 0$.

Proof: This is a consequence of the fact that the operator D from (2.1) generates a quantum dynamical semigroup, that is a semigroup which is completely positive - implying (i) - and trace preserving (ii). We remark that for our problem, requiring positivity instead of complete positivity would be sufficient. For details we refer to [35], [49]. □

With the existence of solutions established, we turn towards the question of stationary states:

Lemma 2.2. *The kernel of the operator D from (2.1) has at least dimension 1.*

Proof: See proposition 5 of [12]. □

Lemma 2.2 implies that the discussion of stationary states need only focus on uniqueness, not on existence of stationary states. In general, the kernel of D can be quite large. It also turns out that the stationary state ρ_∞ is not necessarily invertible; as a simple example consider the case $d = 2$, $H = 0$,

$$L = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then the unique normalised stationary state is $\rho_\infty = \mathrm{diag}(1, 0)$. We give a result on convergence to non-invertible stationary states:

Proposition 2.3. *Let D from (2.1) be given as*

$$D\rho = \frac{1}{2}([L\rho, L^\dagger] + [L, \rho L^\dagger]),$$

where $L \in \mathbb{C}^{d\times d}$ is such that

(1) $LL^\dagger L = \alpha L$ *for some* $0 < \alpha \in \mathbb{R}$,

(2) $L^k = 0$, $L^{k-1} \neq 0$ *for some* $k > 1$.

Then it holds that

(i) The kernel of D is spanned by the density matrices η with $L\eta = 0$. If $k = d$, then there is a unique normalised stationary state $\rho_\infty = (\mathrm{Id} - cL^\dagger L)$ with $c \in \mathbb{R}$ such that $\mathrm{Tr}(\rho_\infty) = 1$.

(ii) *For any hermitian $0 \leq \rho_0 \in \mathbb{C}^{d\times d}$, let ρ be the solution to $\dot\rho = D\rho$, $\rho(t=0) = \rho_0$. Then the projection of ρ onto the orthogonal complement of the kernel of D (in the standard Hilbert-Schmidt space $C^{d\times d}$) decays to 0 exponentially with rate at least $\alpha - \varepsilon$ for any $\varepsilon \in (0,\alpha)$.*

Proof: (i): This is a reformulation of a result from [11]. Let η be a density matrix such that $D\eta = 0$. It follows that

$$\begin{aligned}
0 = L^{k-1}D\eta(L^\dagger)^{k-1} &= L^{k-1}[L\eta L^\dagger - \frac{1}{2}(L^\dagger L\eta + \eta L^\dagger L)](L^\dagger)^{k-1}\\
&= \left[L^k\eta(L^\dagger)^k - \frac{1}{2}(L^{k-1}L^\dagger L\eta(L^\dagger)^{k-1} + L^{k-1}\eta L^\dagger L(L^\dagger)^{k-1})\right]\\
&= -\alpha L^{k-1}\eta(L^\dagger)^{k-1}. \qquad (2.2)
\end{aligned}$$

Thus, $L^{k-1}\eta(L^\dagger)^{k-1} = 0$, and from (2) we have $L^{d-1} \neq 0$. We can now compute $L^jD\eta L^j$ iteratively for $j = k-2, k-3, \ldots, 1$ and obtain, as in (2.2), $L^j\eta(L^\dagger)^j = 0$ for $j \geq 1$. η has a hermitian square root $0 \leq \nu \in \mathbb{C}^{d\times d}$, $\eta = \nu^\dagger\nu$. Then we have

$$L^j\eta(L^\dagger)^j = [L^j\nu^\dagger][L^j\nu^\dagger]^\dagger = 0,$$

which implies $L^j\nu^\dagger = 0$ and thus $L^j\eta = 0$ for all $j \geq 1$, which proves the first part of (i). If $k = d$, (2) implies that there is $v \in \mathbb{R}^d$ such that the $v_j := L^jv$ are linearly independent - or, in other words, the kernel of L is one dimensional. But then $L\eta = 0$ implies that the image of η is in the kernel of L, and trace normalisation yields that η is the projector onto the kernel of L:

$$\eta = P_{\ker L} = \frac{\mathrm{Id} - \frac{1}{\alpha}L^\dagger L}{\mathrm{Tr}(\mathrm{Id} - \frac{1}{\alpha}L^\dagger L)}.$$

(ii): Let ρ be a density matrix. From (i), we know that $L^\dagger L\rho L^\dagger L$ is the projection of ρ onto the orthogonal complement of the kernel of D. We consider the Hilbert-Schmidt norm of this projection, which with (1) is

$$E(\rho) := \mathrm{Tr}(L\rho L^\dagger L\rho L^\dagger).$$

We compute.

$$0 \leq \mathrm{Tr}([\varepsilon A - \frac{1}{\varepsilon}B]^\dagger[\varepsilon A - \frac{1}{\varepsilon}B]) = \varepsilon^2\mathrm{Tr}(A^\dagger A) - \mathrm{Tr}(A^\dagger B + B^\dagger A) + \frac{1}{\varepsilon^2}\mathrm{Tr}(B^\dagger B),$$

which implies the inequality

$$\mathrm{Tr}(A^\dagger B + AB^\dagger) \leq \varepsilon^2\mathrm{Tr}(A^\dagger A) + \frac{1}{\varepsilon^2}\mathrm{Tr}(B^\dagger B). \qquad (2.3)$$

Using (2.3) and (1), we obtain

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{dt}}E(\rho) &= 2\mathrm{Tr}(LD\rho L^\dagger L\rho L^\dagger) \\
&= 2\mathrm{Tr}(L^2\rho(L^\dagger)^2 L\rho L^\dagger) - 2\mathrm{Tr}(L^\dagger L L^\dagger L\rho L^\dagger L\rho) \\
&= 2\mathrm{Tr}(L^2\rho(L^\dagger)^2 L\rho L^\dagger) - 2\alpha E(\rho) \\
&\leq -2(\alpha-\varepsilon)E(\rho) + \frac{2}{\varepsilon}\mathrm{Tr}(L^2\rho(L^\dagger)^2 L^2\rho(L^\dagger)^2)
\end{aligned}$$

If $k = 2$, $L^2 = 0$ and we are done. If $k > 2$, let

$$F_j(\rho) := \mathrm{Tr}(L^j\rho(L^\dagger)^j L^j\rho(L^\dagger)^j) \geq 0.$$

We compute

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{dt}}F_k &= 2\mathrm{Tr}(L^{j+1}\rho(L^\dagger)^{j+1}L^j\rho(L^\dagger)^j) - 2\alpha F_k \\
&\leq -2(\alpha - \frac{\varepsilon}{2})F_j + \tfrac{4}{\varepsilon}F_{j+1}.
\end{aligned}$$

From (2), we have $F_j = 0$ for $j \geq k$. Now let

$$F(\rho) := E(\rho) + \sum_{j=2}^{k-1} C_j F_j \geq E(\rho)$$

with some constants $C_j > 0$ to be specified. Then

$$\tfrac{\mathrm{d}}{\mathrm{dt}}F(\rho) \leq -2(\alpha-\varepsilon)E(\rho) + \sum_{j=2}^{k-1}\left(\left[-2C_j(\alpha - \frac{\varepsilon}{2}) + \frac{4C_{j-1}}{\varepsilon}\right]F_j\right),$$

where $C_1 := 1$. Now we iteratively choose $C_j := \frac{4}{\varepsilon^2}C_{j-1}$ such that

$$-2C_j(\alpha - \frac{\varepsilon}{2}) + \frac{4C_{j-1}}{\varepsilon} = -2C_j(\alpha - \varepsilon).$$

Then

$$\tfrac{\mathrm{d}}{\mathrm{dt}}F(\rho) \leq -2(\alpha-\varepsilon)F(\rho)$$

which implies with Gronwall's Lemma:

$$E(\rho) \leq F(\rho) \leq F(\rho_0)e^{-2(1-\varepsilon)t}$$

and thus gives decay of E with rate $2(1-\varepsilon)$. □

2.3 Transformation

In this section, we show how the equation (2.1) can be rewritten in a form that closely resembles the Fokker-Planck equations discussed in the previous two chapters. The underlying idea is to define discrete derivatives using commutators (see [53], in particular chapter 2, for more on this concept). We then introduce a special basis of $\mathbb{C}^{d\times d}$ inspired by the Pauli matrices. This basis allows the transformation of the operator while identifying some fundamental properties of the Lindblad part.

2.3.1 Setting, preliminary results

As setting for the transformation, we use the Hilbert space $C^{d\times d}$ equipped with the Hilbert-Schmidt norm. We denote by $A^\dagger$ the hermitian adjoint of $A \in \mathbb{C}^{d\times d}$.

Lemma 2.4. *Let $\mathcal{H} := \mathbb{C}^{d\times d}$, and define*

$$\forall A, B \in \mathcal{H}: \quad \langle A, B\rangle_d := \mathrm{Tr}(AB^\dagger)$$

$$\forall A = (A_k)_{k=1,\dots,d^2},\ B = (B_k)_{k=1,\dots,d^2} \in \mathcal{H}^{d^2} : \langle A, B\rangle_{d^2} := \sum_{k=1}^{d^2} \mathrm{Tr}(A_k B_k^\dagger).$$

Let $\{E_j | j = 1, \dots, d^2\} \subset \mathcal{H}$ be given as

$$\begin{aligned}
E_j &= (\delta_{lj}\delta_{kj})_{l,k}, \quad j = 1, \dots, d;\\
E_j &= \tfrac{1}{\sqrt{2}}(\delta_{lr_j}\delta_{ks_j} + \delta_{kr_j}\delta_{ls_j})_{l,k}, \quad j = d+1, \dots, \tfrac{d^2+d}{2};\\
E_j &= \tfrac{1}{\sqrt{2}}(-i\delta_{lr_j}\delta_{ks_j} + i\delta_{kr_j}\delta_{ls_j})_{l,k} \quad j = \tfrac{d^2+d}{2} + 1, \dots, d^2,
\end{aligned}$$

where r_j runs from 2 to d with increasing j, and for fixed r_j, s_j runs from 1 to $r_j - 1$. Then the following holds:

(i) $(\mathcal{H}, \langle\cdot,\cdot\rangle_d)$ and $(\mathcal{H}^{d^2}, \langle\cdot,\cdot\rangle_{d^2})$ are Hilbert spaces.

(ii) $\{E_j | j = 1, \dots, d^2\}$ is an orthonormal basis of $\mathcal{H}$. In addition, it is an $\mathbb{R}$-basis of all hermitian $d \times d$-matrices.

Proof:

(i) is a well-known result, the details of which can be found in many books on matrix algebra. We refer to the classic book by Horn, Johnson [43].

(ii) Since one easily checks that E_j is hermitian for $1 \le j \le d^2$, all we need to show is

$$\mathrm{Tr}(E_j E_k) = \delta_{jk}, \quad 1 \leq j,k \leq d^2.$$

We compute the case $j,k > \frac{d^2+d}{2}$:

$$\begin{aligned}
\mathrm{Tr}(E_j E_k) &= \tfrac{1}{2}\mathrm{Tr}\left(\sum_{l=1}^{d^2} \left[(-i\delta_{mr_j}\delta_{ls_j} + i\delta_{lr_j}\delta_{ms_j})(-i\delta_{lr_k}\delta_{ns_k} + i\delta_{nr_k}\delta_{ls_k})\right]_{m,n}\right) \\
&= \tfrac{1}{2}\sum_{l,m=1}^{d^2} \left(-\delta_{mr_j}\delta_{ls_j}\delta_{lr_k}\delta_{ms_k} + \delta_{lr_j}\delta_{ms_j}\delta_{lr_k}\delta_{ms_k}\right) \\
&\quad + \tfrac{1}{2}\sum_{l,m=1}^{d^2} \left(\delta_{mr_j}\delta_{ls_j}\delta_{mr_k}\delta_{ls_k} - \delta_{lr_j}\delta_{ms_j}\delta_{mr_k}\delta_{ls_k}\right) \\
&= -\delta_{r_j s_k}\delta_{s_j r_k} + \delta_{r_j r_k}\delta_{s_j s_k}.
\end{aligned}$$

Since $r_j = s_k$ implies $s_j < r_j = s_k < r_k$, the first term is always zero. The second term is nonzero iff $r_j = r_k$ and $s_j = s_k$, which implies $j = k$. The computation for the other cases is completely analogous, and thus the proof is complete. □

Example: In $3d$, the basis is formed by the following matrices:

$$E_1 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad E_2 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad E_3 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$E_4 := \tfrac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad E_5 := \tfrac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad E_6 := \tfrac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$
$$E_7 := \tfrac{1}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad E_8 := \tfrac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \qquad E_9 := \tfrac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}.$$

Next, we define discrete versions of differential operators on $\mathcal{H}$.

Definition 2.5. The discrete divergence div_d, discrete gradient ∇_d and discrete Laplacian Δ_d are defined as

$$\begin{aligned}
\mathrm{div}_d : \mathcal{H}^{d^2} &\to \mathcal{H} \\
K = (K_j)_{j=1,\ldots,d^2} &\mapsto i\sum_{j=1}^{d^2}[K_j, E_j],
\end{aligned}$$

$$\begin{aligned}
\nabla_d &: \mathcal{H} \to \mathcal{H}^{d^2} \\
&\quad K \mapsto (i[K, E_j])_{j=1,\dots,d^2}, \\
\Delta_d &: \mathcal{H} \to \mathcal{H} \\
&\quad K \mapsto \operatorname{div}_d(\nabla_d K) = \sum_{j=1}^{d^2} [E_j, [K, E_j]].
\end{aligned}$$

Some fundamental properties of these operators are easily obtained:

Proposition 2.6. *The following holds:*

(i) For any hermitian matrix $\eta \in \mathcal{H}$*, the components of* $\nabla_d \eta \in \mathcal{H}^{d^2}$ *are hermitian. For any matrix* $K = (K_j)_j \in \mathcal{H}^{d^2}$ *with hermitian components* K_j*,* $\operatorname{div}_d K$ *is hermitian.*

(ii) div_d *is the adjoint to* $-\nabla_d$ *and vice versa.*

(iii) A set of eigenvectors (with corresponding eigenvalues) of Δ_d *is the following:*

$$\begin{aligned}
\sum_{k=1}^{d} E_k = \mathrm{Id}_d &\quad (\lambda = 0), \\
E_k (k > d) &\quad (\lambda = -2d), \\
E_l - E_{l+1} (l = 1, \dots, d-1) &\quad (\lambda = -2d).
\end{aligned}$$

These form a basis of $\mathcal{H}$*.*

(iv) There is a discrete Poincaré equality arising from the spectral gap in Δ_d*:*

$$\forall \rho \in \mathcal{H} : \|\nabla_d \rho\|^2_{\mathcal{H}^{d^2}} + 2|\mathrm{Tr}(\rho)|^2 = 2d\|\rho\|^2_{\mathcal{H}} \tag{2.4}$$

Proof:
(i): We compute for hermitian A, B:

$$(i[A, B])^\dagger = -i(AB - BA)^\dagger = -i(BA - AB) = i[A, B].$$

(ii): Let $A \in \mathcal{H}$, $B \in \mathcal{H}^{d^2}$. Then

$$\begin{aligned}
\langle \nabla_d A, B \rangle_{\mathcal{H}^{d^2}} &= \sum_{k=1}^{d^2} i\mathrm{Tr}([A, E_k] B_k) = \sum_{k=1}^{d^2} -i\mathrm{Tr}(A[B_k, E_k]) \\
&= \sum_{k=1}^{d^2} -\langle A, \operatorname{div}_d B \rangle_{\mathcal{H}}.
\end{aligned}$$

(iii): Id is an eigenvector of Δ_d to the eigenvalue 0, since $[\mathrm{Id}_d, E_k] = 0$ for $1 \le k \le d^2$. On the other hand, since the E_k form a basis of $\mathcal{H}$, it follows that $[A, E_k] = 0$ for $1 \le k \le d^2$ implies $[A, B] = 0$ for all $B \in \mathcal{H}$, and thus A is a multiple of Id_d. This shows that Id_d spans the kernel of Δ_d.

To compute the other eigenvectors, we expand $\Delta_d A$ as

$$\begin{aligned}\Delta_d A &= \sum_{j=1}^{d^2}[E_j, [A, E_j]] = \sum_{j=1}^{d^2}\big(2E_jAE_j - (E_j^2A + AE_j^2)\big)\\ &= 2\sum_{j=1}^{d^2} E_jAE_j - \Big([\sum_{j=1}^{d^2} E_j^2]A + A[\sum_{j=1}^{d^2} E_j^2]\Big).\end{aligned} \tag{2.5}$$

First, we look at the term

$$\sum_{j=1}^{d^2} E_j^2.$$

Since $E_j^2 = E_j$, $1 \le j \le d$, it is $\sum\limits_{j=1}^{d} E_j^2 = \mathrm{Id}_d$. For E_j, $\frac{d^2+d}{2} \ge j > d$, we obtain

$$\begin{aligned}(E_j^2)_{l,k} &= \sum_{\alpha=1}^{d}(E_j)_{l,\alpha}(E_j)_{\alpha,k}\\ &= \tfrac{1}{2}\sum_{\alpha=1}^{d}(\delta_{lr}\delta_{\alpha s} + \delta_{ls}\delta_{\alpha r})(\delta_{\alpha r}\delta_{ks} + \delta_{\alpha s}\delta_{kr})\\ &= \tfrac{1}{2}\big(\underbrace{\sum_{\alpha=1}^{d}(\delta_{lr}\delta_{\alpha s}\delta_{\alpha r}\delta_{ks} + \delta_{ls}\delta_{\alpha r}\delta_{\alpha s}\delta_{kr})}_{=0,\ r\neq s} + \sum_{\alpha=1}^{d}(\delta_{lr}\delta_{\alpha s}\delta_{\alpha s}\delta_{kr} + \delta_{ls}\delta_{\alpha r}\delta_{\alpha r}\delta_{ks})\big)\\ &= \tfrac{1}{2}(\delta_{lr}\delta_{kr} + \delta_{ls}\delta_{ks})\end{aligned}$$

Thus

$$E_j^2 = \frac{1}{2}(E_r + E_s),$$

where $d \ge r > s \ge 1$ from the definition of E_j. We conclude that

$$\sum_{j=d+1}^{\frac{d^2+d}{2}} E_j^2 = \sum_{r=2}^{d}\sum_{s=1}^{r-1}\frac{1}{2}(E_r + E_s) = \frac{d-1}{2}\,\mathrm{Id}_d\,.$$

The computations for E_j, $\frac{d^2+d}{2} < j \leq d^2$ are analogous, and we obtain

$$\sum_{j=1}^{d^2} E_j^2 = d\,\mathrm{Id}_d\,.$$

Inserting this into (2.5) yields

$$\Delta_d A = 2\sum_{k=1}^{d^2}(E_k A E_k) - 2dA. \tag{2.6}$$

Now consider the term $2\sum\limits_{k=1}^{d^2}(E_k A E_k)$. We first assume $A = E_j$ for some $j \in \{1,\dots,d^2\}$ and make a distinction by case.
(a) If $A = E_j$, $j \leq d$, we compute for $m \leq d$

$$E_m E_j E_m = \delta_{jm} E_j.$$

For $\frac{d^2+d}{2} \geq m > d$, we obtain

$$\begin{aligned}
(E_j E_m)_{l,k} &= \sum_{\alpha=1}^{d}(E_j)_{l,\alpha}(E_m)_{\alpha,k} \\
&= \tfrac{1}{\sqrt{2}}\sum_{\alpha=1}^{d}\delta_{lj}\delta_{\alpha j}(\delta_{\alpha r}\delta_{ks} + \delta_{kr}\delta_{\alpha s}) \\
&= \tfrac{1}{\sqrt{2}}(\delta_{lj}\delta_{jr}\delta_{ks} + \delta_{lj}\delta_{kr}\delta_{js}) = \tfrac{1}{\sqrt{2}}\begin{cases} 0, & j \neq r,s \\ \delta_{lr}\delta_{ks}, & j = r \\ \delta_{ls}\delta_{kr}, & j = s \end{cases}
\end{aligned} \tag{2.7}$$

and thus

$$\begin{aligned}
(E_m E_j E_m)_{l,k} &= \sum_{\alpha=1}^{d}(E_m)_{l,\alpha}(E_j E_m)_{\alpha,k} \\
&= \tfrac{1}{2}\begin{cases} \sum\limits_{\alpha=1}^{d}(\delta_{lr}\delta_{\alpha s} + \delta_{\alpha r}\delta_{ls})\delta_{\alpha r}\delta_{ks} = \delta_{ks}\delta_{ls}, & j = r \\ \sum\limits_{\alpha=1}^{d}(\delta_{lr}\delta_{\alpha s} + \delta_{\alpha r}\delta_{ls})\delta_{\alpha s}\delta_{kr} = \delta_{kr}\delta_{lr}, & j = s \end{cases},
\end{aligned}$$

which means

$$E_m E_j E_m = \begin{cases} 0, & j \neq r,s \\ \frac{1}{2}E_s, & j = r \\ \frac{1}{2}E_r, & j = s \end{cases}.$$

This implies

$$\sum_{m=d+1}^{\frac{d^2+d}{2}} E_m E_j E_m = \sum_{s=1}^{j-1} \frac{1}{2} E_s + \sum_{r=j+1}^{d} \frac{1}{2} E_r = \frac{1}{2}(\mathrm{Id}_d - E_j),$$

where the term $\sum\limits_{s=1}^{j-1} \frac{1}{2} E_s$ represents all the contributions for $j = r$, and the term $\sum\limits_{r=j+1}^{d} \frac{1}{2} E_r$ all those for $j = s$. The computations for $m > \frac{d^2+d}{2}$ are once again analogous, and we obtain

$$\sum_{k=1}^{d^2} E_k E_j E_k = E_j + \mathrm{Id} - E_j = \mathrm{Id}\,.$$

Thus (2.6) yields

$$\Delta_d E_j = \mathrm{Id}_d - 2d E_j, \quad (j \leq d). \tag{2.8}$$

From (2.8), we see that for $1 \leq l < k \leq d$ it is

$$\Delta_d(E_l - E_k) = -2d(E_l - E_k),$$

so $E_l - E_k$ is an eigenvector of Δ_d to the eigenvalue $-2d$ for any $l \neq k \in \{1, \ldots, d\}$.

(b) Now, consider the case $A = E_j$, $\frac{d^2+d}{2} \geq j > d$. We have for $1 \leq m \leq d$, using (2.7)

$$\begin{aligned}
(E_m E_j E_m)_{l,k} &= \sum_{\alpha=1}^{d} (E_m E_j)_{l\alpha} (E_m)_{\alpha k} \\
&= \tfrac{1}{\sqrt{2}} \begin{cases} \sum\limits_{\alpha=1}^{d} \delta_{lr}\delta_{\alpha s}\delta_{m\alpha}\delta_{mk}, & m = r \\ \sum\limits_{\alpha=1}^{d} \delta_{ls}\delta_{\alpha r}\delta_{m\alpha}\delta_{mk}, & m = s \end{cases} \\
&= 0
\end{aligned} \tag{2.9}$$

For $d < m \leq \frac{d^2+d}{2}$, we obtain

$$(E_m E_j)_{l,k} = \sum_{\alpha=1}^{d} (E_m)_{l,\alpha} (E_j)_{\alpha,k}$$

$$= \frac{1}{2} \sum_{\alpha=1}^{d} (\delta_{lr}\delta_{\alpha s} + \delta_{ls}\delta_{\alpha r})(\delta_{\alpha u}\delta_{kv} + \delta_{\alpha v}\delta_{ku}); \quad u > v, r > s \in \{1, \ldots, d\}$$

$$= \frac{1}{2}\sum_{\alpha=1}^{d}(\delta_{lr}\delta_{\alpha s}\delta_{\alpha u}\delta_{kv} + \delta_{lr}\delta_{\alpha s}\delta_{\alpha v}\delta_{ku} + \delta_{ls}\delta_{\alpha r}\delta_{\alpha u}\delta_{kv} + \delta_{ls}\delta_{\alpha r}\delta_{\alpha v}\delta_{ku})$$
$$= \frac{1}{2}\sum_{\alpha=1}^{d}(\delta_{lr}\delta_{su}\delta_{kv} + \delta_{lr}\delta_{sv}\delta_{ku} + \delta_{ls}\delta_{ru}\delta_{kv} + \delta_{ls}\delta_{rv}\delta_{ku})$$

and thus

$$(E_mE_jE_m)_{l,k} = \sum_{\alpha=1}^{n}(E_mE_j)_{l,\alpha}(E_m)_{\alpha,k}$$
$$= \frac{1}{2\sqrt{2}}\sum_{\alpha=1}^{d}(\delta_{lr}\delta_{su}\delta_{\alpha v} + \delta_{lr}\delta_{sv}\delta_{\alpha u} + \delta_{ls}\delta_{ru}\delta_{\alpha v} + \delta_{ls}\delta_{rv}\delta_{\alpha u})(\delta_{\alpha r}\delta_{ks} + \delta_{\alpha s}\delta_{kr})$$
$$= \frac{1}{2\sqrt{2}}\sum_{\alpha=1}^{d}(\delta_{lr}\delta_{su}\delta_{\alpha v}\delta_{\alpha r}\delta_{ks} + \delta_{lr}\delta_{sv}\delta_{\alpha u}\delta_{\alpha r}\delta_{ks} + \delta_{ls}\delta_{ru}\delta_{\alpha v}\delta_{\alpha r}\delta_{ks} + \delta_{ls}\delta_{rv}\delta_{\alpha u}\delta_{\alpha r}\delta_{ks}$$
$$+ \delta_{lr}\delta_{su}\delta_{\alpha v}\delta_{\alpha s}\delta_{kr} + \delta_{lr}\delta_{sv}\delta_{\alpha u}\delta_{\alpha s}\delta_{kr} + \delta_{ls}\delta_{ru}\delta_{\alpha v}\delta_{\alpha s}\delta_{kr} + \delta_{ls}\delta_{rv}\delta_{\alpha u}\delta_{\alpha s}\delta_{kr})$$
$$= \frac{1}{2\sqrt{2}}\sum_{\alpha=1}^{d}(\underbrace{\delta_{lr}\delta_{su}\delta_{vr}\delta_{ks}}_{=0,u>v\&s<r} + \delta_{lr}\delta_{sv}\delta_{ur}\delta_{ks} + \underbrace{\delta_{ls}\delta_{ru}\delta_{vr}\delta_{ks}}_{=0,u>v} + \underbrace{\delta_{ls}\delta_{rv}\delta_{ur}\delta_{ks}}_{=0,u>v}$$
$$+ \delta_{lr}\underbrace{\delta_{su}\delta_{vs}\delta_{kr}}_{=0,u>v} + \underbrace{\delta_{lr}\delta_{sv}\delta_{us}\delta_{kr}}_{=0,u>v} + \delta_{ls}\delta_{ru}\delta_{vs}\delta_{kr} + \underbrace{\delta_{ls}\delta_{rv}\delta_{us}\delta_{kr}}_{=0,r>s\&u>v})$$
$$= \frac{1}{2\sqrt{2}}\delta_{ur}\delta_{vs}(E_j)_{l,k}. \tag{2.10}$$

Finally, for $d^2 \geq m > \frac{d^2+d}{2}$ we compute

$$(E_mE_j)_{l,k} = \sum_{\alpha=1}^{n}(E_m)_{l,\alpha}(E_j)_{\alpha,k}$$
$$= \frac{1}{2}\sum_{\alpha=1}^{d}(-i\delta_{lr}\delta_{\alpha s} + i\delta_{ls}\delta_{\alpha r})(\delta_{\alpha u}\delta_{kv} + \delta_{\alpha v}\delta_{ku}); \quad u > v, r > s \in \{1,\dots,d\}$$
$$= \frac{1}{2}\sum_{\alpha=1}^{d}(-i\delta_{lr}\delta_{\alpha s}\delta_{\alpha u}\delta_{kv} - i\delta_{lr}\delta_{\alpha s}\delta_{\alpha v}\delta_{ku} + i\delta_{ls}\delta_{\alpha r}\delta_{\alpha u}\delta_{kv} + i\delta_{ls}\delta_{\alpha r}\delta_{\alpha v}\delta_{ku})$$
$$= \frac{1}{2}\sum_{\alpha=1}^{d}(-i\delta_{lr}\delta_{su}\delta_{kv} - i\delta_{lr}\delta_{sv}\delta_{ku} + i\delta_{ls}\delta_{ru}\delta_{kv} + i\delta_{ls}\delta_{rv}\delta_{ku}).$$

Using this, we obtain

$$(E_mE_jE_m)_{l,k} = \sum_{\alpha=1}^{n}(E_mE_j)_{l,\alpha}(E_m)_{\alpha,k}$$
$$= \frac{1}{2\sqrt{2}}\sum_{\alpha=1}^{d}(-i\delta_{lr}\delta_{su}\delta_{\alpha v} - i\delta_{lr}\delta_{sv}\delta_{\alpha u} + i\delta_{ls}\delta_{ru}\delta_{\alpha v} + i\delta_{ls}\delta_{rv}\delta_{\alpha u})(-i\delta_{\alpha r}\delta_{ks} + i\delta_{\alpha s}\delta_{kr})$$

$$= \frac{1}{2\sqrt{2}} \sum_{\alpha=1}^{d} (-\delta_{lr}\delta_{su}\delta_{\alpha v}\delta_{\alpha r}\delta_{ks} - \delta_{lr}\delta_{sv}\delta_{\alpha u}\delta_{\alpha r}\delta_{ks} + \delta_{ls}\delta_{ru}\delta_{\alpha v}\delta_{\alpha r}\delta_{ks} + \delta_{ls}\delta_{rv}\delta_{\alpha u}\delta_{\alpha r}\delta_{ks}$$

$$+ \delta_{lr}\delta_{su}\delta_{\alpha v}\delta_{\alpha s}\delta_{kr} + \delta_{lr}\delta_{sv}\delta_{\alpha u}\delta_{\alpha s}\delta_{kr} - \delta_{ls}\delta_{ru}\delta_{\alpha v}\delta_{\alpha s}\delta_{kr} - \delta_{ls}\delta_{rv}\delta_{\alpha u}\delta_{\alpha s}\delta_{kr})$$

$$= \frac{1}{2\sqrt{2}} \sum_{\alpha=1}^{d} (- \underbrace{\delta_{lr}\delta_{su}\delta_{vr}\delta_{ks}}_{=0, u>v \& s<r} - \delta_{lr}\delta_{sv}\delta_{ur}\delta_{ks} + \underbrace{\delta_{ls}\delta_{ru}\delta_{vr}\delta_{ks}}_{=0, u>v} + \underbrace{\delta_{ls}\delta_{rv}\delta_{ur}\delta_{ks}}_{=0, u>v}$$

$$+ \delta_{lr} \underbrace{\delta_{su}\delta_{vs}\delta_{kr}}_{=0, u>v} + \underbrace{\delta_{lr}\delta_{sv}\delta_{us}\delta_{kr}}_{=0, u>v} - \delta_{ls}\delta_{ru}\delta_{vs}\delta_{kr} - \underbrace{\delta_{ls}\delta_{rv}\delta_{us}\delta_{kr}}_{=0, r>s \& u>v})$$

$$= -\frac{1}{2\sqrt{2}} \delta_{ur}\delta_{vs}(E_j)_{l,k}. \tag{2.11}$$

Adding together (2.9), (2.10) and (2.11) yields

$$\sum_{m=1}^{d^2} E_m E_j E_m = 0.$$

This implies for $d < j \leq \frac{d^2+d}{2}$

$$\Delta_d E_j = -2d E_j, \quad j > d, \tag{2.12}$$

and thus E_j is an eigenvector to the eigenvalue $-2d$. The computations for $\frac{d^2+d}{2} < j \leq d^2$ are analogous.

The eigenvectors Id_d and $E_j - E_k$, $1 \leq j, k \leq d$ span the same space as $\{E_j | 1 \leq j \leq d\}$. This means we have a basis of eigenvectors, which completes the proof of (iii). Note that in the case $d > 2$, the eigenfunctions we have chosen are not orthogonal.

(iv): First, assume $\mathrm{Tr}(\rho) = 0$. If we develop ρ along the eigenfunctions of Δ_d, the part along the identity vanishes (since all other eigenfunctions have trace 0). So ρ is in the eigenspace of Δ_d to the eigenvalue $-2d$, and it follows that

$$\begin{aligned} \|\nabla_d \rho\|^2_{\mathcal{H}^{d^2}} &= \langle \nabla_d \rho, \nabla_d \rho \rangle_{\mathcal{H}^{d^2}} = -\langle \rho, \Delta_d \rho \rangle_{\mathcal{H}} \\ &= 2d \langle \rho, \rho \rangle_{\mathcal{H}} = 2d \|\rho\|^2_{\mathcal{H}}. \end{aligned}$$

Now, if $\mathrm{Tr}(\rho) \neq 0$, we can write $\rho = \alpha \,\mathrm{Id} + \rho_0$, with $\mathrm{Tr}(\rho_0) = 0$, $\alpha = \frac{\mathrm{Tr}(\rho)}{d}$. Since $\nabla_d \,\mathrm{Id} = 0$ we obtain

$$\begin{aligned} \|\nabla_d \rho\|^2_{\mathcal{H}^{d^2}} &= \|\nabla_d \rho_0\|^2_{\mathcal{H}^{d^2}} = 2d \|\rho_0\|^2_{\mathcal{H}}. \\ &= 2d \langle \rho - \alpha \,\mathrm{Id}, \rho - \alpha \,\mathrm{Id} \rangle_{\mathcal{H}} \end{aligned}$$

$$
\begin{aligned}
&= 2d\Big(\|\rho\|_{\mathcal{H}}^2 - \alpha \underbrace{\langle \rho, \mathrm{Id}\rangle}_{=\mathrm{Tr}(\rho^\dagger)} - \overline{\alpha} \underbrace{\langle \mathrm{Id}, \rho\rangle}_{=\mathrm{Tr}(\rho)} + |\alpha|^2 \langle \mathrm{Id}, \mathrm{Id}\rangle\Big) \\
&= 2d\left(\|\rho\|_{\mathcal{H}}^2 - \tfrac{1}{d}|\mathrm{Tr}(\rho)|^2\right) \\
&= 2d\|\rho\|_{\mathcal{H}}^2 - 2|\mathrm{Tr}(\rho)|^2.
\end{aligned}
$$

□

Remarks: 1. There is a reason why this spectral gap gets larger with increasing dimension: While $\nabla_d \rho = 0$ is equivalent to $\rho = c\,\mathrm{Id}$, the same holds if one replaces $\nabla_d \rho$ with the discrete derivatives $(i[\rho, E_{\gamma(l)}])_{l=1,\ldots,l_{\max}}$ along a carefully chosen subset of $\{E_j | j = 1, \ldots, d^2\}$, where $l_{\max} \approx d$. If we were to define the operators ∇_d, div_d as the commutators with only such a subset of $\{E_j | j = 1, \ldots, d^2\}$ that contains d elements, the factor d vanishes. However, there is then no natural way to proceed in Theorem 2.9, as such a subset would not be a basis of $\mathcal{H}$.

2. For $d = 2$ the eigenfunctions of Δ_d are Id and the Pauli matrices

$$
\sigma_1 = E_1 - E_2,\ \sigma_2 = \sqrt{2}E_3,\ \sigma_3 = \sqrt{2}E_4, \tag{2.13}
$$

which fulfill

$$
[\sigma_j, \sigma_l] = c_{jl}\sigma_r, \tag{2.14}
$$

$l \neq r \neq j$, with $c_{jl} \in \{-1, 1\}$ for $j \neq l$ and $c_{jj} = 0$. This raises the valid question why we do not use these matrices as the basis for $\mathcal{H}$.
The answer is that they do not extend as nicely into $d > 2$. For $d = 3$ one gets the following 3 matrices as extensions for σ_1:

$$
\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
$$

As can be easily seen from the fact that their sum is 0, they span a subspace of dimension 2. Thus, for a basis, one would have to choose 2 matrices from that subspace. There is, however, no natural way to decide which matrices to choose. In fact, for $d = 3$, whatever subset one chooses for the basis, there is no "simple" relation to E_j^2 and $[E_j, E_k]$, $3 < j, k \leq 9$. With the basis we chose, the commutators between off-diagonal basis matrices, and the squares of off-diagonal matrices, can always be written as a sum of two diagonal basis matrices – a very homogeneous result, which we exploited in the computation of the eigenvectors to the Laplacian in Proposition 2.6.

Not all properties of continuous derivatives are shared by the discrete versions we introduced. In particular, the next lemma establishes that discrete partial derivatives do not commute:

Lemma 2.7. *Let*

$$\partial_d^{jl}\rho := i[i[\rho, E_j], E_l] = [E_l, [\rho, E_j]]$$

for $1 \le j, l \le d^2$ be the second partial discrete derivative of ρ with respect to j and l. Then for any $l, j \in \{1, \dots, d^2\}$, there exist $r \ge 1$, $1 \le k_1, \dots, k_r \le d^2$ and $s_1, \dots, s_r \in \mathbb{R}$ such that

$$\partial_d^{jl}\rho - \partial_d^{lj}\rho = \sum_{m=1}^{r} s_m \partial_d^{k_m}\rho := \sum_{m=1}^{r} s_m i[\rho, E_{k_m}].$$

That is, the commutator between two second partial derivatives is given by a linear combination of first order partial derivatives.

Proof: Compute

$$\begin{aligned}
& i[i[\rho, E_j], E_l] - i[i[\rho, E_l], E_j] \\
&= -\rho E_j E_l + E_l E_j \rho - E_l \rho E_j - E_j \rho E_l - \rho E_l E_j - E_j E_l \rho + E_j \rho E_l + E_l \rho E_j \\
&= -([E_l, E_j]\rho + \rho[E_j, E_l]) = i[\rho, i[E_l, E_j]].
\end{aligned}$$

Since $i[E_l, E_j]$ is hermitian, its coefficients along the basis $\{E_k\}$ are real. This completes our proof. $\square$

A direct consequence of Lemma 2.7 is that, unlike in the continuous case, in general it holds that $\mathrm{div}_d(\mathcal{R}\nabla_d\rho) \neq 0$ for skew-symmetric matrices $\mathcal{R} \in \mathbb{C}^{d^3 \times d^3}$.

Corollary 2.8. *Let $R = (r_{jk})_{j,k} \in \mathbb{R}^{d^2 \times d^2}$, $R = -R^T$, $\mathcal{R} := R \otimes \mathrm{Id}_d \in \mathbb{R}^{d^3 \times d^3}$. Then there exists $c \in \mathbb{R}^{d^2}$ such that for $\mathcal{C} := c \otimes \mathrm{Id}_d \in \mathbb{R}^{d^3 \times d}$ it holds that*

$$\mathrm{div}_d(\mathcal{R}\nabla_d\rho) = \mathrm{div}_d(\mathcal{C}\rho).$$

Proof: We compute

$$\mathrm{div}_d(\mathcal{R}\nabla_d\rho) = \mathrm{div}_d(\sum_{k=1}^{d^2} r_{jk} i[\rho, E_k])_{j=1,\dots,d^2}$$

$$= \sum_{j,k=1}^{d^2} r_{jk} i[i[\rho, E_k], E_j] = \sum_{j=1}^{d^2} \sum_{k=j+1}^{d^2} r_{jk} \left([E_j[\rho, E_k]] - [E_k, [\rho, E_j]]\right).$$

Now we can use Lemma 2.7 and obtain

$$\begin{aligned} \mathrm{div}_d(\mathcal{R}\nabla_d \rho) &= \sum_{j=1}^{d^2} \sum_{k=j+1}^{d^2} r_{jk} \left(\sum_{m=1}^{r(j,k)} s_m(j,k) i[\rho, E_{l_m(j,k)}] \right) \\ &= \sum_{j=1}^{d^2} \sum_{k=j+1}^{d^2} \sum_{m=1}^{r(j,k)} i[r_{jk} s_m(j,k)\rho, E_{l_m(j,k)}]. \end{aligned}$$

This can be rewritten as

$$\sum_{\tau=1}^{d^2} i[c_\tau \rho, E_\tau]$$

by adding together all the coefficients $r_{jk} s_m(j,k)$ for every $\tau = l_m(j,k)$, $\tau \in \{1, \dots, d^2\}$. □

2.3.2 Transformation of the equation

With the Hilbert space $\mathcal{H}$ and discrete derivatives established, we can now state our main result.

Theorem 2.9. *(2.1) can be written as*

$$\dot{\rho} = \mathcal{L}\rho = \mathrm{div}_d(\mathcal{D}\nabla_d \rho + \{\mathcal{F}, \rho\}_d). \tag{2.15}$$

Here $\mathcal{F} = \mathcal{F}(L_k, H)$ has hermitian components, $\mathcal{D} = \mathcal{D}(L_k)$ is real and positive semidefinite and

$$\{\mathcal{F}, \rho\}_d := (\{F_j, \rho\})_{j=1,\dots,d^2} = (F_j \rho + \rho F_j)_{j=1,\dots,d^2} \in \mathcal{H}^{d^2}$$

is the component-wise anticommutator.

Proof: The proof is structured into three parts. First, we discuss the transformation of a single Lindblad term, than of the Hamiltonian. Finally, we use the linearity of the equation to derive the general case from the first two.
(a) We start with a single Lindblad term

$$D_L(\rho) := [L\rho, L^\dagger] + [L, \rho L^\dagger]$$

Writing $L = \sum_{j=1}^{d^2} l_j E_j,\ l_j \in \mathbb{C}$, we compute

$$[L, \rho L^\dagger] = \sum_{j=1}^{d^2} l_j [E_j, \rho L^\dagger] = \sum_{j,k=1}^{d^2} l_j \overline{l_k} [E_j, \rho E_k],$$
$$[L\rho, L^\dagger] = \sum_{j=1}^{d^2} \overline{l_j} [L\rho, E_j] = -\sum_{j,k=1}^{d^2} \overline{l_j} l_k [E_j, E_k \rho].$$

Let $l_j \overline{l_k} =: d_{jk} + i f_{jk}$, then it follows that

$$\begin{aligned} D_L(\rho) &= \frac{1}{2}([L, \rho L^\dagger] + [L\rho, L^\dagger]) \\ &= \sum_{j,k=1}^{d^2} \left(\frac{1}{2}[d_{jk} + i f_{jk}][E_j, \rho E_k] - \frac{1}{2}[d_{jk} - i f_{jk}][E_j, E_k \rho] \right) \\ &= \sum_{j,k=1}^{d^2} \left(\frac{1}{2} d_{jk} [E_j, [\rho, E_k]] + \frac{1}{2} i f_{jk} [E_j, \{\rho, E_k\}] \right) \\ &= \mathrm{div}_d (\sum_{k=1}^{d^2} i \frac{d_{jk}}{2} [\rho, E_k] - \frac{f_{jk}}{2} \{\rho, E_k\})_{j=1,\ldots,n} \end{aligned}$$

Define the diffusion matrix $\mathcal{D}(L) := (\frac{d_{jk}}{2})_{j,k=1,\ldots,d^2} \otimes \mathrm{Id}_d \in \mathbb{C}^{d^3 \times d^3} \subset \mathcal{L}(\mathcal{H}^{d^2}, \mathcal{H}^{d^2})$, i.e. for $\mathcal{A} = (A_j)_{j=1,\ldots,d^2} \in \mathcal{H}^{d^2}$ it is

$$\mathcal{D}(L)\mathcal{A} := (\sum_{k=1}^{d^2} \frac{d_{jk}}{2} A_k)_{j=1,\ldots,d^2} \in \mathcal{H}^{d^2}.$$

Then we get

$$\begin{aligned} D_L(\rho) &= \mathrm{div}_d (\sum_{k=1}^{d^2} i d_{jk} [\rho, E_k] - f_{jk} \{\rho, E_k\})_{j=1,\ldots,d^2} \\ &= \mathrm{div}_d \left(\mathcal{D}(L) \nabla_d \rho - \sum_{k=1} \{\rho, f_{jk} E_k\} \right)_{j=1,\ldots,d^2} \\ &= \mathrm{div}_d \left(\mathcal{D}(L) \nabla_d \rho + \{\mathcal{F}(L), \rho\} \right) \end{aligned}$$

with the drift term $\mathcal{F}(L) := (F_j)_{j=1,\ldots,d^2} \in \mathcal{H}^{d^3},\ F_j = -\sum_{k=1}^{d^2} \frac{f_{jk}}{2} E_k$.

(b) Now we consider the hamiltionian part $i[H, \rho]$. Writing $H = \sum_{k=1}^{d^2} h_k E_k$,

$\alpha_k \in \mathbb{R}$, we compute

$$\begin{aligned} i[H,\rho] &= \sum_{k=1}^{d^2} h_k i[E_k,\rho] = -\sum_{k=1}^{d^2} i[h_k\rho, E_k] \\ &= \operatorname{div}_d(-h_k\rho)_{k=1,\ldots,d^2} = \operatorname{div}_d\{\mathcal{F}(H),\rho\} \end{aligned}$$

where $\mathcal{F}(H) = (-\frac{h_k}{2}\,\mathrm{Id})_{k=1,\ldots,d^2}$ is an additional drift term.
(c) For a general operator consisting of an Hamiltonian and m Lindblad operators, $D = i[H,\rho] + \sum_{j=1}^{m} D_{L_j}$, we obtain from (a), (b)

$$D(\rho) = -\operatorname{div}_d(\mathcal{D}(\nabla_d\rho) + \{\mathcal{F},\rho\})$$

where

$$\mathcal{D} = \sum_{j=1}^{m} \mathcal{D}(L_j), \quad \mathcal{F} = \sum_{j=1}^{m} \mathcal{F}(L_j) + \mathcal{F}(H).$$

$\mathcal{D}$ is positive semidefinite: Recall that $\mathcal{D} = (d_{jk})_{j,k} \otimes \mathrm{Id}_d$, so $\mathcal{D}$ is positive semidefinite iff $(d_{jk})_{j,k}$ is. However,

$$d_{jk} = \Re(l_j\overline{l_k}) = \tfrac{1}{2}(l_j\overline{l_k} + \overline{l_j}l_k) \Rightarrow D = \tfrac{1}{2}(l\otimes\bar{l} + \bar{l}\otimes l)$$

where $l = (l_j)_j$ is the vector formed by the l_j. We then get for any $u \in \mathbb{C}^{d^2}$, with $u^\dagger = \overline{u}^T$,

$$\begin{aligned} u^\dagger D u &= \tfrac{1}{2}(u^\dagger(l\otimes\bar{l})u + u^\dagger(\bar{l}\otimes l)u) \\ &= \tfrac{1}{2}(\|l^\dagger u\|_2^2 + \|l^T u\|_2^2) \geq 0. \end{aligned}$$

This implies that D is positive semidefinite. Finally, the components of $\mathcal{F}$ are hermitian since all h_k, f_{jk} are real. □

2.3.3 An example

In this section, we study a two-dimensional example for the transformation in Theorem 2.9. This example already highlights some of the differences to the continuous case. It also serves as a reference for the discussion of $\mathcal{L}$ on weighted spaces in Section 2.4.

Proposition 2.10. *Let $H = 0$, $L_1 = iE_1 + \frac{1}{2}E_3$, and*

$$D = i[H,\rho] + \frac{1}{2}[L_1, \rho L_1^\dagger] + [L_1\rho, L_1^\dagger].$$

Then for the transformed operator from Theorem 2.9, we have that

(i)

$$\mathcal{D} = \operatorname{diag}(\frac{1}{2}, 0, \frac{1}{8}, 0) \otimes \mathrm{Id}_2 \in \mathbb{R}^{8\times 8},$$
$$\mathcal{F} = \begin{pmatrix} -\frac{1}{4}E_3 \\ 0 \\ \frac{1}{4}E_1 \\ 0 \end{pmatrix} \in \mathbb{C}^{8\times 2}.$$

(ii)

$$\rho_\infty = \frac{1}{10}\begin{pmatrix} 1 & 2\sqrt{2}i \\ -2\sqrt{2}i & 9 \end{pmatrix} \tag{2.16}$$

spans the kernel of $\mathcal{L}$.

Proof: (i): From the proof of Theorem 2.9, we obtain with $l_1 = i$, $l_3 = \frac{1}{2}$:

$$d_{11} = \frac{1}{2}\Re\{i\bar{i}\} = \frac{1}{2}, \quad d_{33} = \frac{1}{2}\Re\{\frac{1}{2}\overline{\frac{1}{2}}\} = \frac{1}{8},$$
$$d_{13} = \frac{1}{2}\Re\{i\overline{\frac{1}{2}}\} = 0 = d_{31}, f_{13} = \frac{1}{2}\Im\{i\overline{\frac{1}{2}}\} = \frac{1}{4} = -f_{31}.$$

(ii): For a stationary state, we make the ansatz $\rho_\infty = \sum_{j=1}^{4} \alpha_j E_j$, $\alpha_j \in \mathbb{C}$. We expect to find that $\alpha_j \in \mathbb{R}$ holds, since a stationary state should be a density matrix and thus hermitian.
Compute

$$\mathcal{D}\nabla_d\rho = \begin{pmatrix} i\frac{1}{2}[\rho, E_1] \\ 0 \\ \frac{i}{8}[\rho, E_3] \\ 0 \end{pmatrix}$$
$$= \frac{1}{2}\begin{pmatrix} \alpha_3 E_4 - \alpha_4 E_3 \\ 0 \\ \frac{\alpha_2-\alpha_1}{4}E_4 + \frac{\alpha_4}{4}(E_1 - E_2) \\ 0 \end{pmatrix} \tag{2.17}$$

and

$$\{\mathcal{F},\rho\}_d = \begin{pmatrix} -\frac{1}{4}\{\rho,E_3\} \\ 0 \\ \frac{1}{4}\{\rho,E_1\} \\ 0 \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} -\frac{\alpha_1+\alpha_2}{2}E_3 - \frac{\alpha_3}{2}(E_1+E_2) \\ 2\rho \\ \alpha_1 E_1 + \frac{\alpha_3}{2}E_3 + \frac{\alpha_4}{2}E_4 \\ 0 \end{pmatrix}. \tag{2.18}$$

From (2.17), (2.18) we obtain

$$\mathcal{D}\nabla_d\rho + \{\mathcal{F},\rho\}_d = \frac{1}{2}\begin{pmatrix} -\frac{\alpha_3}{2}E_1 - \frac{\alpha_3}{2}E_2 - \frac{\alpha_1+\alpha_2+2\alpha_4}{2}E_3 + \alpha_3 E_4 \\ 0 \\ \frac{4\alpha_1+\alpha_4}{4}E_1 - \frac{\alpha_4}{4}E_2 + \frac{\alpha_3}{2}E_3 + \frac{\alpha_1-\alpha_2+2\alpha_4}{4}E_4 \\ 0 \end{pmatrix}.$$

If $\mathcal{D}\nabla_d\rho + \{\mathcal{F},\rho\}_d = 0$, it follows from the linear independence of the E_j that $\rho = 0$. So the stationary state does not fulfil $\mathcal{D}\nabla_d\rho_\infty + \{\mathcal{F},\rho_\infty\} = 0$. Compute

$$\begin{aligned}
&\mathrm{div}_d(\mathcal{D}\nabla_d\rho + \{\mathcal{F},\rho\}_d) \\
&= \frac{i}{2}[\alpha_3(E_4 - \frac{1}{2}(E_1-E_2)) - \frac{\alpha_1+\alpha_2+2\alpha_4}{2}E_3, E_1] \\
&\quad + \frac{i}{2}[\frac{4\alpha_1+\alpha_4}{4}E_1 - \frac{\alpha_4}{4}E_2 + \frac{\alpha_3}{2}E_3 + \frac{\alpha_2-\alpha_1+2\alpha_4}{4}E_4, E_3] \\
&= -\frac{\alpha_3}{2}E_3 - \frac{\alpha_1+\alpha_2+2\alpha_4}{4}E_4 \\
&\quad - \frac{4\alpha_1+\alpha_4}{8}E_4 - \frac{\alpha_4}{8}E_4 + \frac{\alpha_2-\alpha_1+2\alpha_4}{8}(E_1-E_2) \\
&= \frac{-\alpha_1+\alpha_2+2\alpha_4}{8}(E_1-E_2) - \frac{\alpha_3}{2}E_3 - \frac{3\alpha_1+\alpha_2+3\alpha_4}{4}E_4.
\end{aligned}$$

With the restraint $\alpha_1+\alpha_2 = 1$ from trace normalisation, this has the unique solution

$$\alpha_1 = \frac{1}{10}, \quad \alpha_2 = \frac{9}{10}, \quad \alpha_4 = -\frac{2}{5}.$$

So

$$\rho_\infty = \alpha_1 E_1 + \alpha_2 E_2 + \alpha_4 E_4,$$

which completes the proof. □

This shows two important properties of $\mathcal{L}$: First, we can in general not expect that $\mathcal{D}\nabla_d\rho_\infty + \{\mathcal{F}, \rho_\infty\}_d = 0$, and second, there may be a unique normalised stationary state even if $\mathcal{D}$ does not have full rank. Neither should be surprising, as it is the same for the classic equation, at least in the hypocoercive case discussed in §1. As we will see in the next sections, the first property appears fundamental to our approach; see Conjecture 2.14. The fact that $\mathcal{D}$ need not have full rank for uniqueness of a stationary state should also be expected from the first remark following Theorem 2.6.

2.4 Symmetry, weighted Hilbert space

In this section, we investigate some properties of the transformed equation (2.15) on the Hilbert space $\mathcal{H}$. In the classic case, the natural space to investigate the Fokker-Planck operator is the weighted Hilbert space $L^2(\mathbb{R}^d, f_\infty^{-1})$ (see §1). We discuss the problems that arise in transferring this idea to the operator $\mathcal{L}$, and conditions on $\mathcal{F}$ and $\mathcal{D}$ compared to the symmetric classic operator.

In particular, we provide some results indicating that the only case where $\mathcal{L}$ is symmetric on a weighted Hilbert space is the case $\rho_\infty = \frac{1}{d}\,\mathrm{Id}$ – but then the weight ρ_∞^{-1} has no effect. This indicates that in the quantum mechanical case, the operator $\mathcal{L}$ is essentially non-symmetric in the "interesting" case where its kernel is not spanned by Id. This can be compared to the hypocoercive case in §1, where the non-symmetry of the operator on the weighted space was also essential.

Our first result is for the case $\operatorname{div}\mathcal{F} = 0$, which arises for example when all the Lindblad operators are either hermitian or skew-hermitian.

Proposition 2.11. *Consider the operator $\mathcal{L}$ from (2.15) with* $\operatorname{div}\mathcal{F} = 0$*. Then*

(i) $\mathcal{L}$ is symmetric on $\mathcal{H}$ and $\langle \rho, \mathcal{L}(\rho)\rangle_{\mathcal{H}} \le 0$.

(ii) Id_d is in the kernel of $\mathcal{L}$.

(iii) If $\mathcal{D} \ge \lambda\,\mathrm{Id}_{d^3}$, then the kernel of $\mathcal{L}$ is spanned by Id_d. Furthermore, the solution $\rho(t)$ to (2.15) with hermitian initial condition $0 \le \rho_0 \in \mathbb{C}^{d\times d}$, $\mathrm{Tr}(\rho_0) = 1$, converges to $\rho_\infty = \frac{1}{d}\,\mathrm{Id}$ exponentially with rate at least $2d\lambda$.

Proof: (i): We compute, using $\nabla_d(\rho^2) = \{\rho, \nabla_d \rho\}_d$,

$$\begin{aligned}
\langle \rho, \mathcal{L}(\rho)\rangle_{\mathcal{H}} &= \langle \rho, \mathrm{div}_d(\mathcal{D}\nabla_d \rho + \{\mathcal{F}, \rho\}_d)\rangle_{\mathcal{H}} \\
&= -\langle \nabla_d \rho, \mathcal{D}\nabla_d \rho\rangle_{\mathcal{H}^{d^2}} + \langle \rho, \mathrm{div}_d\{\mathcal{F}, \rho\}_d\rangle_{\mathcal{H}} \\
\langle \rho, \mathrm{div}_d\{\mathcal{F}, \rho\}\rangle_{\mathcal{H}} &= \mathrm{Tr}(\rho \,\mathrm{div}_d\{\mathcal{F}, \rho\}_d) = 2\mathrm{Tr}(\rho^2 \,\mathrm{div}_d \mathcal{F}) + \mathrm{Tr}(\mathcal{F}\{\rho, \nabla_d \rho\}_d) \\
&= 2\mathrm{Tr}(\rho^2 \,\mathrm{div}_d \mathcal{F}) + \mathrm{Tr}(\mathcal{F}\nabla_d \rho^2) \\
&= \mathrm{Tr}(\rho^2 \,\mathrm{div}_d \mathcal{F}) = 0.
\end{aligned}$$

This implies

$$\langle \rho, L(\rho)\rangle_{\mathcal{H}} = -\langle \nabla_d \rho, \mathcal{D}\nabla_d \rho\rangle_{\mathcal{H}^{d^2}} \le 0,$$

since $\mathcal{D} \ge 0$. Since $\mathcal{D} = \mathcal{D}^T \in \mathbb{R}^{d^3 \times d^3}$, $\mathcal{L}$ is self-adjoint.

(ii): Compute

$$\mathcal{L}(\mathrm{Id}_d) = \mathrm{div}_d\left(\mathcal{D}\nabla_d(\mathrm{Id}_d) + \{\mathcal{F}, \mathrm{Id}_d\}_d\right) = 2\,\mathrm{div}_d \mathcal{F} = 0.$$

(iii): We know that the solution fulfils $\rho(t) \ge 0$, is hermitian and has $\mathrm{Tr}(\rho(t)) = \mathrm{Tr}(\rho_0) = 1$. We compute

$$\begin{aligned}
\|\rho(t) - \tfrac{1}{d}\,\mathrm{Id}\|_{\mathcal{H}}^2 &= \mathrm{Tr}((\rho(t) - \tfrac{1}{d}\,\mathrm{Id})^2) = \mathrm{Tr}(\rho(t)^2 - \tfrac{2}{d}\rho(t) + \tfrac{1}{d^2}\,\mathrm{Id}) \\
&= \|\rho(t)\|_{\mathcal{H}}^2 - \tfrac{1}{d}.
\end{aligned}$$

With the discrete Poincaré inequality (2.4), this implies

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{dt}}\|\rho(t) - \tfrac{1}{d}\,\mathrm{Id}\|_{\mathcal{H}}^2 &= 2\langle \rho(t), D(\rho(t))\rangle_{\mathcal{H}} \\
&= -2\langle \nabla_d \rho(t), \mathcal{D}\nabla_d \rho(t)\rangle_{\mathcal{H}^{n^2}} \le -2\lambda\|\nabla_d \rho(t)\|_{\mathcal{H}^{d^2}}^2 \\
&= -4d\lambda\|\rho(t)\|_{\mathcal{H}}^2 + 4\lambda|\mathrm{Tr}(\rho(t))|^2 = -4d\lambda\|\rho(t)\|_{\mathcal{H}}^2 + 4\lambda \\
&= -4d\lambda(\|\rho(t) - \tfrac{1}{d}\,\mathrm{Id}\|_{\mathcal{H}}^2).
\end{aligned}$$

Applying Gronwall's Lemma yields

$$\|\rho(t) - \tfrac{1}{d}\,\mathrm{Id}\|_{\mathcal{H}} \le \|\rho_0 - \tfrac{1}{d}\,\mathrm{Id}\|_{\mathcal{H}} \exp(-2d\lambda t).$$

In particular, this implies that the kernel of $\mathcal{L}$ is spanned by Id. Thus the proof is complete. $\square$

Remark: As mentioned before, the condition $\mathcal{D} \geq \lambda \operatorname{Id}_{d^3}$ is stricter than necessary.

Next, we consider the case where there is a unique, positive definite normalised steady state $\rho_\infty \neq \operatorname{Id}$. Since $\rho_\infty > 0$, there exists a positive definite and symmetric square root $\rho_\infty^{\frac{1}{2}}$ of ρ_∞. From Proposition 2.11, it follows that $\operatorname{div} \mathcal{F} \neq 0$. We aim to compute the adjoint of $\mathcal{L}$ on the weighted space $\mathcal{H}(\rho_\infty^{-1})$. However, since the elements of $\mathcal{H}$ are non-commutative, it is not obvious how to insert the weight into the scalar product:

Lemma 2.12. *Let $\rho_\infty > 0$ be hermitian. Define the scalar products*

$$\begin{aligned}\langle \rho, \eta \rangle_1 &:= \operatorname{Tr}(\rho \rho_\infty^{-1} \eta^\dagger), \\ \langle \rho, \eta \rangle_2 &:= \operatorname{Tr}(\rho \eta^\dagger \rho_\infty^{-1}), \\ \langle \rho, \eta \rangle_{\rho_\infty^{-1}} &:= \operatorname{Tr}(\rho \rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}}).\end{aligned}$$

Then the following holds:

(i) There exist hermitian matrices ρ, η such that

$$\langle \rho, \eta \rangle_1 \notin \mathbb{R}, \quad \langle \rho, \eta \rangle_2 \notin \mathbb{R}. \tag{2.19}$$

(ii) For all hermitian matrices ρ, η: $\langle \rho, \eta \rangle_{\rho_\infty^{-1}} \in \mathbb{R}$.

Proof: (i): We only need to find one example. So take the stationary state from Proposition 2.10. Then

$$\rho_\infty^{-1} = 100(\frac{9}{10} E_1 + \frac{1}{10} E_2 + \frac{2}{5} E_4).$$

Now let $\rho := \frac{1}{2} \operatorname{Id}_2 + E_3$, $\eta := E_4$. Then

$$\begin{aligned}\langle \rho, \eta \rangle_1 &= \operatorname{Tr}(\rho_\infty^{-1} E_4) + \operatorname{Tr}(E_4 E_3 \rho_\infty^{-1}) = -10 + 20i, \\ \langle \rho, \eta \rangle_2 &= \operatorname{Tr}(E_4 \rho_\infty^{-1}) + \operatorname{Tr}(E_3 E_4 \rho_\infty^{-1}) = -10 - 20i.\end{aligned}$$

(ii): Let ρ, η be hermitian. Then

$$\langle \rho, \eta \rangle_{\rho_\infty^{-1}} = \operatorname{Tr}(\rho \rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}}) = \operatorname{Tr}(\rho_\infty^{-\frac{1}{4}} \rho \rho_\infty^{-\frac{1}{2}} \eta \rho_\infty^{-\frac{1}{4}}).$$

Since $\operatorname{Tr}(A^T) = \operatorname{Tr}(A)$ and $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$, it follows that

$$\operatorname{Tr}\left(\rho_\infty^{-\frac{1}{4}} \rho \rho_\infty^{-\frac{1}{2}} \eta \rho_\infty^{-\frac{1}{4}}\right) = \operatorname{Tr}\left(\rho_\infty^{-\frac{1}{4}} \eta \rho_\infty^{-\frac{1}{2}} \rho \rho_\infty^{-\frac{1}{4}}\right) = \operatorname{Tr}\left(\left[\rho_\infty^{-\frac{1}{4}} \rho \rho_\infty^{-\frac{1}{2}} \eta \rho_\infty^{-\frac{1}{4}}\right]^\dagger\right)$$

$$= \overline{\operatorname{Tr}\left(\rho_\infty^{-\frac{1}{4}}\rho\rho_\infty^{-\frac{1}{2}}\eta\rho_\infty^{-\frac{1}{4}}\right)}.$$

Here, $A^\dagger$ denotes the hermitian adjoint of A. This implies

$$\langle \rho, \eta \rangle_{\rho_\infty^{-1}} \in \mathbb{R}.$$

□

The scalar products $\langle\cdot,\cdot\rangle_1$, $\langle\cdot,\cdot\rangle_2$ are unsuited to our agenda, since there is no guarantee that

$$\langle \rho, \mathcal{L}_1\rho\rangle_1 \in \mathbb{R}$$

for some operator $\mathcal{L}_1$ such that $\mathcal{L}_1\rho$ is hermitian. But this is exactly what we would expect for the self-adjoint part of $\mathcal{L}$ on $(\mathcal{H}, \langle\cdot,\cdot\rangle_1)$. So we choose the weighted Hilbert space $\mathcal{H}(\rho_\infty^{-1}) := (\mathcal{H}, \langle\cdot,\cdot\rangle_{\rho_\infty^{-1}})$. This space is also used for non-commutative logarithmic Sobolev inequalities, see for example [60], [54].

We obtain the following result for $\mathcal{L}$ on $\mathcal{H}(\rho_\infty^{-1})$:

Proposition 2.13. *Assume that the kernel of the operator $\mathcal{L}$ from (2.15) is spanned by ρ_∞. Let*

$$\mathcal{K} := \mathcal{F} - \mathcal{D}\mathcal{V}, \quad \mathcal{V} := (\rho_\infty^{\frac{1}{2}} i[\rho_\infty^{-\frac{1}{2}}, E_j])_{j=1,\dots,d^2} = (\rho_\infty^{\frac{1}{2}}(\nabla_d\rho_\infty^{-\frac{1}{2}})_j)_{j=1,\dots,d^2}.$$

Then

(i)

$$\mathcal{L}\rho = \operatorname{div}_d\left(\rho_\infty^{\frac{1}{2}}\left[\mathcal{D}\nabla_d(\rho_\infty^{-\frac{1}{2}}\rho\rho_\infty^{-\frac{1}{2}})\right]\rho_\infty^{\frac{1}{2}} + \mathcal{K}\rho + \rho\mathcal{K}^\dagger\right).$$

(ii) On $\mathcal{H}(\rho_\infty^{-1})$, the adjoint of $\mathcal{L}$ is given by

$$\begin{aligned}\mathcal{L}^*\rho = \operatorname{div}_d &\left(\rho_\infty^{\frac{1}{2}}\left[\mathcal{D}\nabla_d(\rho_\infty^{-\frac{1}{2}}\rho\rho_\infty^{-\frac{1}{2}})\right]\rho_\infty^{\frac{1}{2}}\right)\\ &- \left[\rho_\infty^{\frac{1}{2}}\mathcal{K}^\dagger\nabla_d(\rho_\infty^{-\frac{1}{2}}\rho\rho_\infty^{-\frac{1}{2}})\rho_\infty^{\frac{1}{2}} + \rho_\infty^{\frac{1}{2}}\nabla_d(\rho_\infty^{-\frac{1}{2}}\rho\rho_\infty^{-\frac{1}{2}})\mathcal{K}\rho_\infty^{\frac{1}{2}}\right].\end{aligned}$$

Proof:
(i): We compute

$$\begin{aligned}&\operatorname{div}_d(\mathcal{D}\nabla_d\rho + \mathcal{F}\rho + \rho\mathcal{F}) = \operatorname{div}_d(\mathcal{D}\rho_\infty^{\frac{1}{2}}\rho_\infty^{-\frac{1}{2}}(\nabla_d\rho)\rho_\infty^{-\frac{1}{2}}\rho_\infty^{\frac{1}{2}} + \mathcal{F}\rho + \rho\mathcal{F})\\ &= \operatorname{div}_d\left(\rho_\infty^{\frac{1}{2}}\left[\mathcal{D}\nabla_d(\rho_\infty^{-\frac{1}{2}}\rho\rho_\infty^{-\frac{1}{2}})\right]\rho_\infty^{\frac{1}{2}} + [\mathcal{F} - \mathcal{D}\rho_\infty^{\frac{1}{2}}(\nabla_d\rho_\infty^{-\frac{1}{2}})]\rho + \rho[\mathcal{F} - (\mathcal{D}\nabla_d\rho_\infty^{-\frac{1}{2}})\rho_\infty^{\frac{1}{2}}]\right)\end{aligned}$$

$$= \mathrm{div}_d \left(\rho_\infty^{\frac{1}{2}} \left[\mathcal{D} \nabla_d (\rho_\infty^{-\frac{1}{2}} \rho \rho_\infty^{-\frac{1}{2}}) \right] \rho_\infty^{\frac{1}{2}} + \mathcal{K}\rho + \rho \mathcal{K}^\dagger \right).$$

(ii) We split the computation of the adjoint of $\mathcal{L}$ on $\mathcal{H}(\rho_\infty^{-1})$ into two parts. First,

$$\begin{aligned}
&\langle \mathrm{div}_d \left(\rho_\infty^{\frac{1}{2}} \left[\mathcal{D} \nabla_d (\rho_\infty^{-\frac{1}{2}} \rho \rho_\infty^{-\frac{1}{2}}) \right] \rho_\infty^{\frac{1}{2}} \right), \eta \rangle_{\rho_\infty^{-1}} \\
&= \mathrm{Tr} \left(\mathrm{div}_d \left(\rho_\infty^{\frac{1}{2}} \left[\mathcal{D} \nabla_d (\rho_\infty^{-\frac{1}{2}} \rho \rho_\infty^{-\frac{1}{2}}) \right] \rho_\infty^{\frac{1}{2}} \right) \rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}} \right) \\
&= -\mathrm{Tr} \left(\left(\rho_\infty^{\frac{1}{2}} \left[\mathcal{D} \nabla_d (\rho_\infty^{-\frac{1}{2}} \rho \rho_\infty^{-\frac{1}{2}}) \right] \rho_\infty^{\frac{1}{2}} \right) \cdot \nabla_d (\rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}}) \right) \\
&= -\mathrm{Tr} \left(\nabla_d (\rho_\infty^{-\frac{1}{2}} \rho \rho_\infty^{-\frac{1}{2}}) \cdot \left(\rho_\infty^{\frac{1}{2}} \left[\mathcal{D} \nabla_d (\rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}}) \right] \rho_\infty^{\frac{1}{2}} \right) \right) \\
&= \mathrm{Tr} \left(\rho_\infty^{-\frac{1}{2}} \rho \rho_\infty^{-\frac{1}{2}} \, \mathrm{div}_d \left(\rho_\infty^{\frac{1}{2}} \left[\mathcal{D} \nabla_d (\rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}}) \right] \rho_\infty^{\frac{1}{2}} \right) \right) \\
&= \langle \rho, \mathrm{div}_d \left(\rho_\infty^{\frac{1}{2}} \left[\mathcal{D} \nabla_d (\rho_\infty^{-\frac{1}{2}} \eta \rho_\infty^{-\frac{1}{2}}) \right] \rho_\infty^{\frac{1}{2}} \right) \rangle_{\rho_\infty^{-1}}.
\end{aligned}$$

Further,

$$\begin{aligned}
&\langle \mathrm{div}_d (\mathcal{K}\rho + \rho\mathcal{K}^\dagger), \eta \rangle_{\rho_\infty^{-1}} \\
&= \mathrm{Tr} \left(\mathrm{div}_d (\mathcal{K}\rho + \rho\mathcal{K}^\dagger) \rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}} \right) \\
&= -\mathrm{Tr} \left([\mathcal{K}\rho + \rho\mathcal{K}^\dagger] \nabla_d (\rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}}) \right) \\
&= -\mathrm{Tr} \left(\rho \left[\mathcal{K}^\dagger \nabla_d (\rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}}) + \nabla_d (\rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}}) \mathcal{K} \right] \right) \\
&= -\mathrm{Tr} \left(\rho_\infty^{-\frac{1}{2}} \rho \rho_\infty^{-\frac{1}{2}} \left[\rho_\infty^{\frac{1}{2}} \mathcal{K}^\dagger \nabla_d (\rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}}) \rho_\infty^{\frac{1}{2}} + \rho_\infty^{\frac{1}{2}} \nabla_d (\rho_\infty^{-\frac{1}{2}} \eta^\dagger \rho_\infty^{-\frac{1}{2}}) \mathcal{K} \rho_\infty^{\frac{1}{2}} \right] \right) \\
&= -\langle \rho, \rho_\infty^{\frac{1}{2}} \mathcal{K}^\dagger \nabla_d (\rho_\infty^{-\frac{1}{2}} \eta \rho_\infty^{-\frac{1}{2}}) \rho_\infty^{\frac{1}{2}} + \rho_\infty^{\frac{1}{2}} \nabla_d (\rho_\infty^{-\frac{1}{2}} \eta \rho_\infty^{-\frac{1}{2}}) \mathcal{K} \rho_\infty^{\frac{1}{2}} \rangle_{\rho_\infty^{-1}}.
\end{aligned}$$

We obtain

$$\begin{aligned}
\mathcal{L}^* \rho = {} & \mathrm{div}_d \left(\rho_\infty^{\frac{1}{2}} \left[\mathcal{D} \nabla_d (\rho_\infty^{-\frac{1}{2}} \rho \rho_\infty^{-\frac{1}{2}}) \right] \rho_\infty^{\frac{1}{2}} \right) \\
& - \left[\rho_\infty^{\frac{1}{2}} \mathcal{K}^\dagger \nabla_d (\rho_\infty^{-\frac{1}{2}} \rho \rho_\infty^{-\frac{1}{2}}) \rho_\infty^{\frac{1}{2}} + \rho_\infty^{\frac{1}{2}} \nabla_d (\rho_\infty^{-\frac{1}{2}} \rho \rho_\infty^{-\frac{1}{2}}) \mathcal{K} \rho_\infty^{\frac{1}{2}} \right].
\end{aligned}$$

□

Proposition 2.13 implies that $\mathcal{L}$ is self-adjoint on $\mathcal{H}(\rho_\infty^{-1})$ if $\mathcal{K} = 0$. But we have

Proposition 2.14. *Let $\mathcal{L}$ be given as in Proposition (2.13). Assume $\mathcal{D} = \mathrm{Id}_{d^3}$. Then $\mathcal{K} = 0$ iff $\mathcal{F} = 0$ and $\rho_\infty = c\,\mathrm{Id}$ for some $c \in \mathbb{R}$.*

Proof: If $\mathcal{K} = 0$, then it follows that

$$0 = \mathcal{K} - \mathcal{K}^\dagger = \mathcal{F} - \mathcal{V} - \mathcal{F} + \mathcal{V}^\dagger.$$

This implies $0 = \mathcal{V} - \mathcal{V}^\dagger$, so

$$0 = i\rho_\infty^{\frac{1}{2}}[\rho_\infty^{-\frac{1}{2}}, E_j] - i[\rho_\infty^{-\frac{1}{2}}, E_j]\rho_\infty^{\frac{1}{2}} = i[[\rho_\infty^{-\frac{1}{2}}, E_j], \rho_\infty^{\frac{1}{2}}].$$

for $j = 1, \ldots, d^2$. It follows that

$$\begin{aligned} 0 &= -i\mathrm{Tr}(i[[\rho_\infty^{-\frac{1}{2}}, E_j], \rho_\infty^{\frac{1}{2}}]\rho_\infty^{-\frac{1}{2}} E_j \rho_\infty^{-\frac{1}{2}}) = \mathrm{Tr}([\rho_\infty^{-\frac{1}{2}}, E_j](E_j\rho_\infty^{-\frac{1}{2}} - \rho_\infty^{-\frac{1}{2}} E_j)) \\ &= \mathrm{Tr}([\rho_\infty^{-\frac{1}{2}}, E_j][E_j, \rho_\infty^{-\frac{1}{2}}]) = \mathrm{Tr}((i[\rho_\infty^{-\frac{1}{2}}, E_j])(i[\rho_\infty^{-\frac{1}{2}}, E_j])^\dagger) = \|\nabla_d \rho_\infty^{-\frac{1}{2}}\|^2_{\mathcal{H}^{d^2}}. \end{aligned}$$

This implies $\nabla_d \rho_\infty^{-\frac{1}{2}} = 0$, so $\rho_\infty^{-\frac{1}{2}} = \tilde{c}\,\mathrm{Id}$. □

It follows that the situation of the example from Proposition 2.10, where the stationary state did not fulfil $D\nabla_d\rho_\infty + \mathcal{F}\rho_\infty + \rho_\infty\mathcal{F} = 0$, is the general case for non-constant stationary states:

Corollary 2.15. *Let ρ_∞ be invertible with $\nabla_d\rho_\infty + \mathcal{F}\rho_\infty + \rho_\infty\mathcal{F} = 0$. Then $\rho_\infty = \mathrm{Id}_d$, $\mathcal{F} = 0$.*

Proof: Compute

$$\begin{aligned} \nabla_d\rho_\infty + \mathcal{F}\rho_\infty + \rho_\infty\mathcal{F} &= \rho_\infty^{\frac{1}{2}}\nabla_d\rho_\infty^{\frac{1}{2}} + \rho_\infty\mathcal{F} + (\nabla_d\rho_\infty^{\frac{1}{2}})\rho_\infty^{\frac{1}{2}} + \mathcal{F}\rho_\infty \\ &= \rho_\infty[\rho_\infty^{-\frac{1}{2}}\nabla_d\rho_\infty^{\frac{1}{2}} + \mathcal{F}] + [(\nabla_d\rho_\infty^{\frac{1}{2}})\rho_\infty^{-\frac{1}{2}} + \mathcal{F}]\rho_\infty \\ &= \rho_\infty[\mathcal{F} - (\nabla_d\rho_\infty^{-\frac{1}{2}})\rho_\infty^{\frac{1}{2}}] + [\mathcal{F} - \rho_\infty^{\frac{1}{2}}\nabla_d\rho_\infty^{-\frac{1}{2}}]\rho_\infty \\ &= \rho_\infty\mathcal{K}^\dagger + \mathcal{K}\rho_\infty. \end{aligned}$$

Since ρ_∞ is invertible and positive definite, we conclude that $\rho_\infty\mathcal{K}^\dagger + \mathcal{K}\rho_\infty = 0$ iff $\mathcal{K} = 0$. Applying Proposition 2.14 completes our proof. □

So the only case where $\mathcal{L}$ is symmetric on $\mathcal{H}(\rho_\infty^{-1})$ is the case $\rho_\infty = c\,\mathrm{Id}$, which defeats the purpose of considering $\mathcal{H}(\rho_\infty^{-1})$ instead of $\mathcal{H}$. This means that, if we can transfer the entropy method, we have to do it for the more complicated non-symmetric case.

2.5 The case $d = 2$

In this section, we consider the lowest-dimensional case, $d = 2$. Some properties of $\mathcal{L}$ can already be seen on this level, and the results give some indication what to expect for higher dimensions. This case has, of course, been studied before for (2.1). See for example [20], [21], which discuss the case $d = 2$ from the standpoint of (2.1), with a focus on decoherence.

For $d = 2$, our basis is

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$
$$E_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad E_4 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

We are interested in investigating long-term behaviour of solutions and existence of stationary states for (2.15):

$$\rho_t = \mathrm{div}_d(\mathcal{D}\nabla_d \rho + \{\mathcal{F}, \rho\}).$$

For ease of computation, and since we are mostly interested in the role of $\mathcal{F}$, we assume $\mathcal{D} = \mathrm{diag}(d_{11}, d_{22}, d_{33}, d_{44}) \otimes \mathrm{Id}_2$. We also assume $H = 0$, which means that $\mathcal{F}$ will have no coefficients along Id_2 in any component (see the proof of Theorem 2.9). In light of Proposition 2.6, we develop a density matrix ρ with $\mathrm{Tr}(\rho) = 1$ along the eigenfunctions of Δ_d, which are Id and the Pauli matrices (2.13). The condition $\mathrm{Tr}(\rho) = 1$ implies $\rho(t) = \frac{1}{2}\,\mathrm{Id} + \rho_0(t)$, since the Pauli matrices have trace 0. We obtain:

Proposition 2.16. *A density matrix*

$$\rho(t) = \frac{1}{2}\,\mathrm{Id}_2 + \alpha_1(t)\sigma_1 + \alpha_2(t)\sigma_2 + \alpha_3(t)\sigma_3$$

is a solution to (2.15) iff

$$\frac{d}{dt}\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = (M + R)\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} + \beta,$$

where $\beta = 2\,\mathrm{div}_d(\mathcal{F}) \in \mathbb{R}^3$, $M = M(\mathcal{D})$ is a negative semidefinite diagonal matrix and $R = R(\mathcal{F})$ is antisymmetric.

Proof: We compute

$$\nabla_d \rho = \begin{pmatrix} i[\rho, E_1] \\ i[\rho, E_2] \\ i[\rho, E_3] \\ i[\rho, E_4] \end{pmatrix}$$
$$= \begin{pmatrix} \alpha_2\sigma_3 - \alpha_3\sigma_2 \\ -\alpha_2\sigma_3 + \alpha_3\sigma_2 \\ \sqrt{2}\alpha_3\sigma_1 - \sqrt{2}\alpha_1\sigma_3 \\ -\sqrt{2}\alpha_2\sigma_1 + \sqrt{2}\alpha_1\sigma_2 \end{pmatrix}.$$

This implies

$$\begin{aligned}&\operatorname{div}_d(\mathcal{D}\nabla_d\rho) = d_{11}[E_1,[\rho,E_1]] + d_{22}[E_2,[\rho,E_2]] + d_{33}[E_3,[\rho,E_3]] + d_{44}[E_4,[\rho,E_4]]\\ &= -2(d_{33}+d_{44})\alpha_1\sigma_1 - (d_{11}+d_{22}+2d_{44})\alpha_2\sigma_2 - (d_{11}+d_{22}+2d_{33})\alpha_3\sigma_3\end{aligned}$$

Introducing the notation $d_1 := 2(d_{33}+d_{44})$, $d_2 := d_{11}+d_{22}+2d_{44}$ and $d_3 := d_{11}+d_{22}+2d_{33}$, this simplifies to

$$\operatorname{div}_d(\mathcal{D}\nabla_d\rho) = -d_1\alpha_1\sigma_1 - d_2\alpha_2\sigma_2 - d_3\alpha_3\sigma_3. \tag{2.20}$$

This exemplifies the remark below Proposition 2.6: The operator is fully dissipative even if $\mathcal{D}$ is not regular and $\mathcal{F}=0$, since Δ_d contains more terms than "necessary".

We now compute $\operatorname{div}_d(\{\mathcal{F},\rho\}_d)$. In general,

$$\mathcal{F} = \begin{pmatrix} F_1\\ F_2\\ F_3\\ F_4\end{pmatrix} = \begin{pmatrix} f_{12}E_2 + f_{13}E_3 + f_{14}E_4\\ -f_{12}E_1 + f_{23}E_3 + f_{24}E_4\\ -f_{13}E_1 - f_{23}E_2 + f_{34}E_4\\ -f_{14}E_1 - f_{24}E_2 - f_{34}E_3\end{pmatrix},$$

$f_{jk}\in\mathbb{R}$. It follows that

$$\begin{aligned}&\{\mathcal{F},\rho\}_d\\ &= \begin{pmatrix}\{F_1,\rho\}\\ \{F_2,\rho\}\\ \{F_3,\rho\}\\ \{F_4,\rho\}\end{pmatrix}\\ &= \begin{pmatrix} \left[\sqrt{2}(\alpha_2 f_{13}+\alpha_3 f_{14}) + (\frac{1}{2}-\alpha_1)f_{12}\right]\mathrm{Id}_2 + (\alpha_1-\frac{1}{2})f_{12}\sigma_1\\ \left[\sqrt{2}(\alpha_2 f_{23}+\alpha_3 f_{24}) - (\frac{1}{2}+\alpha_1)f_{12}\right]\mathrm{Id}_2 - (\frac{1}{2}+\alpha_1)f_{12}\sigma_1\\ \left[\sqrt{2}\alpha_3 f_{34} - (\frac{1}{2}+\alpha_1)f_{13} + (\alpha_1-\frac{1}{2})f_{23}\right]\mathrm{Id}_2 + [\frac{f_{23}-f_{13}}{2} - \alpha_1(f_{13}+f_{23})]\sigma_1\\ -\left[(\frac{1}{2}+\alpha_1)f_{14} + \sqrt{2}\alpha_2 f_{34} + (\frac{1}{2}-\alpha_1)f_{24}\right]\mathrm{Id}_2 - [\frac{f_{14}-f_{24}}{2} + (f_{14}+f_{24})\alpha_1]\sigma_1\end{pmatrix}\\ &\quad + \begin{pmatrix} (\frac{f_{13}}{\sqrt{2}} + \alpha_2 f_{12})\sigma_2 + (\frac{f_{14}}{\sqrt{2}} + \alpha_3 f_{12})\sigma_3\\ (\frac{f_{23}}{\sqrt{2}} - \alpha_2 f_{12})\sigma_2 + (\frac{f_{24}}{\sqrt{2}} - \alpha_3 f_{12})\sigma_3\\ -\alpha_2(f_{13}+f_{23})\sigma_2 + \frac{f_{34}}{\sqrt{2}} - \alpha_3(f_{13}+f_{23})\sigma_3\\ -[(f_{14}+f_{24})\alpha_2 + \frac{f_{34}}{\sqrt{2}}]\sigma_2 - (f_{14}+f_{24})\alpha_3\sigma_3\end{pmatrix}.\end{aligned}$$

Since $\nabla_d\rho$ has coefficient 0 along Id_2 in all components, for $\mathcal{D}\nabla_d\rho + \{\mathcal{F},\rho\}_d = 0$ to hold, we would need the coefficients of Id_2 to vanish in all components of $\{\mathcal{F},\rho\}_d$. But since $\mathcal{D}$ is diagonal, and all components of $\nabla_d\rho$ lack σ_j for one

$j \in \{1, 2, 3\}$, we would also need those coefficients to vanish, i.e.

$$f_{12} = 0,\ f_{13} = -f_{23},\ f_{14} = -f_{24}.$$

Numerically solving the remaining equations then leads to $\mathcal{F} = 0$, $\rho = \frac{1}{2}\,\mathrm{Id}$, which is exactly what is claimed in Conjecture 2.18.
We further compute

$$\begin{aligned}
\mathrm{div}_d(\{\mathcal{F}, \rho\}_d) = &\left[2f_{34} - \sqrt{2}\alpha_3(f_{13} + f_{23}) + \sqrt{2}\alpha_2(f_{14} + f_{24})\right]\sigma_1 \\
&+ \left[2\frac{f_{24} - f_{14}}{\sqrt{2}} - \sqrt{2}\alpha_1(f_{14} + f_{24}) - 2\alpha_3 f_{12}\right]\sigma_2 \\
&+ \left[2\frac{f_{13} - f_{23}}{\sqrt{2}} + \sqrt{2}\alpha_1(f_{13} + f_{23}) + 2\alpha_2 f_{12}\right]\sigma_3.
\end{aligned}$$

Introducing the notations

$$\begin{aligned}
r_1 := -\sqrt{2}(f_{13} + f_{23}),\ r_2 := \sqrt{2}(f_{14} + f_{24}),\ r_3 := -2f_{12}, \\
\beta_1 := 2f_{34}, \beta_2 := 2\frac{f_{24} - f_{14}}{\sqrt{2}}, \beta_3 := 2\frac{f_{13} - f_{23}}{\sqrt{2}},
\end{aligned}$$

this simplifies to

$$\begin{aligned}
\mathrm{div}_d(\{\mathcal{F}, \rho\}_d) = &(r_1\alpha_3 + r_2\alpha_2 + \beta_1)\sigma_1 + (-r_2\alpha_1 + r_3\alpha_3 + \beta_2)\sigma_2 \\
&+ (-r_1\alpha_1 - r_2\alpha_2 + \beta_3)\sigma_3.
\end{aligned} \tag{2.21}$$

From comparing the coefficients in (2.20), (2.21) with those in

$$\tfrac{\mathrm{d}}{\mathrm{dt}}\rho = \dot{\alpha}_1\sigma_1 + \dot{\alpha}_2\sigma + \dot{\alpha}_3\sigma_3,$$

we obtain

$$\dot{\alpha}(t) = (M + R)\alpha(t) + \beta,$$

where

$$\alpha(t) = \begin{pmatrix} \alpha_1(t) \\ \alpha_2(t) \\ \alpha_3(t) \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$
$$M = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix}, \quad R = \begin{pmatrix} 0 & r_2 & r_1 \\ -r_2 & 0 & r_3 \\ -r_1 & -r_3 & 0 \end{pmatrix}.$$

It is $\beta = 2\,\mathrm{div}_d(\mathcal{F})$, as can be seen from (2.21) with $\alpha = 0$. This completes the proof. $\square$

For sake of completeness, we remark that letting $H \neq 0$ only modifies the matrix R from Proposition 2.16 (see also Corollary 2.8). From Proposition 2.16, we can extract some properties of (2.15) for $d = 2$:

Corollary 2.17. *Let ρ solve (2.15) with initial condition ρ_0 (ρ_0 a density matrix). Let α, M, R and β be given as in Propostion 2.16. Then it holds that*

(i) There is a unique normalised stationary state $\rho_\infty \neq c\,\mathrm{Id}$ iff $\mathrm{div}(\mathcal{F}) \neq 0$ and $M + R$ is invertible.

(ii) All solutions converge to the set of invariant states with rate at least μ, where μ^2 is the smallest eigenvalue of $(M + R)^\dagger(M + R)$.

Proof: (i): From Proposition 2.16, we have that there is a unique normalised stationary state iff $M + R$ is invertible. If $\beta \neq 0$, this stationary state is not $c\,\mathrm{Id}$. From Proposition 2.16, we have that $\beta \neq 0$ iff $\mathrm{div}_d(\mathcal{F}) \neq 0$. Since $d_{jj} = 0$ implies $f_{jk} = 0$ for all k, any R that would lead to invertibility of $M + R$ would have to come from the Hamiltonian H.
(ii): Let $\gamma_j := (M + R)^{-1}\beta_j$. Then $\rho_\infty = \frac{1}{2}\,\mathrm{Id}_2 + \sum j = 1^3\gamma_j\sigma_j$, and due to the orthogonality of Id_2, σ_j $(j = 1, 2, 3)$ in $\mathcal{H}$ we obtain that

$$\|\rho - \rho_\infty\|^2_{\mathcal{H}} = 4\sum_{j=1}^{3}(\alpha_j - \gamma_j)^2 = 4|\alpha - \gamma|^2,$$

where we have used $\|\sigma_j\|_{\mathcal{H}} = 2$, $1 \leq j \leq 3$. We also get

$$\tfrac{\mathrm{d}}{\mathrm{dt}}(\alpha - \gamma) = \dot{\alpha} = (M + R)\alpha - \beta = (M + R)(\alpha - \gamma).$$

So for the solution ρ, we have

$$(\alpha(t) - \gamma) = e^{(M+R)t}(\alpha_0 - \gamma),$$

where α_0 is the coefficient vector for the initial condition ρ_0. This implies

$$|\alpha(t) - \gamma| = |e^{(M+R)t}(\alpha_0 - \gamma)| \leq \|e^{(M+R)t}\|_2\ |\alpha_0 - \gamma| \leq e^{-\mu t}|\alpha_0 - \gamma|.$$

Here, $\|e^{(M+R)t}\|_2$ is the spectral norm of $e^{(M+R)t}$ and we have used $\|e^A\| \leq e^{\|A\|}$ with

$$\mu = \|M + R\|_2 = \sqrt{\lambda},$$

where $\lambda > 0$ is the smallest eigenvalue of $(M+R)^\dagger(M+R)$. □

These results closely resemble the situation discussed in §1. If there is "full dissipation" ($\mathcal{D}$ has full rank), then the antisymmetric part is not necessary for existence of unique stationary states, though it can influence the stationary state itself as well as the rate of decay. In the case where $\mathcal{D}$ does not have full rank, the antisymmetric part of $\mathcal{F}$ has to be "compatible" for the equation to still have existence of a unique stationary state. This compatibility takes the form of condition (A) in §1, and of the condition "$M+R$ invertible" for this 2-dimensional case.

2.6 Conclusion, open questions

With the results of Section 2.4, we can not expect to establish an entropy method for symmetric operators $\mathcal{L}$, a scenario that would correspond to the "canonic" entropy method for symmetric Fokker-Planck equations. Since chapter 1 shows that the entropy method can be transferred to non-symmetric equations, this is not the end for a quantum entropy method; it merely implies that one would have to take non-symmetric terms into account.

A problem in doing so is the result of Lemma (2.7): discrete partial derivatives do not commute, at least not with the chosen definition. Even worse, we have

Corollary 2.18. *Let $\rho \in \mathbb{C}^{d\times d}$. Then all discrete partial derivatives*

$$\partial_d^j \rho := i[\rho, E_j]$$

commute iff $\rho = c\,\mathrm{Id}_d$ for some $c \in \mathbb{C}$.

Proof: From Lemma (2.7), we have that

$$\partial_d^l(\partial_d^j \rho) - \partial_d^j(\partial_d^l \rho) = i[\rho, \sum_{m=1}^{r} s_m E_{k_m}].$$

From the computations in the proof of Proposition 2.16, we have that $\sum_{m=1}^{r} s_m E_{k_m}$ spans at least E_3, E_4, and that a matrix that commutes with E_3 and E_4 is a multiple of the identity. So the assumption follows for $d = 2$, and by extension for $d > 2$. □

In the classical case, commutation of second partial derivatives is used for computing the Gronwall inequality between first and second time derivative of the relative entropy. Since this does not hold in the quantum case, we can not expect to be able to transfer this computation; this inequality would have to be found another way.

This result and Corollary 2.8 indicate that the approach used in chapter 1 for the entropy method for non-symmetric Fokker Planck equations can not easily be transferred: Neither is $\mathrm{div}_d(\mathcal{R}\nabla_d(\rho)) = 0$ for $\mathcal{R} = -\mathcal{R}^T$, nor can we expect to obtain our stationary state from the equation

$$(\mathcal{D} + \mathcal{R})\nabla_d \rho_\infty + \mathcal{F}\rho_\infty + \rho_\infty \mathcal{F} = 0.$$

So one would have to find another approach to handle the non-symmetric terms that appear in any operator $\mathcal{L}$ with a kernel not spanned by Id_d.

Another fundamental difference is apparent from Theorem 2.11 and Proposition 2.16: The rate of convergence is mainly influenced by the diffusion matrix D; the drift part of 2.15 seems to only contribute a rotation that mixes eigenvalues. In contrast, for the classical case discussed in chapter 1, the diffusion matrix D merely influences the shape of the stationary state; the rate of convergence is solely determined by the eigenvalues of the drift part. This, combined with the fact that there can be constant non-trivial stationary states, indicates that the finite dimensional problem might be more closely connected with Fokker-Planck equations on bounded domains than those on the whole space.

In conclusion, the presented approach allows a discussion of (2.1) in terms that more closely resemble the techniques used in partial differential equations than the statistical approach generally used for (2.1). While ultimately unsuccessful in its endeavour to establish an entropy method for open quantum systems, we are hopeful that it can serve as a basis for future research and to identify some of the similarities and differences between (2.1) and (1).

Bibliography

[1] R. Alicki, *Invitation to quantum dynamical semigroups*, Lecture Notes in Physics **597** (2002), 239–264.

[2] R. Alicki, K. Lendi, *Quantum Dynamical Semigroups and Application*, Springer Verlag (2007).

[3] A. Arnold, E. Carlen, Q. Ju, *Large-time behavior of non-symmetric Fokker-Planck type equations*, Communications in Stochastic Analysis **2** (2008) 153–175.

[4] A. Arnold, J.A. Carrillo, C. Manzini, *Refined long-time asymptotics for some polymeric fluid flow models*, Comm. Math. Sc. **8** (2010), 763–782.

[5] A. Arnold, C. Lasser, D. Matthes, *Wigner-Fokker-Planck equations with perturbation*, preprint, 2014.

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

[7] S. Attal, A. Joye, C.-A. Pillet, *Open Quantum Systems I. The Hamiltonian Approach.*, Springer Verlag (2006).

[8] S. Attal, A. Joye, C.-A. Pillet, *Open Quantum Systems II. The Markovian Approach.*, Springer Verlag (2006).

[9] S. Attal, A. Joye, C.-A. Pillet, *Open Quantum Systems III. Recent Developments.*, Springer Verlag (2006).

[10] F. Baudoin, *Bakry-Emery meet Villani*, http://arxiv.org/abs/1308.4938.

[11] B. Baumgartner, H. Narnhofer, W. Thiring, *Analysis of quantum semigroups with GKS-Lindblad generators: I. Simple Generators*, J. Phys. A.: Math. Gen., **41** 2008, 065201.

[12] B. Baumgartner, H. Narnhofer, W. Thiring, *Analysis of quantum semigroups with GKS-Lindblad generators: II. General*, J. Phys. A.: Math. Gen., **41** 2008, 395303.

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

[14] D. Bakry, M. Émery, *Hypercontractivité de semi-groupes de diffusion*, C. R. Acad. Sci. Paris Sér. I Math. **299** (1984), 775–778.

[15] D. Bakry, M. Émery, *Inégalités de Sobolev pour un semi-group symétrique*, C. R. Acad. Sci. Paris Sér. I Math. **301** (1985), 411–413.

[16] P. Biler, J. Dolbeault, *Long time behavior of solutions to Nernst-drift-diffusion systems*, Annales Henri Poincaré **1** (2000), 461–472.

[17] P. Biler, J. Dolbeault, P.A. Markowich, *Large time asymptotics of nonlinear drift-diffusion systems with Poisson coupling*, Transport Theory and Statistical Physics **30** (2001), 521–536.

[18] H.-P. Breuer, F. Petruccione, *The Theory of Open Quantum Systems*, Oxford University Press (2002).

[19] R. Carbone, *Logarithmic Sobolev inequalities and exponential entropy decay in non-commutative algebras*, http://arxiv.org/abs/1402.6948.

[20] R. Carbone, *Optimal log-sobolev inequality and hypercontractivity for positive semigroups on* $M_2(\mathbb{C})$, Infin. Dimens. Anal. Quantum Prob. Rel. **07** (2004), 317.

[21] R. Carbone, E. Sasso, V. Umanità, *Decoherence for positive semigroups on* $M_2(\mathbb{C})$, J. Math. Phys. **52** (2011), 032202.

[22] R. Carbone, E. Sasso, V. Umanità, *Decoherence for quantum markov semigroups on matrix algebras*, Annales Henri Poincaré **14** (2013), 681–697.

[23] J. A. Carrillo, A. Jüngel, P.A. Markowich, G. Toscani, A. Unterreiter, *Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities*, Monatsh. Math. **133** (1999), 1–82.

[24] J.A. Carrillo, P.A. Markowich, A. Unterreiter, *Large-time asymptotics of porous-medium type equations*, Gakuto International Series Mathematical Sciences and Applications **13** (2000), 24–36.

[25] J.A. Carrillo, G. Toscani, *Asymptotic* L^1*-decay of solutions of the porous medium equation to self-similarity*, Indiana University Mathematics Journal **49** (2000), 113–141.

[26] I. Csiszár, *Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis von Markoffschen Ketten*, Magyar Tud. Akad. Mat. Kutató Probab. **9** (1981), 533–535.

[27] L. Desvillettes, *Hypocoercivity: The example of linear transport*, Contemporary Mathematics **409**: Recent Trends in Partial Differential Equations (2006), 33–53.

[28] J. Dolbeault, M. del Pino, *Generalized Sobolev inequalities and asymptotic behaviour in fast diffusion and porous medium problems*, Preprint Ceremade (1999).

[29] J. Dolbeault, C. Mouhot, C. Schmeiser, *Hypocoercivity for kinetic equations with linear relaxation terms*, Comptes Rendu Mathematics **347** (2009), 511–516.

[30] J. Dolbeault, C. Mouhot, C. Schmeiser, *Hypocoercivity for linear kinetic equations conserving mass*, to appear in Trans. AMS.

[31] R. Duan, *Hypocoercivity of Linear Degenerately Dissipative Kinetic Equations*, Nonlinearity **24** (2011), 2165–2189.

[32] K.-J. Engel, R.Nagel, *A Short Course on Operator Semigroups*, Springer 2006.

[33] J. Erb, A. Arnold, *Sharp entropy decay for hypocoercive Fokker-Planck equations with linear drift*, Preprint 2014.

[34] S. Gadat, L. Miclo, *Spectral decompositions and L^2-operator norms of toy hypocoercive semi-groups*, Kinetic and Related Models **2** (2013), 317–372.

[35] V. Gorini, A. Kossakowsi, E.C.G. Sudar shan, *Completely positive dynamical semigroups of N-level systems*, J. Math. Phys. **17** (1976), 821–825.

[36] L. Gross, *Logarithmic Sobolev inequalities*, American Journal of Mathematics **97** (1975), 1061–1063.

[37] L. Gross, *Logarithmic Sobolev inequalities and Contractivity Properties of Semigroups*, Lecture Notes in Mathematics **1563**, E. Fabes et al. (Eds.) "Dirichlet Forms", Springer (1993).

[38] P. Gurka, *A_r-condition for two weight functions and compact imbeddings of weighted Sobolev spaces*, Czechoslovak Mathematical Journal **38** (1988), 611–617.

[39] B. Helffer, F. Nier, *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*, Lecture Notes in Mathematics 1862 (2005).

[40] F. Hérau, *Short and long time behavior of the Fokker-Planck equation in a confining potential and applications*, Journal of Functional Analysis **244** (2007), 95–118.

[41] C.D. Hill, *A sharp maximum principle for degenerate elliptic-parabolic equations*, Indiana Univ. Math. J. **20** (1970), 213–229.

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

[43] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press (1990, 2012).

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

[45] J.J. Kohn, *pseudo-differential operators and hypoellipicity*, Proc. Symp. Pure Math. **23** (1973), 61–69.

[46] S. Kullback, *Information Theory and Statistics*, John Wiley (1959).

[47] G. Lindblad, *Completely positive maps and entropy inequalities*, Communications in Mathematical Physics **40** (1975), 147–151.

[48] G. Lindblad, *Expectations and entropy inequalities for finite quantum systems*, Communications in Mathematical Physics **39** (1974), 111–119.

[49] G. Lindblad, *On the Generators of Quantum Dynamical Semigroups*, Communications in Mathematical Physics **48** (1976), 119–130.

[50] L. Lorenzi, M. Bertoldi, *Analytical methods for Markov Semigroups*, Pure and Applied Mathematics **283**, Chapman & Hall/CRC (2007).

[51] P. Montmarché, *Hypocoercive relaxation to equilibrium for some kinetic models*, Kinetic and Related Models **7** (2014), 341–360.

[52] C. Mouhot, L. Neumann, *Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus*, Nonlinearity **19** (2006), 969–998.

[53] I. Kolár, P.W. Michor, J. Slovák, *Natural Operations in Differential Geometry*, Springer Verlag (1999).

[54] R. Olkiewicz, B. Zegarlinski, *Hypercontractivity in noncommutative L^p Spaces*, J. Funct. Anal. **161** (1999), 246–285.

[55] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Reprint No. 8 (1999), Max-Planck-Institut Leipzig.

[56] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer (1983).

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

[58] S.F. Huelga, A. Rivas, *Open Quantum Systems. An Introduction.*, Springer Verlag (2012).

[59] L.P. Rothschild, E.M. Stein, *hypoelliptic differential operators and nilpotent groups*, Acta Math. **137** (1976), 247–320.

[60] L. Saloff-Coste, *Lectures on finite Markov chains*, Lect. Notes Math. **1665** (1997).

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

[62] C. Sparber, J.A. Carrillo, J. Dolbeault, P. Markowich, *On the long-term behaviour of the Quantum Fokker-Planck equation*, Monatshefte für Mathematik **141** (2004), 237–257.

[63] H. Spohn, *Approach to equilibrium for completely positive dynamical semigroups of n-level systems*, Reports on Mathematical Physics **10** (1976), 189–194.

[64] H. Spohn, *Entropy production for quantum dynamical semigroups*, J. Math. Phys. **5** (1978), 1227–1230.

[65] D. Tataru, *On The Fefferman-Phong Inequality And Related Problems*, Communications in Partial Differential Equations **26** (2002), 2101–2138.

[66] F. Verstraete, M.M. Wolf, J.I. Cirac, *Quantum computation and quantum-state engineering driven by dissipation*, Nature Physics **5** (2009), 633–636.

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

[68] D. Wittman, *Fisher Matrix for Beginners*, http://www.physics.ucdavis.edu/~dwittman/Fisher-matrix-guide.pdf

[69] W.H. Zurek, *Decoherence, einselection, and the quantum origins of the classical.*, Rev. Mod. Phys. **75** (2003), 715–775.