Finding the jump rate for fastest decay in the Goldstein-Taylor model

Helge Dietert, Josephine Evans

Finding the jump rate for fastest decay in the Goldstein-Taylor model

Helge Dietert* Josephine Evans†

April 8, 2021

Abstract

For hypocoercive linear kinetic equations we first formulate an optimisation problem on a spatially dependent jump rate in order to find the fastest decay rate of perturbations. In the Goldstein-Taylor model we show (i) that for a locally optimal jump rate the spectral gap is determined by multiple, possible degenerate, eigenvectors and (ii) that globally the fastest decay is obtained with a spatially homogeneous jump rate. Our proofs rely on a connection to damped wave equations and a relationship to the spectral theory of Schrödinger operators.

Keywords: Hypocoercivity; spatial weight; optimal control; Goldstein-Taylor model; wave equation

1 Introduction

A typical linear kinetic equation takes the form

$$\partial_t f + T f = \sigma(x)\ C(f) \tag{1}$$

for a density $f = f(t, x, v)$ at time t over the phase space consisting of a spatial position x and a velocity v where T is a transport operator, σ is a spatial weight, and C is a collision operator driving the system to thermal equilibrium.

The theory of hypocoercivity, [22, 11], ensures, by a variety of proofs, that the equilibrium is reached with an exponential rate. The spectral gap λ limiting the decay behaves for a constant σ typically as indicated in Fig. 1. Here we see two distinct regimes:

1. For small jump rates σ the spectral gap scales with σ. Here the decay is limited by the thermalisation rate of the velocity variable so that a faster jump rate improves the spectral gap.
2. For bigger jump rates σ the spectral gap behaves like σ^{-1}. In this regime the decay rate is limited by the spatial diffusion. Here a faster decay rate means slower decay as the effective spatial transport decreases by the law of large numbers.

This motivates the main question of this research.

Question 1. *Can we combine spatial regions of large and small jump rates in order to obtain a faster decay rate? More generally, what is the jump rate σ in order to find the largest spectral gap, i.e. the fastest decay?*

*Email: helge.dietert@imj-prg.fr
Université de Paris and Sorbonne Université, CNRS, Institut de Mathématiques de Jussieu-Paris Rive Gauche (IMJ-PRG), F-75013, Paris, France
Currently on leave and working at
Institut für Mathematik, Universität Leipzig, D-04103 Leipzig, Germany
†Email: josephine.evans@warwick.ac.uk
Warwick Mathematics Institute, University of Warwick, UK

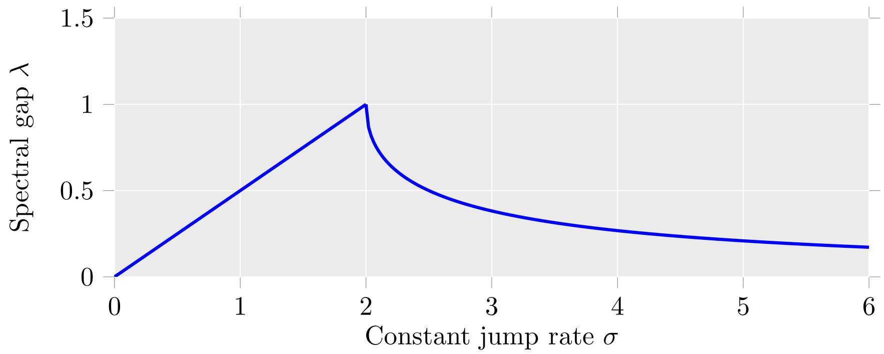

Figure 1: Typical decay rate depending on the noise strength. For a given constant σ, we plot the spectral gap for typical models of (1).

There have been several research works which fix σ and find bounds on the rate of convergence to equilibrium for the system; these works fit into the general framework of *hypocoercivity*. The goal of this work is to understand the dependence of the rate of convergence to equilibrium on σ by studying the optimal control problem of finding the choice of σ which maximises the rate. We believe this provides another direction to understand the precise dependence of the decay rate on the parameters in the equation. We believe this approach has potential in a variety of other kinetic equations:

- For kinetic equations with a confining potential one could investgate the interplay between jump rate and the confining potential.
- For equations posed on a domain with boundary one could investigate, in a similar way, the dependence of the rate on the shape of the domain and boundary conditions. This could produce results similar to the celebrated Faber-Krahn inequality.
- This problem is related to the control of nuclear reactors as for the radiative transfer equation σ is related to the presence or absence of control rods.
- In our perturbation result we show that the optimal σ must occur simultaneously with a degeneracy in the eigenspace associated to the spectral gap eigenvectors. We believe this might point to connections between the optimal σ and symmetries present in the equation.

Apart from applications in kinetic theory, Markov Chain Monte Carlo (MCMC) algorithms are a main motivation. In applications of Bayesian statistics, one needs to calculate the posterior distribution which is given up to a normalisation factor by

$$\mathrm{e}^{-\phi(x)}.$$

For a high-dimensional problems an explicit computation is prohibitively expensive and a common solution is to construct a stochastic process Z which converges to the sought distribution and to sample from that process. One such a process is a diffusion process

$$\mathrm{d}X_t = -\nabla_x \phi(X_t)\,\mathrm{d}t + \mathrm{d}W_t.$$

This procedure can sometimes be slow, and Hamiltonian MCMC (HMCMC) has been developed as a way to increase the speed of convergence of these algorithms, see [12, 7] for a rigorous proof of the increase in speed and references within on HMCMC. The strategy of Hamiltonian Markov Chain Monte Carlo is to look at the related kinetic equation

$$\begin{cases} \mathrm{d}X_t = V_t\,\mathrm{d}t, \\ \mathrm{d}V_t = -\nabla_x \phi(X_t)\,\mathrm{d}t + \sigma(X_t)\,(\mathrm{d}W_t - V_t\,\mathrm{d}t) \end{cases}$$

which has the equilibrium distribution $M(v)\,\mathrm{e}^{-\phi(x)}$ for the velocity equilibrium $M(v)$ so that the sought distribution is obtained by the spatial distribution. Here the intuitive idea is that the kinetic equation yields a faster transport of the distribution over large spatial distances. The previous analyses look at the case of constant σ and we now ask the further question whether the speed of convergence of these processes can be increase by making σ spatially dependent. This has been investigated numerically in statistics literature, for example in [15], where they propose a version of the Metropolis adjusted Langevin algorithm (MALA) which takes into account the geometry of ϕ.

A very simple model to study the exponential decay of kinetic equations is the one-dimensional Goldstein-Taylor model, which is still actively studied as a test case for hypocoercive results [4], and has been studied with σ depending on x in [6], relating it to the work Lebeau [19]. It is a special case of BGK models with only two velocities ± 1. Setting $u = (t,x) = f(t,x,+1)$ and $v(t,x) = f(t,x,-1)$ to the respective spatial densities, the model writes

$$\begin{cases} \partial_t u + \partial_x u = \dfrac{\sigma(x)}{2}\,(v-u), \\ \partial_t v - \partial_x v = \dfrac{\sigma(x)}{2}\,(u-v), \end{cases} \tag{2}$$

where we consider the spatial variable x in the torus $\mathbb{T}$ with length 2π.

As used before [18, 6], the one-dimensional case has the special feature that the kinetic equation (2) is equivalent to a damped wave equation by considering

$$\rho(t,x) := \frac{u(t,x)+v(t,x)}{\sqrt{2}} \qquad \text{and} \qquad j(t,x) := \frac{u(t,x)-v(t,x)}{\sqrt{2}}. \tag{3}$$

Then the Goldstein-Taylor model (2) can be written as

$$\begin{cases} \partial_t \rho + \partial_x j = 0, \\ \partial_t j + \partial_x \rho = -\sigma(x)\; j. \end{cases} \tag{4}$$

In our results we want to characterise the convergence towards the stationary state which is in the formulation (2) given by $u = v = \text{const}$ or in the formulation (4) by $\rho = \text{const}$ and $j = 0$. By the conservation of the mass $\int_{\mathbb{T}} \rho(x)\,\mathrm{d}x$, the limiting state can be characterised and using the linearity it therefore suffices to study the perturbation from the limiting space.

Working in L^2 we therefore consider the evolution in the space

$$\mathrm{L}^2_p = \{(\rho,j) \in \mathrm{L}^2(\mathbb{T}) : \int_{\mathbb{T}} \rho(x)\,\mathrm{d}x = 0\} \tag{5}$$

with the natural norm given by

$$\|(\rho,j)\|_2^2 = \|\rho\|_2^2 + \|j\|_2^2.$$

The evolution in L^2_p can be understood with a semigroup with the generator A_σ (see Proposition 4 below). As a first result we characterise a possible jump rate $\sigma = \sigma(x)$ by considering perturbations.

Theorem 2. *Suppose $\sigma \in \mathrm{L}^\infty(\mathbb{T})$ is such that the spectral gap in L^2_p is locally maximised. Then the spectral gap is not determined by a simple eigenvalue.*

Such a wave equation of a string has been studied in Cox and Zuazua [9] with fixed ends. One aspect of their work is to characterise the eigenvalues along the real axis by the spectrum of Schrödinger operators which allows them to find the spatial damping σ minimising the largest real eigenvalues, but they cannot consider the full spectral gap. However, in our situation, and with our aims in mind, we are able to go further by associating a different Schrödinger operator. In this newly associated Schrödinger operator we can obtain the result by looking at the second eigenvalue and exploiting the translation symmetry; this is key to our bound when $\|\sigma\|_1$ is large. Due to the different boundary values we capture the true spectral gap in contrast to Cox and Zuazua [9] where their result for the eigenvalues along the real axis does not capture the spectral gap. This yields the following theorem:

Theorem 3. *For the Goldstein-Taylor model* (2)*, the largest spectral gap in* L^2_p *is obtained with the constant* $\sigma = 2$ *giving the spectral gap* 1*.*

In the context of the wave equation, the corresponding question of the spatially dependent damping has been studied before and shows that the competing effects of the jump rate are more intricate as the result of a constant damping in Theorem 3 might suggests. So it is noted by Castro and Cox [8] that an arbitrary large decay rate can be obtained in the case of fixed-ends and a spatialy dependent damping diverging towards the boundary. Taking the damping as an indicator function of a set ω and optimising the set ω, the competition between the effects is non-trivial and yields in general non-existence of optimal sets [17, 20]. Phrased in terms of the related observability condition it has been further studied in [21]. The problem has also been formulated in terms of the overall energy [10] and from a numerical side the problem is also studied by e.g. [14]. It is also studied in more complex geometries in Lebeau [19].

The study of the decay rate for the presented class of systems is a wide field ranging from works in kinetic theory [1, 2, 4] to the wave equation [16] to numerical methods [3].

As a first step, we formulate in Section 2 the spectral problem precisely and also shows that the spectral gap determines the decay rate of the L^2 norm under the flow of our equation. Furthermore, we find eigenvalues corresponding to the decay rate $\|\sigma\|_1/(4\pi)$ of the velocity distribution alone.

Proposition 4. *For non-negative* $\sigma \in \mathrm{L}^\infty$ *the closed linear operator* A_σ *defined by*

$$A_\sigma \begin{pmatrix} \rho \\ j \end{pmatrix} = \begin{pmatrix} 0 & -\partial_x \\ -\partial_x & -\sigma(x) \end{pmatrix} \begin{pmatrix} \rho \\ j \end{pmatrix} \tag{6}$$

on L^2_p *generates a contraction semigroup* $(\mathrm{e}^{tA_\sigma})_t$ *matching the evolution* (4)*.*

For any $\epsilon > 0$ *there exists an eigenvalue* $\lambda \in \mathbb{C}$ *with* $|\Re\lambda - \|\sigma\|_1/(4\pi)| \le \epsilon$ *and there are at most finitely many eigenvalues in the halfspace* $\{\lambda \in \mathbb{C} : \Re\lambda \ge \|\sigma\|_1/(4\pi) + \epsilon\}$*.*

If for $a > -\|\sigma\|_1/(4\pi)$ *there exists no eigenvalue with* $\Re\lambda \ge a$*, then we have the growth bound*

$$\|\mathrm{e}^{tA_\sigma}\| \lesssim \mathrm{e}^{ta} \qquad \forall t \ge 0.$$

Note that the corresponding results for the wave equation with fixed ends have been shown in Cox and Zuazua [9].

The main implication of our study is that, for the Goldstein-Taylor model, the rate of convergence to equilibrium cannot be improved by making σ depend on x. We believe this suggests that the rate of convergence to equilibrium is unlikely to be increased, in kinetic models, by adding local oscillations to the jump rate σ. We note that this does not exclude the possiblility that spatially dependent jump rates cannot improve the rate of convergence to equilibrium in the presence of more complex geometries. In fact, it is proposed in [15] to vary σ on large scales in a way that is sympathetic to the confining function ϕ. Therefore we close the introduction by the following open question:

Question 5. *Does there exist a linear kinetic equation posed on* $\mathbb{R}^d$ *for some* d *with a confining potential for which a strictly faster rate of convergence can be achieved by allowing the collision rate* σ *to depend on* x *than is achieved by constant* σ*?*

2 Semigroup and spectral problem

In this section we prove Proposition 4 in two parts. First we show that it we have a well defined semigroup for the flow.

Proof of Proposition 4 (first part). The generator (6) formally gives the required PDE (4). Without the σ the solution is given explicitly by the characteristics defined by the transport and this explicit representation shows that it generates a semigroup. As $\sigma \in \mathrm{L}^\infty$, the contribution of σ in A_σ is a bounded perturbation so that it defines a semigroup.

The required mass conservation follows from the estimate that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{T}} \rho(t,x)\,\mathrm{d}x = -\int_{T} \partial_x j(t,x)\,\mathrm{d}x = 0$$

and the contraction property by the estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{T}} |\rho(t,x)|^2 + |j|^2\,\mathrm{d}x = -\int_{T} \left(\overline{\rho}\partial_x j + j(\partial_x\overline{\rho} + \sigma(x)\overline{j})\right)\mathrm{d}x = -\int_{T} \sigma(x)\,|j|^2 \mathrm{d}x \le 0.$$

This gives the first part of Proposition 4. □

In the spectral property the central object is the resolvent

$$R(\lambda, A_\sigma) := (\lambda - A_\sigma)^{-1}. \tag{7}$$

For $(a,b) \in \mathrm{L}^2_p$ the image $(\rho, j) = R(\lambda, A_\sigma)(a,b)$ is the solution to

$$(\lambda - A_\sigma)\begin{pmatrix}\rho \\ j\end{pmatrix} = \begin{pmatrix}a \\ b\end{pmatrix} \tag{8}$$

as long it has a unique solution. This equation can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}x}\begin{pmatrix}\rho \\ j\end{pmatrix} + \begin{pmatrix}0 & \sigma+\lambda \\ \lambda & 0\end{pmatrix}\begin{pmatrix}\rho \\ j\end{pmatrix} = \begin{pmatrix}b \\ a\end{pmatrix}. \tag{9}$$

Associated to (9) we define the solution operator for the homogeneous part and an operator $M(\sigma,\lambda)$ yielding the solvability condition.

Definition 6. *Fix $\sigma \in \mathrm{L}^\infty$ and $\lambda \in \mathbb{C}$. For $y \in \mathbb{T}$ and given $(\rho_{\mathrm{in}}, j_{\mathrm{in}})$ consider the linear ODE*

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}x}\begin{pmatrix}\rho(x) \\ j(x)\end{pmatrix} + \begin{pmatrix}0 & \sigma(x)+\lambda \\ \lambda & 0\end{pmatrix}\begin{pmatrix}\rho(x) \\ j(x)\end{pmatrix} = 0 \\ \begin{pmatrix}\rho(y) \\ j(y)\end{pmatrix} = \begin{pmatrix}\rho_{\mathrm{in}} \\ j_{\mathrm{in}}\end{pmatrix} \end{cases} \tag{10}$$

and define $\mathcal{S}^{y\to x}_{\sigma,\lambda}$ as solution operator so that $\mathcal{S}^{y\to x}_{\sigma,\lambda}((\rho_{\mathrm{in}}, j_{\mathrm{in}})) = (\rho(x), j(x))$. For the resolvent define the operator

$$M(\sigma,\lambda) = \mathrm{Id} - \mathcal{S}^{0\to 2\pi}_{\sigma,\lambda}. \tag{11}$$

By Duhamel's principle the resolvent equation (9) can be written with the solution operator as

$$\begin{pmatrix}\rho(y) \\ j(y)\end{pmatrix} = \mathcal{S}^{0\to y}_{\sigma,\lambda}\begin{pmatrix}\rho_0 \\ j_0\end{pmatrix} + \int_0^y \mathcal{S}^{x\to y}_{\sigma,\lambda}\begin{pmatrix}b(x) \\ a(x)\end{pmatrix}\mathrm{d}x \tag{12}$$

for constants ρ_0, j_0. By the periodic boundary condition $\rho(2\pi) = \rho_0$ and $j(2\pi) = j_0$ we must have

$$M(\sigma,\lambda)\begin{pmatrix}\rho_0 \\ j_0\end{pmatrix} = \int_0^{2\pi} \mathcal{S}^{x\to 2\pi}_{\sigma,\lambda}\begin{pmatrix}b(x) \\ a(x)\end{pmatrix}\mathrm{d}x \tag{13}$$

which yields a unique solution if $M(\sigma,\lambda)$ is invertible. If not the kernel gives an eigenvalue solving

$$\begin{cases} \lambda\rho + \partial_x j = 0, \\ \partial_x\rho + (\lambda + \sigma(x))j = 0. \end{cases} \tag{14}$$

The characterisation by the operator M corresponds to the well-known shooting method for solving eigenvalue problems as done in [9]. For the further analysis we note basic properties.

Lemma 7. *For a fixed $\sigma \in \mathrm{L}^\infty$ and $x, y \in \mathbb{T}$, the solution operator $\mathcal{S}_{\sigma,\lambda}^{x\to y}$ and $M(\sigma,\lambda)$ are analytic with respect to $\lambda \in \mathbb{C}$.*

Moreover, the solution operator is bounded as

$$\|\mathcal{S}_{\sigma,\lambda}^{x\to y}\| \leq \exp(|\Re(\lambda)|\,|y-x| + \|\sigma\|_1/2)$$

and the solution operator $\mathcal{S}_{\sigma,\lambda}^{x\to y}$ and $M(\sigma,\lambda)$ are continuous differentiable with respect to $\|\sigma\|_1$.

Proof. The existence and analyticity of the solution operator $\mathcal{S}_{\sigma,\lambda}^{x\to y}$ and $M(\sigma,\lambda)$ follow from standard ODE theory.

For the growth bound and the dependence with respect to $\|\sigma\|_1$ note the following a priori estimate

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{2}\big(|\rho|^2 + |j|^2\big) &= \overline{\rho}\partial_x\rho + j\partial_x\overline{j} \\
&= -\overline{\rho}(\sigma+\lambda)j - j\overline{\lambda}\overline{\rho} \\
&= -\overline{\rho}j(\sigma+\lambda+\overline{\lambda}) \\
&\leq \Big[2|\Re(\lambda)| + \sigma(x)\Big]\frac{1}{2}\big(|\rho|^2+|j|^2\big),
\end{aligned}$$

which yields the result. □

Hence we can precisely describe the resolvent.

Lemma 8. *The resolvent set consist of $\lambda \in \mathbb{C}$ for which the matrix $M(\sigma,\lambda)$ is invertible and the resolvent is bounded in the operator norm as*

$$\|R(\lambda, A)\| \lesssim \|M(\sigma,\lambda)^{-1}\|\ \exp\Big(4\pi|\Re(\lambda)| + \|\sigma\|_1\Big).$$

Moreover, for every λ in the spectrum of A_σ there exists at least one eigenvector.

Proof. The result follows directly from the representation (12) once we determine the constants ρ_0, j_0 by (13) and use the bound from Lemma 7.

If λ is in the spectrum of A_σ, then there exists an element in the kernel of $M(\sigma,\lambda)$ which yields an eigenvector. □

We now look at the asymptotic form of M as $|\Im\lambda| \to \infty$ over a finite range of $\Re\lambda$.

Lemma 9. *Fix a bounded interval $I \in \mathbb{R}$ and $\sigma \in \mathrm{L}^\infty(\mathbb{T})$, then*

$$M(\sigma,\lambda) \to \begin{pmatrix} 1-\cosh\Big(\lambda\Big(2\pi + \frac{\|\sigma\|_1}{2\lambda}\Big)\Big) & \sinh\Big(\lambda\Big(2\pi+\frac{\|\sigma\|_1}{2\lambda}\Big)\Big) \\ \sinh\Big(\lambda\Big(2\pi+\frac{\|\sigma\|_1}{2\lambda}\Big)\Big) & 1-\cosh\Big(\lambda\Big(2\pi+\frac{\|\sigma\|_1}{2\lambda}\Big)\Big) \end{pmatrix}$$

uniformly over $\Re\lambda \in I$ as $|\Im\lambda| \to \infty$.

This matches [10, Theorem 5.1] where the result has been proven with explicit error bounds by a series solution. For being self-contained, we give another shorter proof.

Proof. The idea is like in the proof of the Riemann-Lebesgue lemma that we can approximate σ by a piecewise constant function $\tilde{\sigma}$, i.e. there exists $x_0 = 0 < x_1 < x_2 < \cdots < x_K = 2\pi$ and $\sigma_1, \sigma_2, \ldots, \sigma_K$ such that

$$\tilde{\sigma}(y) = \sigma_j \qquad \forall y \in [x_{j-1}, x_j).$$

On each constant part we find the explicit solution

$$\mathcal{S}_{\tilde{\sigma},\lambda}^{x_{j-1}\to x_j} = \begin{pmatrix} \cosh\Big((x_j - x_{j-1})\sqrt{\lambda(\lambda+\sigma_j)}\Big) & -\frac{\lambda+\sigma_j}{\lambda}\sinh\Big((x_j-x_{j-1})\sqrt{\lambda(\lambda+\sigma_j)}\Big) \\ -\frac{\lambda}{\lambda+\sigma_j}\sinh\Big((x_j-x_{j-1})\sqrt{\lambda(\lambda+\sigma_j)}\Big) & \cosh\Big((x_j-x_{j-1})\sqrt{\lambda(\lambda+\sigma_j)}\Big) \end{pmatrix}$$

and for $|\Im\lambda| \to \infty$ this is converging uniformly to

$$\tilde{S}_j := \begin{pmatrix} \cosh\left(\lambda\left((x_j - x_{j-1}) + \frac{(x_j - x_{j-1})\sigma_j}{2\lambda}\right)\right) & -\sinh\left(\lambda\left((x_j - x_{j-1}) + \frac{(x_j - x_{j-1})\sigma_j}{2\lambda}\right)\right) \\ -\sinh\left(\lambda\left((x_j - x_{j-1}) + \frac{(x_j - x_{j-1})\sigma_j}{2\lambda}\right)\right) & \cosh\left(\lambda\left((x_j - x_{j-1}) + \frac{(x_j - x_{j-1})\sigma_j}{2\lambda}\right)\right) \end{pmatrix}.$$

By the group property of the solution operator we find

$$\mathcal{S}_{\tilde{\sigma},\lambda}^{0\to 2\pi} = \mathcal{S}_{\tilde{\sigma},\lambda}^{x_{j-1}\to 2\pi} \circ \mathcal{S}_{\tilde{\sigma},\lambda}^{x_{j-2}\to x_{j-1}} \circ \cdots \circ \mathcal{S}_{\tilde{\sigma},\lambda}^{x_0\to x_1}$$

which therefore converges uniformly to

$$\tilde{S}_j \circ \tilde{S}_{j-1} \circ \cdots \circ \tilde{S}_1 = \begin{pmatrix} \cosh\left(\lambda\left(2\pi + \frac{\|\tilde{\sigma}\|_1}{2\lambda}\right)\right) & -\sinh\left(\lambda\left(2\pi + \frac{\|\tilde{\sigma}\|_1}{2\lambda}\right)\right) \\ -\sinh\left(\lambda\left(2\pi + \frac{\|\tilde{\sigma}\|_1}{2\lambda}\right)\right) & \cosh\left(\lambda\left(2\pi + \frac{\|\tilde{\sigma}\|_1}{2\lambda}\right)\right) \end{pmatrix}.$$

As we can approximate any function $\sigma \in \mathrm{L}^\infty$ arbitrary well by the piecewise function in L^1 the result follows from the stability of Lemma 7. □

We can now prove the remaining parts of Proposition 4.

Proof of Proposition 4 (remaining part). The asymptotic expression of M in Lemma 9 is invertible except when $\lambda = -\|\sigma\|_1/(4\pi) + 2\pi i n$ for $n \in \mathbb{Z}$. By considering the uniform convergence over $\Re\lambda \in [-\|\sigma\|_1/(4\pi) - \epsilon, -\|\sigma\|_1/(4\pi) + \epsilon]$ then yields the existence of root of limiting expression for $\det M$ in the strip for $|\Im\lambda|$ large enough. As $\det M$ is analytic, Rouge's theorem ensures then the existence of a root for $\det M$, i.e. an eigenvalue for the generator.

As A_σ generates a contracting semigroup so that there are no eigenvalues λ with $\Re\lambda > 0$. By the asymptotic expression there are no eigenvalues in $-\|\sigma\|_1/(4\pi) + \epsilon$ for large enough $|\Im\lambda|$. As M is analytic there can be only finitely many eigenvalues in the remaining bounded region.

For the last part assume $a > -\|\sigma\|_1/(4\pi)$ such that there is no eigenvalue λ with $\Re\lambda \geq a$. For $\Re\lambda > 1$ we can use the fact that $(\mathrm{e}^{tA_\sigma})_t$ is a contraction semigroup to find a uniform bound on the resolvent $R(\lambda, A_\sigma)$. By the asymptotic expression of Lemma 9 we have a uniform bound of $\|M(\sigma,\lambda)^{-1}\|$ for large enough $|\Im\lambda|$ and $\Re\lambda \in [a, 1]$. As there are no eigenvalue with real part equal to a and M is continuous, this shows

$$\sup_{\Re\lambda\in[-a,1]} \|M(\sigma, a + \mathrm{i}b)^{-1}\| < \infty.$$

By Lemma 8 this shows the same bound for the resolvent. As our function space L^2_p is a Hilbert space, we can thus apply Gearhart-Prüss-Greiner theorem [13, Thm 1.11 in Chapter V] to find the claimed growth bound. □

Remark 10. *An alternative for using the Gearhart-Prüss-Greiner theorem is an adaptation of the theory of positive semigroups as done in Bernard and Salvarani [5]. Cox and Zuazua [9] establish the growth bound by studying the eigenvector system in more detail.*

We close this section noting the eigenvalues for the case that σ is constant.

Lemma 11. *Suppose that σ is constant. Then the spectrum in L^2_p consists of*

$$-\frac{\|\sigma\|_1}{4\pi} \pm \sqrt{\frac{\|\sigma\|_1^2}{(4\pi)^2} - n^2} \qquad \textit{for } n = 1, 2, 3, \dots$$

and

$$-\frac{\|\sigma\|_1}{2\pi}.$$

Proof. In this setting the spatial Fourier modes decouple and the result follows directly by solving the eigenvalue problem for each mode, where we exclude the stationary state as it is done in L^2_p. Alternatively, we can use the explicit expression of M for a constant σ as in the proof of Lemma 9. □

For a detailed study of the decay behaviour in the case of constant σ we refer to Achleitner, Arnold, and Signorello [2].

3 Perturbation analysis

Given some $\sigma_0 \in \mathrm{L}^\infty$ and a perturbation $\eta \in \mathrm{L}^\infty$ we can compute how the eigenvalues of $A_{\sigma_0+\epsilon\eta}$ are changing for varying ϵ. The eigenvalues $(\rho_\epsilon, j_\epsilon)$ satisfy the equation

$$(\lambda_\epsilon - A_{\sigma_0+\epsilon\eta}) \begin{pmatrix} \rho_\epsilon \\ j_\epsilon \end{pmatrix} = 0$$

and formally taking the derivative with respect to ϵ yields

$$(\lambda'_\epsilon - A'_{\sigma_0+\epsilon\eta}) \begin{pmatrix} \rho_\epsilon \\ j_\epsilon \end{pmatrix} + (\lambda_\epsilon - A_{\sigma_0+\epsilon\eta}) \begin{pmatrix} \rho'_\epsilon \\ j'_\epsilon \end{pmatrix} = 0.$$

Testing with a suitable adjoint $(\rho^*_\epsilon, j^*_\epsilon)$ yields

$$\lambda'_\epsilon \left\langle \begin{pmatrix} \rho^*_\epsilon \\ j^*_\epsilon \end{pmatrix}, \begin{pmatrix} \rho_\epsilon \\ j_\epsilon \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} \rho^*_\epsilon \\ j^*_\epsilon \end{pmatrix}, A'_{\sigma+\epsilon\eta} \begin{pmatrix} \rho_\epsilon \\ j_\epsilon \end{pmatrix} \right\rangle.$$

Hence assuming that both inner products are non-zero we find that formally λ_ϵ is differentiable with a non-zero derivative and thus by varying ϵ we can change the eigenvalue and improve the spectral gap if the spectral gap is determined by a single eigenvalue.

The key point for this perturbation analysis is to understand when the weight in front of λ'_ϵ can be vanishing. By using the formulation with respect to the operator $M(\sigma, \lambda)$ from Definition 6 we can show that for a simple eigenvalue this prefactor never vanishes.

As a first step we sharpen the condition for the spectrum from Lemma 8.

Lemma 12. *Suppose λ_0 is seperated from the spectrum, then λ_0 is determined by a simple eigenvalue if and only if $\det(M, \sigma, \lambda)$ has a simple root.*

Proof. Recall the formula (12) and (13) for the resolvent map. The inverse of M can be written as product of $\det(M)^{-1}$ and the adjugate. The spectral projection Π_{λ_0} to λ_0 is given by the contour integral

$$\int_C R(\lambda, \sigma) \mathrm{d}\lambda,$$

where C is a simple curve separating λ_0 from the rest of the specturum. This integral can be computed using the Laurent series. This gives a one-dimensional image if and only if $\det(M)$ has a simple root. □

Remark 13. *The fact that we have only finitely many eigenvalues in any strip $a \leq \Re(\lambda) \leq b$ with $a > -\|\sigma\|_1/(4\pi)$ means that every eigenvalue is separated from the rest of the spectrum.*

This allows us to determine M around an eigenvector.

Lemma 14. *Suppose that $\lambda_0 \in \mathbb{C}$ is a simple eigenvalue. Then let (ρ_0, j_0) be in $\ker M(\sigma, \lambda_0)$ and normalised to $\|(\rho_0, j_0)\|_2 = 1$. In the basis*

$$V_1 = \begin{pmatrix} \rho_0 \\ j_0 \end{pmatrix} \qquad \text{and} \qquad V_2 = \begin{pmatrix} -\overline{j_0} \\ \overline{\rho_0} \end{pmatrix}$$

the operator $M(\sigma,\lambda)$ takes the form

$$\tilde{M}(\sigma,\lambda_0+\delta\lambda) := \begin{pmatrix} O(\delta\lambda) & b+O(\delta\lambda) \\ c\,\delta\lambda + O((\delta\lambda)^2) & O(\delta\lambda) \end{pmatrix}$$

for small $\delta\lambda$ and constants $b,c\neq 0$.

Proof. By construction V_1 and V_2 form an orthonormal basis. Using the solution operator $\mathcal{S}^{\cdot\to\cdot}_{\sigma,\lambda_0}$ from Definition 6 we define

$$\begin{pmatrix} \rho_1(x) \\ j_1(x) \end{pmatrix} = \mathcal{S}^{0\to x}_{\sigma,\lambda_0} \begin{pmatrix} \rho_0 \\ j_0 \end{pmatrix} \tag{15}$$

and

$$\begin{pmatrix} \rho_2(x) \\ j_2(x) \end{pmatrix} = \mathcal{S}^{0\to x}_{\sigma,\lambda_0} \begin{pmatrix} -\overline{j_0} \\ \overline{\rho_0} \end{pmatrix}. \tag{16}$$

The idea of the Wronskian is to consider $\rho_1(x)j_2(x)-\rho_2(x)j_1(x)$ and by (10) we find

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(\rho_1(x)j_2(x)-\rho_2(x)j_1(x)\Big)=0.$$

As $(\rho_0,j_0)\in\ker M(\sigma,\lambda_0)$ we have that $(\rho_1(2\pi),j_1(2\pi))=(\rho_0,j_0)$. For V_2 we find with the Wronskian

$$\begin{aligned}\left\langle \begin{pmatrix} \rho_2(2\pi) \\ j_2(2\pi) \end{pmatrix}, V_2 \right\rangle &= \rho_0\, j_2(2\pi) - j_0\,\rho_2(2\pi) = \rho_1(2\pi)\,j_2(2\pi) - j_1(2\pi)\,\rho_2(2\pi) \\ &= \rho_1(0)j_2(0)-\rho_2(0)j_1(0) = |\rho_0|^2+|j_0|^2 = 1.\end{aligned}$$

Hence we have found that $\langle M(\sigma,\lambda_0)V_2,V_2\rangle = 0$. Further recalling that $V_1\in\ker M(\sigma,\lambda_0)$ and that M and $\tilde{M}$ are analytic with respect to λ shows that

$$\tilde{M}(\sigma,\lambda_0+\delta\lambda) = \begin{pmatrix} O(\delta\lambda) & b+O(\delta\lambda) \\ c\,\delta\lambda + O((\delta\lambda)^2) & O(\delta\lambda) \end{pmatrix}$$

for some constants $b,c\in\mathbb{C}$. By Lemma 12, the determinant $M(\sigma,\lambda)=\tilde{M}(\sigma,\lambda)$ must have a single root at λ_0 so that $b,c\neq 0$. ∎

We can now prove the perturbation result.

Proof of Theorem 2. We argue by contradiction and assume that for a $\sigma\in\mathrm{L}^\infty$ the spectral gap is determined by the simple eigenvalue λ_0.

For a perturbation $\eta\in\mathrm{L}^\infty$, we then find for $M(\sigma,\lambda)$ around λ_0 in the form of $\tilde{M}(\sigma,\lambda)$ defined in Lemma 14 that

$$\tilde{M}(\sigma+\epsilon\eta,\lambda_0+\delta\lambda) = \begin{pmatrix} O(\epsilon,\delta\lambda) & b+O(\epsilon,\delta\lambda) \\ c\,\delta\lambda + d\epsilon + O((\epsilon,\delta\lambda)^2) & O(\epsilon,\delta\lambda) \end{pmatrix}.$$

By Duhamel's principle we have that

$$\mathcal{S}^{0\to x}_{\sigma+\epsilon\eta,\lambda_0} = \mathcal{S}^{0\to x}_{\sigma,\lambda_0} + \int_0^x \mathcal{S}^{y\to x}_{\sigma,\lambda_0}\,\epsilon\eta(y)\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}\mathcal{S}^{0\to y}_{\sigma+\epsilon\eta,\lambda_0}\,\mathrm{d}y.$$

And the bottom left hand term of the matrix $\mathcal{S}^{0\to 2\pi}_{\sigma+\epsilon\eta,\lambda_0}$ will be $\langle \mathcal{S}^{0\to 2\pi}_{\sigma+\epsilon\eta,\lambda_0}V_1,V_2\rangle$. Repeating the Wronskian argument as in the proof of Lemma 14 shows that

$$d = \int_0^{2\pi}\eta(y)\langle \mathcal{S}^{x\to 2\pi}_{\sigma,\lambda_0}\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}\mathcal{S}^{0\to x}_{\sigma,\lambda_0},V_2\rangle\,\mathrm{d}y = \int_0^{2\pi}\begin{pmatrix} -j_0 \\ \rho_0 \end{pmatrix}\cdot\mathcal{S}^{y\to 2\pi}_{\sigma,\lambda_0}\begin{pmatrix} -j_1(y) \\ 0 \end{pmatrix}\mathrm{d}y = \int_0^{2\pi} j_1(y)^2\,\eta(y)\,\mathrm{d}y$$

with $j_1(y)$ from (15)[1]. As the eigenvalue is determined by $\det \tilde{M}$ this shows that the eigenvalue behaves as

$$\lambda_0 - \frac{d}{c}\epsilon + O(\epsilon^2).$$

Hence if we can find some η such that $\Re(d/c) \neq 0$, we can choose a small $\epsilon \in \mathbb{R}$ so that $\sigma + \epsilon\eta$ would have a bigger spectral gap. This shows the result if $\sigma + \epsilon\eta$ is a valid jump rate, i.e. non-negative.

As σ is a non-trivial jump rate, it is strictly positive in a subset I of positive measure. If $j_1(y)^2$ has not constant complex phase, we can always construct $\eta \in \mathrm{L}^\infty$ with support in I such that $\Re(d/c) \neq 0$ and thus $\sigma + \epsilon\eta$ yields a valid perturbation of the jump rate. In the case that $\lambda \in \mathbb{R}$, the real and imaginary part decouple and the constants d and c must be real. In the other case the system is translation invariant and thus the spectral gap cannot determined by a simple eigenvalue. □

4 Global optimum by associated Schrödinger equation

The eigenvector equation (14) can be written as the following second order equation

$$-\partial_x^2 j + \lambda(\lambda + \sigma) j = 0. \tag{17}$$

For a fixed $\lambda \in \mathbb{R}$ we then consider the Hamiltonian $H_{\sigma,\lambda}$ by

$$H_{\sigma,\lambda} j = -\partial_x^2 j + \lambda(\lambda + \sigma(x))\, j. \tag{18}$$

By the construction, our generator A_σ has an eigenvector with eigenvalue λ if $H_{\sigma,\lambda}$ has a zero eigenvector. By looking at the evolution of the spectrum, similar to Cox and Zuazua [9], we can show a slowly decaying eigenvector in the diffusive regime.

Proposition 15. *Suppose that $\|\sigma\|_1 > 4\pi$ and let*

$$\lambda_s = -\frac{\|\sigma\|_1}{4\pi} + \sqrt{\left(\frac{\|\sigma\|_1}{4\pi}\right)^2 - 1}.$$

Then there exists a $\lambda \in [\lambda_s, 0)$ such that λ is an eigenvalue of the generator A_σ from (6) of the Goldstein-Taylor system (4).

Proof. The Hamiltonian $H_{\sigma,\lambda}$ from (18) is for $\lambda \in \mathbb{R}$ a self-adjoint operator and has real eigenvalues $\mu_1, \mu_2, \dots$ (chosen in increasing order) converging to ∞.

For $\lambda = 0$, the Hamiltonian $H_{\sigma,\lambda}$ is just the Laplacian so that $\mu_1 = 0$ and $\mu_2 = 1$. By considering the perturbation around $\lambda = 0$, we find for a small $|\lambda_s| > \epsilon > 0$ that for $\lambda = 0 - \epsilon$ that $\mu_1 < 0$ and $\mu_2 > 0$.

For the given σ find a shift ϕ such that

$$\int_0^{2\pi} \sigma(x) \cos(2(x - \phi))\, \mathrm{d}x = 0$$

and consider the two test functions

$$j_1(x) = 1$$

and

$$j_2(x) = \sin(x - \phi)$$

which are linearly independent.

For these test functions we find

$$\langle j_1, H_{\sigma,\lambda_s} j_1 \rangle = \lambda_s (2\pi\lambda_s + 2\|\sigma\|_1) < 0$$

[1] One could also prove $c = \int_0^{2\pi} (j_1(y)^2 - \rho_1(y)^2)\, \mathrm{d}y$.

and

$$\langle j_2, H_{\sigma,\lambda_s} j_2 \rangle = \frac{2\pi}{2}\left(1 + \lambda_s\left(\lambda_s + \frac{\|\sigma\|_1}{2\pi}\right)\right) \le 0.$$

Hence we must have $\mu_2(\lambda_s) \le 0$. As $\mu_2(0-\epsilon) > 0$ this means that there exist some $\lambda \in [\lambda_s, -\epsilon)$ such that $\mu_2 = 0$ and by the previous discussion this means that there exist an eigenvalue for the generator A_σ. □

We can now conclude that the constant $\sigma = 2$ yields the fastest decay rate.

Proof of Theorem 3. By Lemma 11 the choice $\sigma = 2$ yields the spectral gap 1.

For another σ with $\|\sigma\|_1 \le 4\pi$, the bound from the velocity relaxation in Proposition 4 shows that the spectral gap is not bigger.

For another σ with $\|\sigma\|_1 > 4\pi$, we can apply Proposition 15 to show that the spectral gap needs to be worse than the constant case $\sigma = 2$. □

Acknowledgements

The authors gratefully acknowledge the support of the Hausdorff Research Institute for Mathematics (Bonn), through the Junior Trimester Program on Kinetic Theory. The authors thank Clément Mouhot, Laurent Desvillettes, Iván Moyano for the discussions on this project.

References

[1] Franz Achleitner, Anton Arnold, and Eric A. Carlen. "On Linear Hypocoercive BGK Models". In: *Springer Proceedings in Mathematics & Statistics* (2016), 1–37. ISSN: 2194-1017. DOI: 10.1007/978-3-319-32144-8_1. URL: http://dx.doi.org/10.1007/978-3-319-32144-8_1.

[2] Franz Achleitner, Anton Arnold, and Beatrice Signorello. "On Optimal Decay Estimates for ODEs and PDEs with Modal Decomposition". In: *Springer Proceedings in Mathematics & Statistics* (2019), 241–264. ISSN: 2194-1017. DOI: 10.1007/978-3-030-15096-9_6. URL: http://dx.doi.org/10.1007/978-3-030-15096-9_6.

[3] Kaïs Ammari and Carlos Castro. "Numerical Approximation of the Best Decay Rate for Some Dissipative Systems". In: *SIAM Journal on Numerical Analysis* 57.2 (2019), 681–701. ISSN: 1095-7170. DOI: 10.1137/17m1160057. URL: http://dx.doi.org/10.1137/17m1160057.

[4] Anton Arnold, Amit Einav, Beatrice Signorello, and Tobias Wöhrer. "Large Time Convergence of the Non-homogeneous Goldstein-Taylor Equation". In: *Journal of Statistical Physics* 182.2 (2021). ISSN: 1572-9613. DOI: 10.1007/s10955-021-02702-8. URL: http://dx.doi.org/10.1007/s10955-021-02702-8.

[5] Étienne Bernard and Francesco Salvarani. "On the exponential decay to equilibrium of the degenerate linear Boltzmann equation". In: *Journal of Functional Analysis* 265.9 (2013), 1934–1954. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2013.06.012. URL: http://dx.doi.org/10.1016/j.jfa.2013.06.012.

[6] Étienne Bernard and Francesco Salvarani. "Optimal Estimate of the Spectral Gap for the Degenerate Goldstein-Taylor Model". In: *Journal of Statistical Physics* 153.2 (2013). Correction published in 181, 1470-1471(2020), 363–375. ISSN: 1572-9613. DOI: 10.1007/s10955-013-0825-6. URL: http://dx.doi.org/10.1007/s10955-013-0825-6.

[7] Nawaf Bou-Rabee, Andreas Eberle, and Raphael Zimmer. "Coupling and convergence for Hamiltonian Monte Carlo". In: *Ann. Appl. Probab.* 30.3 (2020), pp. 1209–1250. ISSN: 1050-5164. DOI: 10.1214/19-AAP1528. URL: https://0-doi-org.pugwash.lib.warwick.ac.uk/10.1214/19-AAP1528.

[8] Carlos Castro and Steven J. Cox. "Achieving Arbitrarily Large Decay in the Damped Wave Equation". In: *SIAM Journal on Control and Optimization* 39.6 (2001), 1748–1755. ISSN: 1095-7138. DOI: 10.1137/s0363012900370971. URL: http://dx.doi.org/10.1137/s0363012900370971.

[9] Steven Cox and Enrique Zuazua. "The rate at which energy decays in a damped String". In: *Communications in Partial Differential Equations* 19.1-2 (1994), 213–243. ISSN: 1532-4133. DOI: 10.1080/03605309408821015. URL: http://dx.doi.org/10.1080/03605309408821015.

[10] Steven J. Cox. "Designing for Optimal Energy Absorption II, The Damped Wave Equation". In: *Control and Estimation of Distributed Parameter Systems* (1998), 103–109. DOI: 10.1007/978-3-0348-8849-3_8. URL: http://dx.doi.org/10.1007/978-3-0348-8849-3_8.

[11] Jean Dolbeault, Clément Mouhot, and Christian Schmeiser. "Hypocoercivity for linear kinetic equations conserving mass". In: *Trans. Amer. Math. Soc.* 367.6 (2015), pp. 3807–3828. ISSN: 0002-9947. DOI: `10.1090/S0002-9947-2015-06012-` URL: `http://dx.doi.org/10.1090/S0002-9947-2015-06012-7`.

[12] Andreas Eberle, Arnaud Guillin, and Raphael Zimmer. "Couplings and quantitative contraction rates for Langevin dynamics". In: *Ann. Probab.* 47.4 (2019), pp. 1982–2010. ISSN: 0091-1798. DOI: `10.1214/18-AOP1299`. URL: `https://0-doi-org.pugwash.lib.warwick.ac.uk/10.1214/18-AOP1299`.

[13] Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*. Vol. 194. Graduate Texts in Mathematics. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt. Springer-Verlag, New York, 2000, pp. xxii+586. ISBN: 0-387-98463-1.

[14] F. Fahroo and K. Ito. "Optimal absorption design for damped elastic systems". In: *2006 American Control Conference* (2006). DOI: `10.1109/acc.2006.1655331`. URL: `http://dx.doi.org/10.1109/acc.2006.1655331`.

[15] Mark Girolami and Ben Calderhead. "Riemann manifold Langevin and Hamiltonian Monte Carlo methods". In: *J. R. Stat. Soc. Ser. B Stat. Methodol.* 73.2 (2011). With discussion and a reply by the authors, pp. 123–214. ISSN: 1369-7412. DOI: `10.1111/j.1467-9868.2010.00765.x`. URL: `https://0-doi-org.pugwash.lib.warwick.ac.uk/10.1111/j.`

[16] Alain Haraux, Thibault Liard, and Yannick Privat. "How to estimate observability constants of one-dimensional wave equations? Propagation versus spectral methods". In: *Journal of Evolution Equations* 16.4 (2016), 825–856. ISSN: 1424-3202. DOI: `10.1007/s00028-016-0321-y`. URL: `http://dx.doi.org/10.1007/s00028-016-0321-y`.

[17] Pascal Hébrard and Antoine Henrott. "Optimal shape and position of the actuators for the stabilization of a string". In: *Systems & Control Letters* 48.3-4 (2003), 199–209. ISSN: 0167-6911. DOI: `10.1016/s0167-6911(02)00265-7`. URL: `http://dx.doi.org/10.1016/s0167-6911(02)00265-7`.

[18] Mark Kac. "A stochastic model related to the telegrapher's equation". In: *Rocky Mountain Journal of Mathematics* 4.3 (1974), 497–510. ISSN: 0035-7596. DOI: `10.1216/rmj-1974-4-3-497`. URL: `http://dx.doi.org/10.1216/rmj-1974-4-3-`

[19] G. Lebeau. "Équations des ondes amorties". In: *Séminaire sur les Équations aux Dérivées Partielles, 1993–1994*. École Polytech., Palaiseau, 1994, Exp. No. XV, 16.

[20] Arnaud Münch, Pablo Pedregal, and Francisco Periago. "Optimal design of the damping set for the stabilization of the wave equation". In: *Journal of Differential Equations* 231.1 (2006), 331–358. ISSN: 0022-0396. DOI: `10.1016/j.jde.2006.06.009`. URL: `http://dx.doi.org/10.1016/j.jde.2006.06.009`.

[21] Yannick Privat, Emmanuel Trélat, and Enrique Zuazua. "Optimal Observation of the One-dimensional Wave Equation". In: *Journal of Fourier Analysis and Applications* 19.3 (2013), 514–544. ISSN: 1531-5851. DOI: `10.1007/s00041-013-9267-4`. URL: `http://dx.doi.org/10.1007/s00041-013-9267-4`.

[22] Cédric Villani. "Hypocoercivity". In: *Mem. Amer. Math. Soc.* 202.950 (2009), pp. iv+141. ISSN: 0065-9266. DOI: `10.1090/S0065-9266-09-00567-5`. URL: `http://dx.doi.org/10.1090/S0065-9266-09-00567-5`.