ANALYSIS & PDE

Volume 10 No. 7 2017

msp

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2017 is US $265/year for the electronic version, and $470/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 10, No. 7, 2017
dx.doi.org/10.2140/apde.2017.10.1539

A VECTOR FIELD METHOD FOR RELATIVISTIC TRANSPORT EQUATIONS WITH APPLICATIONS

DAVID FAJMAN, JÉRÉMIE JOUDIOUX AND JACQUES SMULEVICI

We adapt the vector field method of Klainerman to the study of relativistic transport equations. First, we prove robust decay estimates for velocity averages of solutions to the relativistic massive and massless transport equations, without any compact support requirements (in x or v) for the distribution functions. In the second part of this article, we apply our method to the study of the massive and massless Vlasov–Nordström systems. In the massive case, we prove global existence and (almost) optimal decay estimates for solutions in dimensions $n \geq 4$ under some smallness assumptions. In the massless case, the system decouples and we prove optimal decay estimates for the solutions in dimensions $n \geq 4$ for arbitrarily large data, and in dimension 3 under some smallness assumptions, exploiting a certain form of the null condition satisfied by the equations. The 3-dimensional massive case requires an extension of our method and will be treated in future work.

MSC2010: 35B40, 35Q83, 83C30.
Keywords: relativistic kinetic equations, wave equation, vector-field method, asymptotic behaviour, nonlinear stability, Vlasov–Nordström system.

1. Introduction

The vector field method of Klainerman [1985b] provides powerful tools which are at the core of many fundamental results in the study of nonlinear wave equations, such as the famous proof of the stability of the Minkowski space [Christodoulou and Klainerman 1993]. In essence, the method takes advantage of the symmetries of a linear evolution equation to derive in a robust way boundedness and decay estimates of solutions. The robustness is crucial, as the final aim is typically to prove the nonlinear stability of some stationary solution, so that the method should be stable when perturbed by the nonlinearities of the equations.

In this paper, we are interested in the massive and massless relativistic transport equations[1]

$$\boldsymbol{T}_m(f) \equiv \left(\sqrt{m^2+|v|^2}\partial_t + v^i\partial_{x^i}\right)(f) = 0, \tag{1}$$

where $m \geq 0$ is the mass[2] of the particles and f is a function of (t, x, v) defined on $\mathbb{R}_t \times \mathbb{R}^n_x \times \mathbb{R}^n_v$ if $m > 0$ and $\mathbb{R}_t \times \mathbb{R}^n_x \times (\mathbb{R}^n_v \setminus \{0\})$ otherwise, with $n \geq 1$ being the dimension of the physical space.

Decay estimates via the method of characteristics for relativistic transport equations. For transport equations, the standard method to prove decay estimates is the method of characteristics. The origin of these decay estimates goes back in the nonrelativistic case to the work of Bardos and Degond [1985] on the Vlasov–Poisson system. Recall that if f is a regular solution to say $\boldsymbol{T}_1(f) = 0$ then, for all $(t, x, v) \in \mathbb{R}_t \times \mathbb{R}^n_x \times \mathbb{R}^n_v$,

$$f(t,x,v) = f\left(0, x - \frac{vt}{\sqrt{1+|v|^2}}, v\right),$$

and assuming that f has initially compact support in v, one can easily infer the velocity average estimate, for all $t > 0$ and all $x \in \mathbb{R}^n$,

$$\int_{v\in\mathbb{R}^n} |f|(t,x,v)\,dv \lesssim \frac{C(V)}{t^n}\|f(t=0)\|_{L^1_x L^\infty_v}, \tag{2}$$

[1] We will be using in the whole article the Einstein summation convention. For instance, $v^i\partial_{x^i}$ in (1) stands for $\sum_{i=1}^n v^i\partial_{x^i}$.

[2] In the remainder of this article, we will often normalize the mass to be either 1 or 0 and thus consider mostly $\boldsymbol{T}_1$ and $\boldsymbol{T}_0$. We will however sometimes keep the mass $m > 0$ so that the reader can see how some of the estimates would degenerate as $m \to 0$.

where V is an upper bound on the size of the support in v of f at the initial time and $C(V) \to +\infty$ as $V \to +\infty$.

These estimates, while being relatively easy to derive, suffer from several significant drawbacks when applied to a nonlinear system:

(1) They require a strong control of the characteristics of the system.

(2) The constant $C(V)$ in (2) depends on the size of the v-support of the solutions. Similar, more refined estimates, which do not require a compact support assumption, can nonetheless be derived (see [Schaeffer 2004]), but they are based on an even finer analysis of the characteristics.[3] This explains (partially) why most of the previous works assumed compact support in v. One therefore typically needs to bound an extra quantity, the size of the v-support at time t. In particular, this approach enlarges the number of variables of the system that need to be controlled.

Concerning the first problem, we note that there are many evolution problems for which the characteristics in a neighbourhood of a stationary solution will eventually diverge from the original ones, introducing extra difficulties in the analysis. A famous example of that is the stability of Minkowski space, where there is a logarithmic divergence; see [Christodoulou and Klainerman 1993; Lindblad and Rodnianski 2010]. Moreover, to prove decay estimates such as (2), one needs to control the Jacobian associated with the differential of the characteristic flow[4] and in order to obtain improved decay estimates for derivatives, one also needs estimates on the derivatives of the Jacobian. See, for instance, [Hwang et al. 2011], where such a program is carried out for the Vlasov–Poisson system. In other words, one needs strong control on the characteristics to be able to prove sharp decay estimates via this method in a nonlinear setting. Finally, note that there are many interesting models where the correct assumption, from the point of view of physics (see, for instance, the end of the Introduction in [Villani 2010]), is to allow arbitrarily large velocities.

Decay estimates for the wave equation. In the context of the wave equation

$$\Box\phi \equiv \left[-\partial_t^2 + \sum_{i=1}^{n} \partial_{x^i}^2\right]\phi = 0,$$

several methods exist to prove decay estimates of solutions. For instance, one standard way is to use the Fourier representation of the solution together with estimates for oscillatory integrals. In his fundamental paper, Klainerman [1985b] introduced what is now referred to as *the vector field method*.[5] Instead of relying on an explicit integral representation of the solutions, it uses:

(1) A coercive conservation law. In the case of the wave equation, this is simply the conservation of the energy.

[3] Note also that in [Schaeffer 2004], there is a loss of decay for the velocity averages of f compared to the linear case, which is directly related to the polynomial decay in v of the initial data and independent of the smallness assumptions.

[4] In the context of the Vlasov equation on a curved Lorentzian manifold, this means that one needs estimates on the differential of the exponential map, or at least on its restriction to certain submanifolds.

[5] Let us also mention that, complementary to the method of Klainerman, which uses vector fields as commutators, one can also use vector fields as *multipliers*, in the style of the work of Morawetz [1962; 1968].

(2) A set of vector fields which commute with the equations. In the case of the wave equation, these are the Killing and conformal Killing fields of the Minkowski space.

(3) Weighted Sobolev $L^2 - L^\infty$ inequalities. The standard derivatives ∂_t, ∂_{x^i} are rewritten in terms of the commutator vector fields before applying the usual Sobolev inequalities. The weights in these decompositions together with those arising from the conservation laws are then translated into decay rates.

This leads to the decay estimate

$$|\partial\phi(t,x)| \lesssim \frac{\mathcal{E}^{\frac{1}{2}}(\phi)}{\left(1+\left|t-|x|\right|\right)^{\frac{1}{2}}\left(1+\left|t+|x|\right|\right)^{\frac{n-1}{2}}} \tag{3}$$

for solutions of the wave equation $\square\phi = 0$, where $E(\phi)$ is an energy norm obtained by integrating ϕ and derivatives of ϕ (with weights) at time $t = 0$.

These types of estimates, being based on conservation laws and commutators, are quite robust, and as a consequence, are applicable in strongly nonlinear settings, such as the Einstein equations or the Euler equations; see, for instance, [Christodoulou 2007] for such an application.

A vector field method for transport equations. In our opinion, the decay estimate (2), being based on an explicit representation of the solutions, given by the method of characteristics, should be compared to the decay estimates for the wave equation obtained via the Fourier or other integral representations. In this paper, we derive an analogue of the vector field method for the massive and massless relativistic transport equations (1). The coercive conservation law is given by the conservation of the L^1-norm of the solution, while the vector fields commuting with the operators are essentially obtained by taking the *complete lifts* of the Killing and conformal Killing fields, a classical operation in differential geometry which takes a vector field on a manifold M to a vector field on the tangent bundle TM. The weighted Sobolev inequalities are slightly more technical. One of the main ingredients is that averages in v possess good commutation properties with the Killing vector fields and their complete lifts. Our decay estimates can then be stated as:

Theorem 1 (decay estimates for velocity averages of massless distribution functions). *For any regular distribution function f, solution to $\boldsymbol{T}_0(f) = 0$, and any $(t,x) \in \mathbb{R}^+_t \times \mathbb{R}^n_x$, we have*

$$\begin{aligned}\int_{v\in\mathbb{R}^n_v\setminus\{0\}} |f|(t,x,v)\frac{dv}{|v|} \\ \lesssim \frac{1}{\left(1+\left|t-|x|\right|\right)\left(1+\left|t+|x|\right|\right)^{n-1}} \sum_{\substack{|\alpha|\le n\\ \widehat{Z}^\alpha\in\widehat{\mathbb{K}}^{|\alpha|}}} \left\||v|^{-1}\widehat{Z}^\alpha(f)(t=0)\right\|_{L^1(\mathbb{R}^n_x\times(\mathbb{R}^n_v\setminus\{0\}))},\end{aligned} \tag{4}$$

where the α are multi-indices of length $|\alpha|$ and the $\widehat{Z}^\alpha$ are differential operators of order $|\alpha|$ obtained as a composition of $|\alpha|$ vector fields of the algebra $\widehat{\mathbb{K}}$.

The detailed list of the vector fields and their complete lifts used here is given in Section 2G1. For the massive transport equation, we prove:

Theorem 2 (decay estimates for velocity averages of massive distribution functions). *For any regular distribution function f, solution to $\boldsymbol{T}_1(f) = 0$, any $x \in \mathbb{R}^n$ and any $t \geq \sqrt{1+|x|^2}$, we have*

$$\int_{v \in \mathbb{R}^n_v} |f|(t,x,v) \frac{dv}{\sqrt{1+|v|^2}} \lesssim \frac{1}{(1+t)^n} \sum_{\substack{|\alpha| \leq n \\ \widehat{Z}^\alpha \in \widehat{\mathbb{P}}^{|\alpha|}}} \left\| \widehat{Z}^\alpha(f)_{|H_1 \times \mathbb{R}^n_v} \, v_\alpha \nu_1^\alpha \right\|_{L^1(H_1 \times \mathbb{R}^n_v)}, \tag{5}$$

where H_1 denotes the unit hyperboloid $H_1 \equiv \{(t,x) \in \mathbb{R}_t \times \mathbb{R}^n_x \mid 1 = t^2 - x^2\}$, $\widehat{Z}^\alpha(f)_{|H_1 \times \mathbb{R}^n_v}$ is the restriction to $H_1 \times \mathbb{R}^n_v$ of $\widehat{Z}^\alpha(f)$, $v_\alpha \nu_1^\alpha$ is the contraction of the 4*-velocity* $(\sqrt{1+|v|^2}, v^i)$ *with the unit normal ν_1 to H_1 and where the $\widehat{Z}^\alpha$ are differential operators obtained as a composition of $|\alpha|$ vector fields of the algebra $\widehat{\mathbb{P}}$.*

Remark 1.1. No compact support assumptions in x or in v are required for the statements of Theorems 1 or 2, but the norm on the right-hand side of (5) has two extra powers of v compared to the left-hand side of (5). Of course, for the norms on the right-hand sides of (4) or (5) to be finite, some amount of decay in x and v for the data is needed. Note that from the point of view of nonlinear applications, it is sufficient to propagate bounds for the norms appearing on the right-hand sides of (4) or (5), without any need to control pointwise the decay in x or v of the solutions, to get the desired decay estimates for the velocity averages.

Remark 1.2. In (4), the decay is worse near the light cone $t = |x|$. This of course is an analogue of the decay estimate (3) as traditionally obtained for the wave equation by the vector field method.

Remark 1.3. Estimate (5) is the analogue of the decay estimate for Klein–Gordon fields ϕ, solutions to $\Box\phi = \phi$, for which, for all $t \geq \sqrt{1+|x|^2}$,

$$|\partial\phi(t,x)| \lesssim \frac{\mathcal{E}^{\frac{1}{2}}[\phi]}{(1+t)^{\frac{n}{2}}},$$

where $E[\phi]$ is an energy norm obtained by integrating ϕ and derivatives of ϕ (with weights) on an initial hyperboloid, as obtained by Klainerman [1993].

Remark 1.4. As in the case of the Klein–Gordon equation, one can easily prove that for regular solutions f to $\boldsymbol{T}_1(f) = 0$ with data given at $t = 0$ and decaying sufficiently fast as $|x| \to +\infty$ (in particular, solutions arising from data with compact support in x) the norm on the right-hand side of (5) is finite, so that the decay estimate applies.[6] Thus, the use of hyperboloids is merely a technical issue. The restriction "$t \geq \sqrt{1+|x|^2}$" simply means in the future of the unit hyperboloid. We provide a classical construction in Appendix A, which explains how Theorem 2 can be applied to solutions arising from initial data with compact x-support given at $t = 0$ to obtain a $1/t^n$ decay of velocity averages in the whole future of the $t = 0$ hypersurface.

Remark 1.5. The reader might wonder whether the same types of techniques could be applied for the classical transport operator $\boldsymbol{T}_{\mathrm{cl}} = \partial_t + v^i \partial_{x^i}$. This question was addressed in [Smulevici 2016], where

[6] See also [Georgiev 1992], where decay estimates for the Klein–Gordon operator were obtained starting from noncompactly supported data at $t = 0$ using (mostly) vector-field-type methods.

decay estimates for velocity averages of solutions to the classical transport operator were obtained. As an application, that paper considered the study of small data solutions of the Vlasov–Poisson system and provided an alternative proof (with some additional information on the asymptotic behaviour of the solutions, concerning in particular the decay in $|x|$ and uniform bounds on some global norms) of the optimal time decay for derivatives of velocity averages obtained first in [Hwang et al. 2011]. One of the nice features of the vector field method is that improved decay estimates for derivatives follow typically easily from the main estimates, and [Smulevici 2016] was no exception. In the relativistic case, our vector field method also provides improved decay for derivatives. See Propositions 3.2 and 3.4 in Sections 3B and 3C, respectively.

Applications to the massive and massless Vlasov–Nordström systems. In the second part of this paper, we will apply our vector field method to the massive and massless Vlasov–Nordström systems

$$\Box\phi = m^2 \int_v f \frac{dv}{\sqrt{m^2+|v|^2}}, \tag{6}$$

$$\boldsymbol{T}_m(f) - (\boldsymbol{T}_m(\phi)v^i + m^2\nabla^i\phi)\frac{\partial f}{\partial v^i} = (n+1) f\, \boldsymbol{T}_m(\phi), \tag{7}$$

where $m = 0$ in the massless case and $m > 0$ in the massive case, $\Box \equiv -\partial_t^2 + \sum_{i=1}^n \partial_{x^i}^2$ is the standard wave operator of Minkowski space, ϕ is a scalar function of (t, x) and f is, as before, a function of (t, x, v^i) with $x \in \mathbb{R}^n$, $v \in \mathbb{R}^n$ if $m > 0$, $v \in \mathbb{R}^n \setminus \{0\}$ if $m = 0$. A good introduction to this system can be found in [Calogero 2003]. See also the classical works [Calogero 2006; Pallard 2006].

Roughly speaking, the Vlasov–Nordström system can be derived, in the context of scalar gravitation metric theory, by considering only a special class of metrics (that of metrics conformal to the Minkowski metric) and by neglecting some of the nonlinear terms in the equations for the gravitational field (see [Calogero 2003, Section 2] for a detailed discussion on the derivation). Since most of the simplifications concern difficulties which we already know how to handle, in the style of [Christodoulou and Klainerman 1993] or [Lindblad and Rodnianski 2010], and since the method that we are using here is of the same type as the one used to study the Einstein vacuum equations, we believe it is a good model problem before addressing the full Einstein–Vlasov system via vector field methods.

Before presenting our main results for the massive and massless Vlasov–Nordström systems, let us explain the main differences between the $m = 0$ and $m > 0$ cases. First, as easily seen from (6)–(7), when $m = 0$, the system degenerates to a partially decoupled system[7]

$$\Box\phi = 0, \tag{8}$$

$$\boldsymbol{T}_0(f) - (\boldsymbol{T}_0(\phi)v^i)\frac{\partial f}{\partial v^i} = (n+1) f\, \boldsymbol{T}_0(\phi). \tag{9}$$

Because of the decoupling, the first equation is simply the wave equation on Minkowski space and the second can be viewed as a linear transport equation, where the transport operator is the massless

[7]In fact, using $e^{-(n+1)\phi} f$ as an unknown, we can obtain an even simpler form of the equations where the right-hand side of (9) is put to 0. See (51).

relativistic transport operator plus some perturbations. In particular, all solutions are necessarily global as long as the initial data is sufficiently regular that the linear equations can be solved. Thus, our objective is solely to derive sharp asymptotics for the solutions of the transport equation. Moreover, since we have in mind future applications to more nonlinear problems, the only estimates that we will use on ϕ will be those compatible with what can be derived via a standard application of the vector field method.

Apart from the decoupling just explained, let us mention also two important pieces of structure present in the above equations. First, another great simplification comes from the existence of an *extra scaling symmetry* present only in the massless case: the vector field $v^i\partial_{v^i}$ commutes with the massless transport operator $\boldsymbol{T}_0$ and it is precisely this combination of derivatives in v which appears in equation (9). This fact will make all the error terms obtained after commutations much better than if a random set of derivatives in v was present in (9). Another property of (8)–(9) is the existence of a *null structure*, similar to the null structure of Klainerman for wave equations. More precisely, we show that $\boldsymbol{T}_0(\phi)$ has roughly the structure

$$\boldsymbol{T}_0(\phi) \simeq |v|\bar{\partial}\phi + \frac{1}{t}\,\partial\phi\cdot z(t,x,v),$$

where $\bar{\partial}\phi$ denotes derivatives tangential to the outgoing cone, $\partial\phi$ denotes arbitrary derivatives of ϕ and $z(t,x,v)$ are weights which are bounded along the characteristics of the linear massless transport operator. Since $\bar{\partial}\phi$ has better decay properties than a random derivative $\partial\phi$, we see that products of the form $\boldsymbol{T}_0(\phi)g$, where g is a solution to $\boldsymbol{T}_0(g)=0$, have better decay properties than expected.[8] Similar to the study of 3-dimensional wave equations with nonlinearities satisfying the null condition, the extra decay obtained means that in dimension 3 (or greater), all the error terms in the (approximate) conservation laws are now integrable.

We now state our main results for the massless Vlasov–Nordström system.

Theorem 3 (asymptotics in the massless case for dimension $n\geq 4$). *Let $n\geq 4$ and $N\geq \frac{3}{2}n+1$. Let ϕ be a solution of (8) satisfying $\phi(t=0)=\phi_0$ and $\partial_t\phi(t=0)=\phi_1$ for some sufficiently regular functions (ϕ_0,ϕ_1). Then, if $\mathcal{E}_N[\phi_0,\phi_1]<+\infty$, where $\mathcal{E}_N[\phi_0,\phi_1]$ is an energy norm containing up to N derivatives of $(\partial\phi_0,\phi_1)$ and if $E_N[f_0]<+\infty$, where $E_N[f_0]$ is a norm containing up to N derivatives of f_0, then the unique classical solution f to (9) satisfying $f(t=0)=f_0$ also satisfies*:

(1) *Global bounds. For all $t\geq 0$,*

$$E_N[f](t)\leq e^{C\mathcal{E}_N^{1/2}[\phi_0,\phi_1]}E_N[f_0],$$

where $C>0$ is a constant depending only on N, n.

[8]It is interesting to compare this form of the null condition to the one uncovered in [Dafermos 2006] for the massless Einstein–Vlasov system in spherical symmetry. In fact, two null conditions were used there. The obvious one consists essentially in understanding why null components of the energy momentum tensor of f decay better than expected. A more secret null condition is used in the analysis of the differential equation satisfied by the part of the velocity vector tangent to the outgoing cone. Our null condition is closely related to this one, even though we exploit it in a different manner since we are not using directly the characteristic system of ordinary differential equations associated with the transport equations.

(2) *Pointwise estimates for velocity averages. For all* $(t,x) \in [0,+\infty) \times \mathbb{R}^n_x$ *and all multi-indices* α *satisfying* $|\alpha| \leq N - n$,

$$\int_{v \in \mathbb{R}^n_v \setminus \{0\}} |\widehat{Z}^\alpha f|(t,x,v)|v|\,dv \lesssim \frac{e^{C\mathcal{E}_N^{1/2}[\phi_0,\phi_1]} E_N[f_0]}{\big(1+\big|t-|x|\big|\big)\big(1+\big|t+|x|\big|\big)^{n-1}}.$$

In dimension $n \geq 3$, similar results can be obtained provided an extra smallness assumption on the initial data for the wave function as well as stronger decay for the initial data of f hold.

Theorem 4 (asymptotics in the massless case for dimension $n=3$). *Let* $n \geq 3$, $N \geq 7$ *and* $q \geq 1$. *Let* ϕ *be a solution of* (8) *satisfying* $\phi(t=0) = \phi_0$ *and* $\partial_t \phi(t=0) = \phi_1$ *for some sufficiently regular functions* (ϕ_0, ϕ_1). *Then, if* $\mathcal{E}_N[\phi_0,\phi_1] \leq \varepsilon$, *where* $\mathcal{E}_N[\phi_0,\phi_1]$ *is an energy norm containing up to* N *derivatives of* $(\partial\phi_0, \phi_1)$ *and if* $E_{N,q}[f_0] < +\infty$, *where* $E_{N,q}[f_0]$ *is a norm*[9] *containing up to* N *derivatives of* f_0, *then the unique classical solution* f *to* (9) *satisfying* $f(t=0) = f_0$ *also satisfies*:

(1) *Global bounds with loss. For all* $t \geq 0$,

$$E_{N,q}[f](t) \leq (1+t)^{C\varepsilon^{1/2}} E_{N,q}[f_0],$$

where $C > 0$ *depends only on* N, n.

(2) *Improved global bounds for lower orders. For any* $M \leq N - \frac{1}{2}(n+2)$ *and any* $t \geq 0$,

$$E_{M,q-1}[f](t) \leq e^{C\varepsilon^{1/2}} E_{N,q}[f_0].$$

(3) *Pointwise estimates for velocity averages. For all* $(t,x) \in [0,+\infty) \times \mathbb{R}^n_x$ *and all multi-indices* $|\alpha| \leq N - \frac{1}{2}(3n+2)$,

$$\int_{v \in \mathbb{R}^n_v \setminus \{0\}} |\widehat{Z}^\alpha f|(t,x,v)|v|\,dv \lesssim \frac{E_{N,q}[f_0]}{\big(1+\big|t-|x|\big|\big)\big(1+\big|t+|x|\big|\big)^{n-1}}.$$

Perhaps counterintuitively, the massive case turned out to be harder to treat. While it is true that in the massive case the pointwise decay of velocity averages is not weaker along the null cone, there are two important extra difficulties, namely:

- The equations are now fully coupled. In particular, one cannot close an energy estimate for (6) unless we have some decay for the right-hand side. On the other hand, our decay estimates, being based on commutators, necessarily lose some derivatives. In turn, this would imply commuting (9) more, but we would then fail to close the estimates at the top order. We resolve this issue by another decay estimate for inhomogeneous transport equations with rough source terms satisfying certain product structures. This other type of decay estimate only provides L^2_x time decay of the velocity averages, which is precisely what is required to close the energy estimate for (6). The proof of this L^2_x decay estimate itself can be reduced to our L^∞ estimates, so that it can also be obtained using purely vector-field-type methods.

[9] The index q refers to powers of certain weights. See (63) for a precise definition of the norms.

- The vector field $v^i \partial_{v^i}$ does not commute with the massive transport operator. This implies that commuting with (some of) the standard vector fields will lose a power of t of decay compared to the massless case.

Because of the last issue, the results that we will present here are restricted to dimension $n \geq 4$. One way to treat the 3-dimensional case would be to improve upon the commutation formulae to eliminate the most dangerous terms. For instance, one could try to use *modified vector fields* in the spirit of [Smulevici 2016]. We plan to address the 3-dimensional case in future work.

A slightly more technical consequence of this last issue is that it introduces t weights in the estimates, which are not constant on the leaves of the hyperboloidal foliation that we wish to use. Together with the fact that the energy estimates are weaker on hyperboloids, this implies that the error terms arising in the top-order approximate conservation laws can be shown to be space-time integrable only in dimension $n \geq 5$. To address the dimension $n = 4$, instead of estimating directly $\widehat{Z}^\alpha(f)$, where f is the unknown distribution function and $\widehat{Z}^\alpha$ is a combination of α vector fields, we estimate instead a renormalized quantity of the form $\widehat{Z}^\alpha(f) + g^\alpha$, where the g^α is a (small) nonlinear term constructed from the solution. The extra terms appearing in the equation when the transport operator hits g^α will then cancel some of the worst terms in the equations.

Our main result in the massive case can then be stated as follows.

Theorem 5. *Let $n \geq 4$ and $m > 0$. Let $N \in \mathbb{N}$ be sufficiently large depending only on n. For any $\rho \geq 1$, denote by H_ρ the hyperboloid*

$$H_\rho = \{(t,x) \in \mathbb{R}_t \times \mathbb{R}^n_x \mid \rho^2 = t^2 - x^2\}.$$

For any sufficiently regular function ψ defined on $\mathbb{R}_t \times \mathbb{R}^n_x$, denote by $\psi_{|H_\rho}$ its restriction to H_ρ. Similarly, for any sufficiently regular function g defined on $\bigcup_{1 \leq \rho \leq +\infty} H_\rho \times \mathbb{R}^n_v$, denote by $g_{|H_\rho \times \mathbb{R}^n_v}$ its restriction to $H_\rho \times \mathbb{R}^n_v$. Then, there exists an $\varepsilon_0 > 0$ such that for any $0 < \varepsilon \leq \varepsilon_0$, if $E_{N+n}[f_0] + \mathcal{E}_N[\phi_0, \phi_1] \leq \varepsilon$, where $E_{N+n}[f_0]$ and $\mathcal{E}_N[\phi_0, \phi_1]$ are norms depending on respectively $N+n$ derivatives of f_0 and N derivatives of $(\partial\phi_0, \phi_1)$, then there exists a unique classical solution (f, ϕ) to (6)–(7) satisfying the initial conditions

$$\phi_{|H_1} = \phi_0, \quad \partial_t \phi_{|H_1} = \phi_1, \quad f_{|H_1 \times \mathbb{R}^n_v} = f_0$$

such that (f, ϕ) exists globally[10] *and satisfies the following estimates*:

(1) *Global bounds. For all $\rho \geq 1$,*

$$\mathcal{E}_N[\phi](\rho) \lesssim \varepsilon \quad \textit{and} \quad E_N[f][\rho] \lesssim \varepsilon \rho^{C\varepsilon^{1/4}},$$

where $C = 1$ when $n = 4$ and $C = 0$ when $n > 4$.

(2) *Pointwise decay for $\partial Z^\alpha \phi$. For all multi-indices $|\alpha|$ such that $|\alpha| \leq N - \frac{1}{2}(n+2)$ and all (t,x) with $t \geq \sqrt{1+|x|^2}$, we have*

$$|\partial Z^\alpha \phi| \lesssim \frac{\varepsilon}{(1+t)^{\frac{n-1}{2}}(1+(t-|x|))^{\frac{1}{2}}}.$$

[10] Here, globally means at every point lying in the future of the initial hyperboloid H_1. In 3 dimensions, this, of course, would already follow from [Calogero 2006] for regular initial data of compact support given on a constant-time slice.

(3) *Pointwise decay for $\rho(|\partial Z^\alpha f|)$. For all multi-indices α and β such that $|\alpha| \leq N-n$ and $|\beta| \leq \lfloor \frac{N}{2} \rfloor - n$ and all (t,x) with $t \geq \sqrt{1+|x|^2}$, we have*
$$\int_v |\widehat{Z}^\alpha f| \frac{dv}{v^0} \lesssim \frac{\varepsilon}{(1+t)^{n-C\varepsilon^{1/4}}} \quad \textit{and} \quad \int_v v^0 |\widehat{Z}^\beta f|\, dv \lesssim \frac{\varepsilon}{(1+t)^{n-C\varepsilon^{1/4}}},$$
where $C=1$ when $n=4$ and $C=0$ when $n>4$.

(4) *Finally, the following L^2-estimates on f hold. For all multi-indices α with $\lfloor \frac{N}{2} \rfloor - n + 1 \leq |\alpha| \leq N$, and all (t,x) with $t \geq \sqrt{1+|x|^2}$, we have*
$$\int_{H_\rho} \frac{t}{\rho} \left(\int_v |\widehat{Z}^\alpha f| \frac{dv}{v^0} \right)^2 d\mu_{H_\rho} \lesssim \varepsilon^2 \rho^{C\varepsilon^{1/4}-n},$$
where $C=2$ when $n=4$ and $C=0$ when $n>4$.

Remark 1.6. As for the linear decay estimates of Theorem 2, it is not essential to start on an initial hyperboloid for the conclusions of Theorem 5 to hold. In particular, an easy argument based on finite speed of propagation, similar to that given in Appendix A, shows that our method and results apply to the case of sufficiently small initial data with compact x-support given at $t=0$.

Remark 1.7. In [Friedrich 2004], solutions of the massive Vlasov–Nordström system in dimension 3 arising from small, regular, compactly supported (in x and v) data given at $t=0$ were studied and the asymptotics of velocity averages of the Vlasov field and up to two derivatives of the wave function were obtained. However, no estimates were obtained for derivatives of the Vlasov field or for higher derivatives of the wave function. Thus, [Friedrich 2004] is the analogue of [Bardos and Degond 1985] for the Vlasov–Nordström system, while we obtained here (in dimension 4 and greater) results more in the spirit of [Hwang et al. 2011; Smulevici 2016].

Remark 1.8. A posteriori, it is straightforward to propagate higher moments of the solutions in any of the situations of Theorems 3, 4 and 5, provided that these moments are finite initially. Moreover, we recall that improved decay for derivatives of f and ϕ follows from the statements of Theorems 3, 4 and 5. See, for instance, Propositions 3.2 and 3.4 below.

Aside: the Einstein–Vlasov system. As explained above, the Vlasov–Nordström system is a model problem for the more physically relevant Einstein–Vlasov system. We refer to the recent book[11] [Ringström 2013] for a thorough introduction to this system. The small data theory around the Minkowski space is still incomplete for the Einstein–Vlasov system. The spherically symmetric cases in dimension $(3+1)$ have been treated in [Rein and Rendall 1992] for the massive case and in [Dafermos 2006] for the massless case with compactly supported initial data. A proof of stability for the massless case without spherical symmetry but with compact support in both x and v was recently given by M. Taylor [2017]. As in [Dafermos 2006], the compact support assumptions and the fact that the particles are massless are important as they allow one to reduce the proof to that of the vacuum case outside from a strip going to null infinity. Interestingly, Taylor's argument is quite geometric, relying for instance on the double null

[11] Apart from a general introduction to the Einstein–Vlasov system, the main purpose of [Ringström 2013] is to present a proof of stability of exponentially expanding space-times for the Einstein–Vlasov system.

foliation, in the spirit of [Klainerman and Nicolò 2003], as well as several structures associated with the tangent bundle of the tangent bundle of the base manifold.

We hope to address the stability of the Minkowski space for the Einstein–Vlasov system in the massive and massless case (without the compact support assumptions) using the method developed in this paper in future works.

Structure of the article. Section 2 contains preliminary materials, such as basic properties of the transport operators, the definition and properties of the foliation by hyperboloids used for the analysis of the massive distribution function, the commutation vector fields and elementary properties of these vector fields. In Section 3, we introduce the vector field method for relativistic Vlasov fields and prove Theorems 1 and 2. In Section 4, we apply our method first to the massless case in dimension $n \geq 4$ (Section 4B3) and $n = 3$ (Section 4B4) and then to the massive case in dimension $n \geq 4$ (Section 4C). In Appendix A, we provide a classical construction which explains how our decay estimates in the massive case can be applied to data of compact support in x given at $t = 0$. Some integral estimates useful in the course of the paper are proven in Appendix B. Finally, Appendix C contains a general geometrical framework for the analysis of the Vlasov equation on a Lorentzian manifold.

2. Preliminaries

2A. *Basic notations.* Throughout this paper we work on the $(n+1)$-dimensional Minkowski space $(\mathbb{R}^{n+1}, \eta)$, where the standard Minkowski metric η is globally defined in Cartesian coordinates (t, x^i) by $\eta = \mathrm{diag}\{-1, 1, \dots, 1\}$. We denote space-time indices by Greek letters $\alpha, \beta, \dots \in \{0, \dots, n\}$ and spatial indices by Latin letters $i, j, \dots \in \{1, \dots, n\}$. We will sometimes use $\partial_{x^\alpha}, \partial_t, \partial_{x^i}, \partial_{v^i}, \dots$ to denote the partial derivatives $\partial/\partial x^\alpha, \partial/\partial t, \dots$.

Since we will be interested in either massive particles with $m = 1$ or massless particles $m = 0$, the velocity vector $(v^\beta)_{\beta=0,\dots,n}$ will be parametrized by $(v^i)_{i=1,\dots,n}$ and $v^0 = |v|$ in the massless case, $v^0 = \sqrt{1 + |v|^2}$ in the massive case.

The indices 0 and $m > 0$ will be used to denote objects corresponding to the massless and massive cases respectively, such as the massless transport operator $\boldsymbol{T_0}$ and the massive one $\boldsymbol{T_m}$, and should not be confused with spatial or space-time indices for tensor components (we use bold letters on the transport operators to avoid this confusion).

The notation $A \lesssim B$ will be used to denote an inequality of the form $A \leq CB$ for some constant $C > 0$ independent of the solutions (typically C will depend on the number of dimensions, the maximal order of commutations N, the value of the mass m).

2B. *The relativistic transport operators.* For any $m > 0$ and any $v \in \mathbb{R}^n$, let us define the *massive relativistic transport operator* $\boldsymbol{T_m}$ by

$$\boldsymbol{T_m} = v^0 \partial_t + v^i \partial_{x^i}, \quad \text{with } v^0 = \sqrt{m^2 + |v|^2}. \tag{10}$$

Similarly, we define for any $v \in \mathbb{R}^3 \setminus \{0\}$, the *massless transport operator* $\boldsymbol{T_0}$ by

$$\boldsymbol{T_0} = v^0 \partial_t + v^i \partial_{x^i}, \quad \text{with } v^0 = |v|. \tag{11}$$

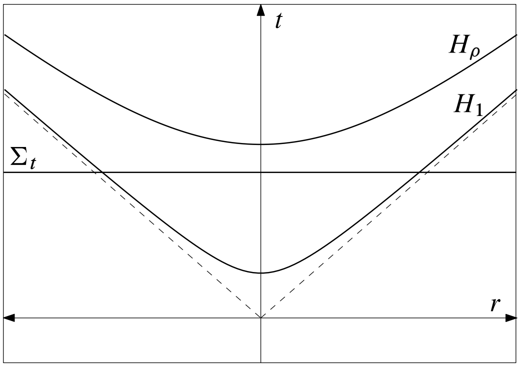


Figure 1. The H_ρ foliations in the (t, r) plane, $\rho > 1$.

For the sake of comparison, let us recall that the *classical* transport operator is given by

$$\boldsymbol{T}_{\mathrm{cl}} = \partial_t + v^i \partial_{x^i}.$$

In the remainder of this work, we will normalize the mass to be either 1 or 0, so that the massive transport operators we will study are

$$\boldsymbol{T}_1 = \sqrt{1+|v|^2}\partial_t + v^i \partial_{x^i} \quad \text{and} \quad \boldsymbol{T}_0 = |v|\partial_t + v^i \partial_{x^i}.$$

2C. *The foliations.* We will consider two distinct foliations of (some subsets of) the Minkowski space.

Let us fix global Cartesian coordinates (t, x^i), $1 \le i \le n$, on $\mathbb{R}^{n+1}$ and denote by Σ_t the hypersurface of constant t. The hypersurfaces Σ_t, $t \in \mathbb{R}$, then give a complete foliation of $\mathbb{R}^{n+1}$. The second foliation is defined as follows. For any $\rho > 0$, define H_ρ by

$$H_\rho = \{(t, x) \mid t \ge |x| \text{ and } t^2 - |x|^2 = \rho^2\}.$$

See Figure 1. For any $\rho > 0$, H_ρ is thus only one sheet of a two-sheeted hyperboloid.[12]

Note that

$$\bigcup_{\rho \ge 1} H_\rho = \{(t, x) \in \mathbb{R}^{n+1} \mid t \ge (1+|x|^2)^{\frac{1}{2}}\}.$$

The above subset of $\mathbb{R}^{n+1}$ will be referred to as *the future of the unit hyperboloid*; see Figure 2. On this set, we will use as an alternative to the Cartesian coordinates (t, x) the following two other sets of coordinates:

- Spherical coordinates. We first consider spherical coordinates (r, ω) on $\mathbb{R}^n_x$, where ω denotes spherical coordinates on the $(n-1)$-dimensional spheres and $r = |x|$. Then (ρ, r, ω) defines a coordinate system on the future of the unit hyperboloid. These new coordinates are defined globally on the future of the unit hyperboloid apart from the usual degeneration of spherical coordinates and at $r = 0$.

[12]The hyperboloidal foliation was originally introduced in [Klainerman 1985a] in the context of the nonlinear Klein–Gordon equation. For more recent applications, see [Wang 2015a; LeFloch and Ma 2016], which concern the stability of the Minkowski space for the Einstein–Klein–Gordon system.

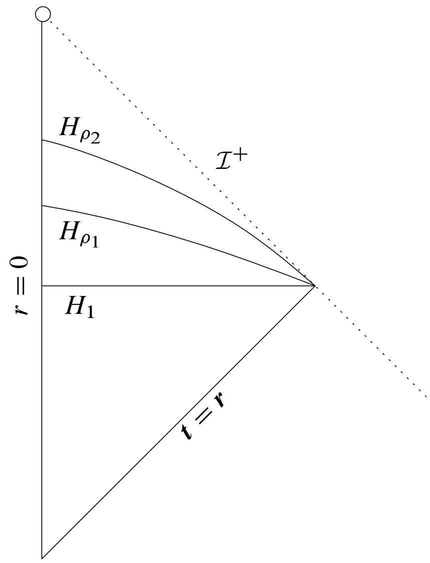


Figure 2. The H_ρ foliations in a Penrose diagram of Minkowski space, $\rho_2 > \rho_1 > 1$.

- Pseudo-Cartesian coordinates. These are the coordinates $(y^0, y^j) \equiv (\rho, x^j)$. These new coordinates are also defined globally on the future of the unit hyperboloid.

For any function defined on (some part of) the future of the unit hyperboloid, we will move freely between these three sets of coordinates.

2D. ***Geometry of the hyperboloids.*** The Minkowski metric η is given in (ρ, r, ω) coordinates by

$$\eta = -\frac{\rho^2}{t^2}(d\rho^2 - dr^2) - \frac{2\rho r}{t^2}\, d\rho\, dr + r^2 \sigma_{\mathbb{S}^{n-1}},$$

where $\sigma_{\mathbb{S}^{n-1}}$ is the standard round metric on the $(n-1)$-dimensional unit sphere, so that, for instance,

$$\sigma_{\mathbb{S}^2} = \sin\theta^2\, d\theta^2 + d\phi^2$$

in standard (θ, ϕ) spherical coordinates for the 2-sphere. The 4-dimensional volume form is thus given by

$$\frac{\rho}{t} r^{n-1}\, d\rho\, dr\, d\sigma_{\mathbb{S}^{n-1}},$$

where $d\sigma_{\mathbb{S}^{n-1}}$ is the standard volume form of the $(n-1)$-dimensional unit sphere.

The Minkowski metric induces on each of the H_ρ a Riemannian metric given by

$$ds^2_{H_\rho} = \frac{\rho^2}{t^2}\, dr^2 + r^2 \sigma_{\mathbb{S}^{n-1}}.$$

A normal differential form to H_ρ is given by $t\, dt - r\, dr$, while $t\partial_t + r\,\partial_r$ is a normal vector field. Since

$$\eta(t\partial_t + r\partial_r,\, t\partial_t + r\partial_r) = -\rho^2,$$

the future unit normal vector field to H_ρ is given by the vector field

$$\nu_\rho \equiv \frac{1}{\rho}(t\,\partial_t + r\,\partial_r). \tag{12}$$

Finally, the induced volume form on H_ρ, denoted by $d\mu_{H_\rho}$, is given by

$$d\mu_{H_\rho} = \frac{\rho}{t} r^{n-1}\,dr\,d\sigma_{\mathbb{S}^{n-1}}.$$

2E. *Regular distribution functions.* For the massive transport operator, we will consider distribution functions f as functions of (t, x, v) or (ρ, r, ω, v) defined on

$$\bigcup_{1\le\rho<P} H_\rho \times \mathbb{R}^n_v, \qquad P \in [1, +\infty];$$

i.e., we are looking at the future of the unit hyperboloid, or a subset of it, times $\mathbb{R}^n_v$.

For the massless transport operator, we need to exclude $|v| = 0$ and we will only use the Σ_t foliation so that we will consider distribution functions f as functions of (t, x, v) defined on $[0, T) \times \mathbb{R}^n_x \times (\mathbb{R}^n_v \setminus \{0\})$, $T \in [0, +\infty]$.

In the remainder of this article, we will denote by *regular* distribution function any such function f that is sufficiently regular that all the norms appearing on the right-hand sides of the estimates are finite. For simplicity, the reader might assume that f is smooth and decays fast enough in x and v at infinity and in the massless case, that f is integrable near $v = 0$ and similarly for the distribution functions obtained after commutations.

In physics, distribution functions represent the number of particles and are therefore required to be nonnegative. This will play no role in the present article, so we simply assume that distribution functions are real-valued.

2F. *The linear equations.* In the first part of this paper, we will study, for any $\boldsymbol{T} = \boldsymbol{T}_0, \boldsymbol{T}_1$, the homogeneous transport equation

$$\boldsymbol{T} f = 0, \tag{13}$$

as well as the inhomogeneous transport equation

$$\boldsymbol{T} f = v^0 h, \tag{14}$$

where $v^0 = \sqrt{1 + |v|^2}$ in the massive case and $v^0 = |v|$ in the massless case and where the source term h is a regular distribution function, as explained in Section 2E.

In the massless case, we will study the solution f to (13) or (14) with the initial data condition $f(t = 0, \cdot) = f_0$, where f_0 is a function defined on $\mathbb{R}^n_x \times (\mathbb{R}^n_v \setminus \{0\})$.

In the massive case, we will study the solution f to (13) or (14) in the future of the unit hyperboloid with the initial data condition $f_{|H_1\times\mathbb{R}^n_v} = f_0$, where f_0 is a function defined on $H_1 \times \mathbb{R}^n_v$.

Equations (13) and (14) are transport equations and can therefore be solved explicitly (at least for C^1 initial data) via the method of characteristics. If f solves (13), then

$$f(t, x, v) = f\left(0, x - \frac{v}{v^0}t, v\right),$$

where $v^0=\sqrt{1+|v|^2}$ for the massive case and $v^0=|v|$ for the massless case. In the inhomogeneous case, we obtain via the Duhamel formula that if f solves (14) with 0 initial data, then

$$f(t,x,v)=\int_0^t h\left(s,x-(t-s)\frac{v}{v^0},v\right)ds.$$

2G. *The commutation vector fields.*

2G1. *Complete lifts of isometries and conformal isometries.* Let us recall that the set of generators of isometries of the Minkowski space, that is to say, the set of Killing fields, denoted by $\mathbb{P}$, consists of the translations, the rotations and the hyperbolic rotations; i.e.,

$$\mathbb{P}=\{\partial_t,\partial_{x^1},\dots,\partial_{x^n}\}\cup\{\Omega_{ij}=x^i\partial_{x^j}-x^j\partial_{x^i}\mid 1\le i,j\le n\}\cup\{\Omega_{0i}=t\partial_{x^i}+x^i\partial_t\mid 1\le i\le n\}.$$

Mostly in the case of the massless transport operator, it will be useful, as in the study of the wave equation, to add the scaling vector field $S=t\partial_t+x^i\partial_i$ to our set of commutator vector fields. Let us thus define the set

$$\mathbb{K}=\mathbb{P}\cup\{S\}.$$

The vector fields in $\mathbb{K}$ and $\mathbb{P}$ lie in the tangent bundle of the Minkowski space. To any vector field on a manifold, we can associate a *complete lift*, which is a vector field lying on the tangent bundle, to the tangent bundle of the manifold. In Appendix C, we recall the general construction on a Lorentzian manifold. For the sake of simplicity, let us here give a working definition of the complete lifts only in coordinates.

Definition 2.1. Let W be a vector field of the form $W=W^\alpha\partial_{x^\alpha}$. Then let

$$\widehat{W}=W^\alpha\partial_{x^\alpha}+v^\beta\frac{\partial W^i}{\partial x^\beta}\partial_{v^i}, \tag{15}$$

where $(v^\beta)_{\beta=0,\dots,n}=(v^0,v^1,\dots,v^n)$ with $v^0=|v|$ in the massless case and $v^0=\sqrt{1+|v|^2}$ in the massive case, be called the *complete lift*[13] of W.

We will denote by

$$\widehat{\mathbb{K}}\equiv\{\widehat{Z}\mid Z\in\mathbb{K}\}\quad\text{and}\quad\widehat{\mathbb{P}}\equiv\{\widehat{Z}\mid Z\in\mathbb{P}\}$$

the sets of the complete lifts of $\mathbb{K}$ and $\mathbb{P}$.

Finally, let us also define $\widehat{\mathbb{P}}_0$ and $\widehat{\mathbb{K}}_0$ as the sets composed respectively of $\widehat{\mathbb{P}}$ and $\widehat{\mathbb{K}}$ and a scaling vector field[14] in (t,x) only:

$$\widehat{\mathbb{P}}_0\equiv\widehat{\mathbb{P}}\cup\{t\partial_t+x^i\partial_{x^i}\}=\widehat{\mathbb{P}}\cup\{S\}, \tag{16}$$

$$\widehat{\mathbb{K}}_0\equiv\widehat{\mathbb{K}}\cup\{t\partial_t+x^i\partial_{x^i}\}=\widehat{\mathbb{K}}\cup\{S\}. \tag{17}$$

[13]This is in fact a small abuse of notation, as, with the above definition, $\widehat{W}$ actually corresponds to the restriction of the complete lift of W to the submanifold corresponding to $v^0=\sqrt{1+|v|^2}$ in the massive case and $v^0=|v|$ in the massless case. See again Appendix C for a more precise definition of $\widehat{W}$.

[14]Here, by a small abuse of notation, we denote with the same letter S, the vector field $t\,\partial_t+x^i\partial_{x^i}$ irrespectively of whether we consider it as a vector field on $\mathbb{R}^{n+1}$ or a vector field on (some subsets of) $\mathbb{R}^{n+1}\times\mathbb{R}^n_v$.

Lemma 2.2. *In Cartesian coordinates, the complete lifts of the elements of* $\mathbb{P}$ *and* $\mathbb{K}$ *are given by the following formulae*:

$$\begin{gathered}\hat{\partial}_t = \partial_t, \quad \hat{\partial}_{x^i} = \partial_{x^i}, \\ \widehat{\Omega}_{ij} = x^i \partial_{x^j} - x^j \partial_{x^i} + v^i \partial_{v^j} - v^j \partial_{v^i}, \quad \widehat{\Omega}_{0i} = t\partial_{x^i} + x^i \partial_t + v^0 \partial_{v^i}, \\ \widehat{S} = t\partial_t + x^i \partial_{x^i} + v^i \partial_{v^i}.\end{gathered}$$

2G2. *Commutation properties of the complete lifts.* As for the wave equation, the symmetries of the Minkowski space are reflected in the transport operators (10) and (11) through the existence of commutation vector fields. More precisely,

Lemma 2.3.

- *Commutation rules for the massive transport operator*:

$$[\boldsymbol{T}_1, \widehat{Z}] = 0 \qquad \forall \widehat{Z} \in \widehat{\mathbb{P}}, \tag{18}$$

$$[\boldsymbol{T}_1, S] = \boldsymbol{T}_1, \tag{19}$$

where $S = t\partial_t + x^i \partial_{x^i}$ *is the usual scaling vector field.*

- *Commutation rules for the massless transport operator*:

$$[\boldsymbol{T}_0, \widehat{Z}] = 0 \qquad \forall \widehat{Z} \in \widehat{\mathbb{K}}, \tag{20}$$

$$[\boldsymbol{T}_0, S] = \boldsymbol{T}_0. \tag{21}$$

Proof. The identities can be verified directly using the explicit expressions for the elements in $\widehat{\mathbb{P}}$ and $\widehat{\mathbb{K}}$, but also follow from the general formula given in Appendix C (see Lemma C.7). □

Remark 2.4. Note that from the expression of $\widehat{S}$ and the two commutation rules for $\boldsymbol{T}_0$ and $\widehat{S}$ and for $\boldsymbol{T}_0$ and S, it follows that

$$[\boldsymbol{T}_0, v^i \partial_{v^i}] = -\boldsymbol{T}_0.$$

Thus, we have in a certain sense two scaling symmetries, one in x and one in v.

Remark 2.5. It is interesting to note that while the Klein–Gordon operator $\Box - m^2$ $(m > 0)$ does not commute with the scaling vector field, the massive transport equation does commute in the form of equation (19). What does not commute is the second scaling vector field $v^i \partial_{v^i}$.

We also have the following commutation relation within $\widehat{\mathbb{P}}_0$ and $\widehat{\mathbb{K}}_0$.

Lemma 2.6. *For any* $Z, Z' \in \widehat{\mathbb{P}}_0$, *there exist constant coefficients* $C_{ZZ'W}$ *such that*

$$[Z, Z'] = \sum_{W \in \widehat{\mathbb{P}}} C_{ZZ'W} W.$$

Similarly, for any $Z, Z' \in \widehat{\mathbb{K}}_0$, *there exist constant coefficients* $D_{ZZ'W}$ *such that*

$$[Z, Z'] = \sum_{W \in \widehat{\mathbb{K}}} D_{ZZ'W} W.$$

2H. *Weights preserved by the flow.* Recall that in a general Lorentzian manifold with metric g, if γ is a geodesic with tangent vector $\dot\gamma$ and K denotes a Killing field, then $g(\dot\gamma, K)$ is preserved along γ. In this section, we explain how to transpose this fact to the transport operators on Minkowski space.

We define the sets of weights

$$\mathbb{k}_m \equiv \{v^\alpha x^\beta - x^\alpha v^\beta, v^\alpha\}, \tag{22}$$

$$\mathbb{k}_0 \equiv \{x^\alpha v_\alpha, v^\alpha x^\beta - x^\alpha v^\beta, v^\alpha\}. \tag{23}$$

The following lemma can be easily checked.

Lemma 2.7. (1) *For all $\mathfrak{z} \in \mathbb{k}_0$, we have $[\boldsymbol{T}_0, \mathfrak{z}] = 0$.*

(2) *For all $\mathfrak{z} \in \mathbb{k}_m$, we have $[\boldsymbol{T}_m, \mathfrak{z}] = 0$.*

The weights in $\mathbb{k}_m$ and $\mathbb{k}_0$ also have good commutation properties with the vector fields in $\widehat{\mathbb{P}}_0$ and $\widehat{\mathbb{K}}_0$.

Lemma 2.8. *For any $\mathfrak{z} \in \mathbb{k}_m$ and any $\widehat{Z} \in \widehat{\mathbb{P}}_0$,*

$$[\widehat{Z}, \mathfrak{z}] = \sum_{\mathfrak{z}' \in \mathbb{k}_m} c_{\mathfrak{z}'} \mathfrak{z}',$$

where the $c_{\mathfrak{z}'}$ are constant coefficients.

Similarly for any $\mathfrak{z} \in \mathbb{k}_0$ and any $\widehat{Z} \in \widehat{\mathbb{K}}_0$,

$$[\widehat{Z}, \mathfrak{z}] = \sum_{\mathfrak{z}' \in \mathbb{k}_m} d_{\mathfrak{z}'} \mathfrak{z}'$$

for some constant coefficients $d_{\mathfrak{z}'}$.

Proof. This follows from straightforward computations. □

2I. *Multi-index notations.* Recall that a multi-index α of length $|\alpha|$ is an element of $\mathbb{N}^r$ for some $r \in \mathbb{N} \setminus \{0\}$ such that $\sum_{i=1}^r \alpha_i = |\alpha|$.

Let Z^i, $i = 1, \dots, 2n+2+\frac{1}{2}n(n-1)$, be an ordering of $\mathbb{K}$. For any multi-index α, we will denote by Z^α the differential operator of order $|\alpha|$ given by the composition $Z^{\alpha_1} Z^{\alpha_2} \cdots$.

In view of the above discussion, the complete lift operation defines a bijection between $\mathbb{K}$ and $\widehat{\mathbb{K}}$. Thus, to any ordering of $\mathbb{K}$, we can associate an ordering of $\widehat{\mathbb{K}}$. One extends this ordering to $\widehat{\mathbb{K}}_0$ by setting[15] $\widehat{Z}^{2n+3+\frac{1}{2}n(n-1)} = S$. We will again write $\widehat{Z}^\alpha$ to denote the differential operator of order $|\alpha|$ obtained by the composition $\widehat{Z}^{\alpha_1} \widehat{Z}^{\alpha_2} \cdots$.

Similarly, we consider an ordering of $\mathbb{P}$, which gives us an ordering of $\widehat{\mathbb{P}}$ which can be extended to give an ordering of $\widehat{\mathbb{P}}_0$, and we write Z^α and $\widehat{Z}^\alpha$ for a composition of $|\alpha|$ vector fields in $\mathbb{P}$, $\widehat{\mathbb{P}}$ or $\widehat{\mathbb{P}}_0$.

The notation $\mathbb{K}^N$ will be used to denote the set of all the differential operators of the form Z^α, with $|\alpha| = N$. Similarly, we will use the notations $\mathbb{P}^{|\alpha|}$, $\widehat{\mathbb{K}}_0^{|\alpha|}$ and $\widehat{\mathbb{P}}_0^{|\alpha|}$.

We will also write $\partial_{t,x}^\alpha$ to denote a differential operator of order $|\alpha|$ obtained as a composition of $|\alpha|$ translations among the ∂_t, ∂_{x^i} vector fields.

As for the sets of vector fields, we will also consider orderings of the sets of weights $\mathbb{k}_m$ and $\mathbb{k}_0$ and we will write $\mathfrak{z}^\alpha \in \mathbb{k}_m^{|\alpha|}$ or $\mathfrak{z}^\alpha \in \mathbb{k}_0^{|\alpha|}$ to denote a product of $|\alpha|$ weights in $\mathbb{k}_m$ or $\mathbb{k}_0$.

[15] Note that this is a small abuse of notation, since S is not obtained via the complete lift construction.

2J. *Vector field identities.* The following classical vector field identities will be used later in the paper.

Lemma 2.9. *The following identities hold*:

$$(t^2-r^2)\partial_t = tS - x^i\Omega_{0i},\quad (t^2-r^2)\partial_i = -x^j\Omega_{ij} + t\Omega_{0i} - x^iS,\quad (t^2-r^2)\partial_r = t\frac{x^i}{r}\Omega_{0i} - rS.$$

Furthermore,

$$\begin{aligned}\partial_s &\equiv \tfrac{1}{2}(\partial_t+\partial_r) = \frac{S+\omega^i\Omega_{0i}}{2(t+r)},\\ \bar{\partial}_i &\equiv \partial_i - \omega_i\partial_r = \frac{\omega^j\Omega_{ij}}{r} = \frac{-\omega_i\omega^j\Omega_{0j}+\Omega_{0i}}{t}.\end{aligned}\tag{24}$$

2K. *The particle vector field and the stress energy tensor of Vlasov fields.* Recall that in the Vlasov–Poisson or Einstein–Vlasov systems, the transport equation for f is coupled to an elliptic equation or a set of evolution equations, via integrals in v of f, often referred to as velocity averages in the classical case. In the relativistic cases, the volume forms[16] in these integrals are defined as

$$d\mu_m \equiv \frac{dv^1\wedge\cdots\wedge dv^n}{v^0} = \frac{dv}{\sqrt{m^2+|v|^2}},\tag{25}$$

where as usual $m=0$ in the massless case.

Remark 2.10. In the massless case, the volume form $dv/|v|$ is singular near $v=0$. In the remainder of this article, we will however study mostly energy densities, which introduce an additional factor of $|v|^2$ in the relevant integrals and thus remove this singular behaviour near $v=0$.

We now define the particle vector field in the case of massive particles as

$$N_m^\mu \equiv \int_{\mathbb{R}^n} f v^\mu\, d\mu_m,$$

and in the case of massless particles as

$$N_0^\mu \equiv \int_{\mathbb{R}^n\setminus\{0\}} f v^\mu\, d\mu_0,$$

as well as the energy momentum tensors

$$T_m^{\mu\nu} \equiv \int_{\mathbb{R}^n} f v^\mu v^\nu\, d\mu_m \quad\text{and}\quad T_0^{\mu\nu} \equiv \int_{\mathbb{R}^n\setminus\{0\}} f v^\mu v^\nu\, d\mu_0,$$

where $d\mu_m$ and $d\mu_0$ are the volume forms defined in (25). More generally, we can define the higher moments

$$M_m^{\alpha_1\cdots\alpha_p} \equiv \int_{\mathbb{R}^n} f v^{\alpha_1}\cdots v^{\alpha_p}\, d\mu_m,$$

and similarly for the massless system.

[16]For the interested reader, they can be interpreted geometrically as the natural volume forms associated with an induced metric on the manifold on which the averages are computed, together with a choice of normal in the massless case.

The interest in any of the above quantities is that if f is a solution to the associated massless or massive transport equations, then these quantities are divergence free. Indeed, we have

$$\partial_\mu T_{\mathbf{0}}^{\mu\nu} = \int_{\mathbb{R}^n \setminus \{0\}} \boldsymbol{T}_0(f) v^\nu \, d\mu_0, \tag{26}$$

$$\partial_\mu T_{\boldsymbol{m}}^{\mu\nu} = \int_{\mathbb{R}^n} \boldsymbol{T}_{\boldsymbol{m}}(f) v^\nu \, d\mu_m. \tag{27}$$

We will be interested in particular in the energy densities

$$\rho_0(f) \equiv T_0(\partial_t, \partial_t) = \int_{\mathbb{R}^n \setminus \{0\}} f |v| \, dv \tag{28}$$

for the massless case, while for the massive case we define

$$\rho_m(f) \equiv T_m(\partial_t, \partial_t) = \int_{\mathbb{R}^n} f v^0 \, dv. \tag{29}$$

In the following, we will denote by $\rho(f)$ either of the quantities $\rho_m(f)$ or $\rho_0(f)$ depending on whether we are looking at the massive or the massless relativistic operator.

In the massive case, we will also make use of the energy density

$$\chi_m(f) \equiv T_m(\partial_t, \nu_\rho), \tag{30}$$

where ν_ρ is the future unit normal to H_ρ introduced in Section 2D. We compute

$$\chi_{\mathrm{m}}(f) = \int_{v \in \mathbb{R}^n} f v_0 \left(\frac{t}{\rho} v_0 + \frac{r}{\rho} v^r \right) d\mu_m = \int_{v \in \mathbb{R}^n} f \left(\frac{t}{\rho} v^0 - \frac{x^i}{\rho} v_i \right) dv.$$

The following lemma will be used later.

Lemma 2.11 (coercivity of the energy density normal to the hyperboloids). *Assuming that $t \geq r$, we have*

$$\chi_m(f) \geq \frac{t}{2\rho} \int_{v \in \mathbb{R}^n} f \left[\left(1 - \frac{r}{t}\right) ((v^0)^2 + v_r^2) + r^2 \sigma_{AB} v^A v^B + m^2 \right] \frac{dv}{v^0}. \tag{31}$$

Proof. Using that

$$(v^0)^2 = v_r^2 + r^2 \sigma_{AB} v^A v^B + m^2,$$

where σ_{AB} denotes the components of the metric $\sigma_{\mathbb{S}^n}$ and v^A, v^B are the angular velocities, we have

$$(v^0)^2 = \tfrac{1}{2}(v^0)^2 + \tfrac{1}{2}(v_r^2 + r^2 \sigma_{AB} v^A v^B + m^2),$$

and thus

$$v^0 \left(\frac{t}{\rho} v^0 - \frac{x^i}{\rho} v_i \right) = \frac{t}{2\rho} \left((v^0)^2 + v_r^2 + r^2 \sigma_{AB} v^A v^B + m^2 - 2 \frac{x^i}{t} v_i v^0 \right).$$

The lemma now follows from

$$(v^0)^2 + v_r^2 - 2 \frac{x^i}{t} v_i v^0 \geq \left(1 - \frac{r}{t}\right) ((v^0)^2 + v_r^2),$$

assuming $t \geq r$. $\square$

Remark 2.12.

- Since $(v^0)^2 \geq v_r^2$, we will use (31) in the form

$$\chi_m(f) \geq \frac{t}{2\rho} \int_{v\in\mathbb{R}^n} f\left[\left(1-\frac{r}{t}\right)(v^0)^2 + r^2\sigma_{AB}v^A v^B + m^2\right] d\mu_m.$$

- We also remark that

$$\chi_m(|f|) \geq \tfrac{1}{2}m^2 \int_v |f| \frac{dv}{v^0} = \tfrac{1}{2}m^2 \rho_m\left(\frac{|f|}{(v^0)^2}\right),$$

since $t/(2\rho) \geq \frac{1}{2}$, and, furthermore,

$$\chi_m(|f|) \geq \frac{t-r}{2\rho}\rho_m(|f|) = \frac{\rho}{2(t+r)}\rho_m(|f|).$$

- Finally, independently of Lemma 2.11, since by the Cauchy–Schwarz inequality for Lorentzian metrics, as the vectors ν_ρ and v are both timelike future directed,

$$\left|\frac{tv^0 - x^i v_i}{\rho}\right| = |\langle v, \nu_\rho\rangle| \geq |v||\nu_\rho| = m, \quad \text{where } |v| = |g(v,v)|^{\frac{1}{2}},$$

we get immediately

$$\int_v |f|\,dv \leq \int_v \frac{1}{m}\left|\frac{tv^0 - x^i v_i}{\rho}\right| |f|\,dv = \frac{1}{m}\chi_m(|f|).$$

2L. *Commutation vector fields and energy densities.* Vector fields and the operator of averaging in v essentially commute in the following sense.

Lemma 2.13. *Let f be a regular distribution function for the massless case. Then*:

- *For any translation ∂_{x^α}, we have*

$$\partial_{x^\alpha}[\rho_0(f)] = \rho_0(\partial_{x^\alpha}(f)) = \rho_0(\hat{\partial}_{x^\alpha}(f)).$$

- *For any rotation Ω_{ij}, $1 \leq i, j, \leq n$, we have*

$$\Omega_{ij}[\rho_0(f)] = \rho_0(\widehat{\Omega}_{ij}(f)),$$

where $\widehat{\Omega}_{ij}$ is the complete lift of the vector field Ω_{ij}.

- *For any Lorentz boost Ω_{0i}, $1 \leq i \leq n$, we have*

$$\Omega_{0i}[\rho_0(f)] = \rho_0(\widehat{\Omega}_{0i}(f)) + 2\rho_0\left(\frac{v^i}{|v|}f\right).$$

- *For the scaling vector field S, we have*

$$S[\rho_0(f)] = \rho_0(\widehat{S}(f)) + (n+1)\rho_0(f).$$

- *Finally, all the above equalities hold (almost everywhere) with f replaced by $|f|$.*

Proof. Let us consider, for instance, a Lorentz boost $\Omega_{0i} = t\partial_{x^i} + x^i\partial_t$. Then

$$\Omega_{0i}[\rho_0(f)] = \int_v (t\partial_{x^i} + x^i\partial_t)(f)|v|\,dv. \tag{32}$$

On the other hand,

$$\begin{aligned}\int_v (t\partial_{x^i} + x^i\partial_t)(f)|v|\,dv &= \int_v (t\partial_{x^i} + x^i\partial_t + |v|\partial_{v^i})(f)|v|\,dv - \int_v |v|^2\partial_{v^i}(f)\,dv\\ &= \int_v \widehat{\Omega}_{0i}(f)|v|\,dv + 2\int_v \frac{v^i}{|v|}(f)|v|\,dv\\ &= \rho_0(\widehat{\Omega}_{0i}(f)) + 2\rho_0\left(\frac{v^i}{|v|}f\right),\end{aligned}$$

using an integration by parts in v^i. The other cases can all be treated similarly, the translations being trivial since $\hat{\partial}_{x^\alpha} = \partial_{x^\alpha}$. That f can be replaced by $|f|$ follows from the standard property of differentiation of the absolute value.[17] □

In the massive case, we have the following lemma, whose proof is left to the reader since it is very similar to the above.

Lemma 2.14. *Let f be a regular distribution function for the massive case. Then*:

- *For any translation ∂_{x^α}, we have*
$$\partial_{x^\alpha}[\rho_m(f)] = \rho_m(\partial_{x^\alpha}(f)) = \rho_m(\hat{\partial}_{x^\alpha}(f)).$$
- *For any rotation Ω_{ij}, $1 \le i, j, \le n$, we have*
$$\Omega_{ij}[\rho_m(f)] = \rho_m(\widehat{\Omega}_{ij}(f)),$$
where $\widehat{\Omega}_{ij}$ is the complete lift of the vector field Ω_{ij}.
- *For any Lorentz boost Ω_{0i}, $1 \le i \le n$, we have*
$$\Omega_{0i}[\rho_m(f)] = \rho_0(\widehat{\Omega}_{0i}(f)) + 2\rho_m\left(\frac{v^i}{v^0}f\right).$$
- *Finally, all the above equalities holds with f replaced by $|f|$.*

Remark 2.15. Although we do not have for all commutation vector fields $Z\rho = \rho\widehat{Z}$, we do have that $|Z\rho(|f|)| \lesssim \rho(|\widehat{Z}(f)|) + \rho(|f|)$ and this is all we shall need from the above. Note also that if we were looking at other moments, then similar formulae would hold with different coefficients. For instance, we have $\Omega_{0i}\int_v f\,d\mu_m = \int_v \widehat{\Omega}_{0i} f\,d\mu_m$ for sufficiently regular f.

Remark 2.16. In the massless case, we included the scaling vector field, but recall that $\boldsymbol{T}_0$ actually commutes with S (in the sense that $[\boldsymbol{T}_0, S] = \boldsymbol{T}_0$) so we will not really need to replace S by $\widehat{S}$. Note also that S enjoys good commutation properties with $\boldsymbol{T}_m$ and that $S\rho_m = \rho_m S$.

[17] Recall that $f \in W^{1,1}$ implies that $|f| \in W^{1,1}$ with $\partial|f| = (f/|f|)\,\partial f$ almost everywhere. See, for instance, [Lieb and Loss 1997, Chapter 6.17].

2M. ***(Approximate) conservation laws for Vlasov fields.*** The following lemma is easily verified.[18]

Lemma 2.17 (massless case). *Let h be a regular distribution function for the massless case in the sense of Section 2E. Let f be a regular solution to $\boldsymbol{T}_0(f) = v^0 h$, with $v^0 = |v|$, defined on $[0,T]\times\mathbb{R}^n_x\times(\mathbb{R}^n_v\setminus\{0\})$ for some $T>0$. Then, for all $t\in[0,T]$,*

$$\int_{\Sigma_t}\rho_0(f)(t,x)\,dx\left(\equiv\int_{x\in\mathbb{R}^n}\int_{v\in\mathbb{R}^n\setminus\{0\}}|v|f(t,x,v)\,dx\,dv\right)$$
$$=\int_{\Sigma_0}\rho_0(f)(0,x)\,dx+\int_0^t\int_{\Sigma_s}\rho_0(h)(s,x)\,dx\,ds, \tag{33}$$

and

$$\int_{\Sigma_t}\rho_0(|f|)(t,x)\,dx\le\int_{\Sigma_0}\rho_0(|f|)(0,x)\,dx+\int_0^t\int_{\Sigma_s}\rho_0(|h|)(s,x)\,dx\,ds. \tag{34}$$

Proof. The proof of (33) follows from an easy integration by parts (or an application of Stokes' theorem) and (26). A standard regularization argument of the absolute value allows to derive (34) in a similar manner. □

A similar identity holds for the massive case, but we shall need the following variant where we replace the Σ_t foliation by the H_ρ one.

Lemma 2.18 (massive case). *Let h be a regular distribution function for the massive case in the sense of Section 2E. Let f be a regular solution to $\boldsymbol{T}_m(f)=v^0h$, with $v^0=\sqrt{m^2+|v|^2}$, $m>0$, defined on $\bigcup_{\rho\in[1,P]}H_\rho\times\mathbb{R}^n_v$ for some $P>1$. Then, for all $\rho\in[1,P]$,*

$$\int_{H_\rho}\chi_m(f)(\rho,r,\omega)\,d\mu_{H_\rho}=\int_{H_1}\chi_m(f)(1,r,\omega)\,d\mu_{H_1}+\int_1^\rho\int_{H_\rho}\rho_m(h)(s,r,\omega)\,d\mu_{H_s}\,ds, \tag{35}$$

$$\int_{H_\rho}\chi_m(|f|)(\rho,r,\omega)\,d\mu_{H_\rho}\le\int_{H_1}\chi_m(|f|)(1,r,\omega)\,d\mu_{H_1}+\int_1^\rho\int_{H_\rho}\rho_m(|h|)(s,r,\omega)\,d\mu_{H_s}\,ds, \tag{36}$$

Proof. Again, the proof of (35) just follows from (27) and an integration by parts, while that of (36) follows similarly after a standard regularization argument. □

3. The vector field method for Vlasov fields

3A. ***The norms.*** We define, in the following, norms of distribution functions obtained from the standard conservation laws for the transport equations and the commutation vector fields introduced in the previous section.

Definition 3.1. • Let f be a regular distribution function for the massless case in the sense of Section 2E defined on $[0,T]\times\mathbb{R}^n_x\times(\mathbb{R}^n_v\setminus\{0\})$. For $k\in\mathbb{N}$, we define, for all $t\in[0,T]$,

$$\|f\|_{\mathbb{K},k}(t)\equiv\sum_{|\alpha|\le k}\sum_{\widehat{Z}^\alpha\in\widehat{\mathbb{K}}^{|\alpha|}}\int_{\Sigma_t}\rho_0(|\widehat{Z}^\alpha f|)(t,x)\,dx. \tag{37}$$

[18] Recall that if f is a regular solution to $\boldsymbol{T}(f)=v^0h$, then $|f|$ is a solution, in the sense of distributions, of $\boldsymbol{T}(|f|)=(f/|f|)v^0h$.

- Similarly, let f be a regular distribution function for the massive case in the sense of Section 2E defined on $\bigcup_{1\le\rho\le P} H_\rho \times \mathbb{R}^n_v$. For $k \in \mathbb{N}$, we define, for all $\rho \in [1, P]$,

$$\|f\|_{\mathbb{P},k}(\rho) \equiv \sum_{|\alpha|\le k} \sum_{\widehat{Z}^\alpha \in \widehat{\mathbb{P}}^{|\alpha|}} \int_{H_\rho} \chi_m(|\widehat{Z}^\alpha f|)\, d\mu_{H_\rho}. \tag{38}$$

3B. *Klainerman–Sobolev inequalities and decay estimates: massless case.* We are now ready to prove to following variant of the Klainerman–Sobolev inequalities.[19]

Theorem 6 (Klainerman–Sobolev inequalities for velocity averages of massless distribution functions). *Let f be a regular distribution function for the massless case defined on $[0, T] \times \mathbb{R}^n_x \times (\mathbb{R}^n_v \setminus \{0\})$ for some $T > 0$. Then, for all $(t, x) \in [0, T] \times \mathbb{R}^n_x$,*

$$\rho_0(|f|)(t, x) \lesssim \frac{1}{\big(1 + \big|t - |x|\big|\big)\big(1 + \big|t + |x|\big|\big)^{n-1}} \|f\|_{\mathbb{K},n}(t). \tag{39}$$

Proof. Let $(t, x) \in [0, T] \times \mathbb{R}^n_x$ and assume first that $|x| \notin \big[\tfrac{1}{2}t, \tfrac{3}{2}t\big]$ and $t + |x| \ge 1$. Let ψ be defined as

$$\psi : y \to \rho_0\big(|f|(t, x + (t + |x|)y)\big),$$

where $y = (y_1, y_2, \dots, y_n)$. Note that

$$\partial_{y_i}\psi(y) = \partial_{y^i}\big[\rho_0\big(|f|(t, x + (t + |x|)y\big)\big] = (t + |x|)\, \partial_{x^i}(\rho_0[|f|])(t, x + (t + |x|)y).$$

Assume now that $|y| \le \frac{1}{4}$. Using the fact that we are away from the light cone and the condition on $|y|$, it follows that

$$\frac{1}{C} \le \frac{\big|t + |x|\big|}{\big|t - |x + (t + |x|)y|\big|} \le C$$

for some $C > 0$. It then follows from the vector field identities of Lemma 2.9 that

$$\big|\partial_{y^i}\rho_0\big(|f|(t, x + (t + |x|)y)\big)\big| \lesssim \sum_{Z\in\mathbb{K}} \big|Z(\rho_0[|f|])\big|(t, x + (t + |x|)y).$$

From Lemma 2.13, we then obtain that

$$\begin{aligned}\big|\partial_{y^i}\rho_0\big[|f|(t,x+(t+|x|)y)\big]\big| &\lesssim \sum_{\substack{|\alpha|\le 1\\ \widehat{Z}^\alpha\in\widehat{\mathbb{K}}^{|\alpha|}}} \big|\rho_0[\widehat{Z}(|f|)]\big|(t,x+(t+|x|)y)+\rho_0(|f|)(t,x+(t+|x|)y)\\ &\lesssim \sum_{\substack{|\alpha|\le 1\\ \widehat{Z}^\alpha\in\widehat{\mathbb{K}}^{|\alpha|}}} \big|\rho_0[\widehat{Z}^\alpha(|f|)]\big|(t,x+(t+|x|)y)\end{aligned}$$

[19]Note that (in more than one spatial dimension) we cannot apply directly the standard Klainerman–Sobolev inequalities, in fact not even the usual Sobolev inequalities, to quantities such as $\rho(|f|)$ because of the lack of regularity of the absolute value. The aim of this section is therefore to explain how to circumvent this technical issue.

For a very clear introduction to Klainerman–Sobolev inequalities in the classical case of the wave equation, the interested reader may consult [Wang 2015b]. Some of the arguments below have been adapted from those notes.

$$\lesssim \sum_{\substack{|\alpha|\le 1\\ \widehat{Z}^\alpha\in\widehat{\mathbb{K}}^{|\alpha|}}} \rho_0\big[\big|\widehat{Z}^\alpha(|f|)\big|\big](t,x+(t+|x|)y)$$
$$\lesssim \sum_{\substack{|\alpha|\le 1\\ \widehat{Z}^\alpha\in\widehat{\mathbb{K}}^{|\alpha|}}} \rho_0\big[|\widehat{Z}^\alpha(f)|\big](t,x+(t+|x|)y),$$

where we have used in the last line that for any vector field $\widehat{Z}$, we have $|\widehat{Z}(|f|)| = |\widehat{Z}(f)|$ (almost everywhere and provided f is sufficiently regular), which essentially follows from the fact that $\partial|f| = (f/|f|)\,\partial f$ almost everywhere if $f\in W^{1,1}$. Let now $\delta = 1/(16n)$, so that if $|y_i|\le\delta^{\frac{1}{2}}$ for all $1\le i\le n$, we then have $|y|\le\frac{1}{4}$. Applying now a 1-dimensional Sobolev inequality in the variable y_1, we have

$$\begin{aligned}|\psi(0)| = \rho_0[|f|](t,x) &\lesssim \int_{|y_1|\le\delta^{1/2}} \big(\big|\partial_{y_1}\psi(y_1,0,\dots,0)\big| + \big|\psi(y_1,0,\dots,0)\big|\big)\,dy_1\\ &\lesssim \int_{|y_1|\le\delta^{1/2}} \Bigg(\sum_{\substack{|\alpha|\le 1\\ \widehat{Z}^\alpha\in\widehat{\mathbb{K}}^{|\alpha|}}} \rho_0\big[|\widehat{Z}^\alpha(f)|\big]\big(t,x+(t+|x|)(y_1,0,\dots,0)\big)\Bigg)\,dy_1.\end{aligned}$$

We can now apply a 1-dimensional Sobolev inequality in the variable y_2 and repeat the previous argument, with $|Z^\alpha(f)|$ replacing $|f|$, to obtain

$$|\psi(0)| \lesssim \int_{|y_1|\le\delta^{1/2}}\int_{|y_2|\le\delta^{1/2}} \Bigg(\sum_{\substack{|\alpha|\le 2\\ \widehat{Z}^\alpha\in\widehat{\mathbb{K}}^{|\alpha|}}} \rho_0\big[|\widehat{Z}^\alpha(f)|\big]\big(t,x+(t+|x|)(y_1,y_2,\dots,0)\big)\Bigg)\,dy_1\,dy_2.$$

Repeating the argument up to exhaustion of all variables, we obtain that

$$\begin{aligned}&\rho_0[|f|](t,x)\\ &\lesssim \int_{|y_1|\le\delta^{1/2}}\int_{|y_2|\le\delta^{1/2}}\cdots\int_{|y_n|\le\delta^{1/2}} \Bigg(\sum_{\substack{|\alpha|\le n\\ \widehat{Z}^\alpha\in\widehat{\mathbb{K}}^{|\alpha|}}} \rho_0\big[|\widehat{Z}^\alpha(f)|\big]\big(t,x+(t+|x|)(y_1,y_2,\dots,y_n)\big)\Bigg)\,dy_1\,dy_2\cdots dy_n.\end{aligned}$$

Applying the change of variable $z = (t+|x|)y$ gives us a $(t+|x|)^n$ factor which completes the proof of the inequality in this particular case. The case where $(t+|x|)\le 1$ follows from simpler considerations and is therefore left to the reader.

Let us thus turn to the case where $x\in\big[\frac{1}{2}t,\frac{3}{2}t\big]$ and $(t+|x|)\ge 1$. Note that it then follows that $t>\frac{2}{5}$ and $|x|>\frac{1}{3}$. Let us introduce spherical coordinates $(r,\omega)\in[0,+\infty)\times\mathbb{S}^{n-1}$ such that $x=r\omega$ and denote by q the optical function $q\equiv r-t$. Let $v(t,q,\omega)\equiv\rho_0(f)(t,(t+q)\omega)$.

Note that $\partial_q v = \partial_r\rho_0$ and $q\,\partial_q v = (r-t)\partial_r$ and that there exist constants C_{ij} such that

$$\partial_\omega v = \partial_\omega\big(\rho_0(f)(t,(q+t)\omega)\big) = \sum_{i<j} C_{ij}\Omega_{ij}\rho_0(f),$$

where the Ω_{ij} are the rotation vector fields.

Let $q_0 = |x| - t$. We need to prove that

$$t^{n-1}(1+|q_0|)\,|v(t,q_0,\omega)| \lesssim \|f\|_{\mathbb{K},n}(t).$$

Using a 1-dimensional Sobolev inequality, we have for any $\omega \in \mathbb{S}^{n-1}$,

$$|v(t,q_0,\omega)| \lesssim \int_{|q|<\frac{1}{4}t} \sum_{|\alpha|\le 1} \big|(\partial_q^\alpha v)(t,q+q_0,\eta)\big|\,dq.$$

Now

$$(\partial_q v)(t,q+q_0,\omega) = (\partial_r \rho_0(f))(t,q+q_0,\omega) = \rho_0(\partial_r(|f|)),$$

and thus

$$|\partial_q v(t,q+q_0,\omega)| \lesssim \rho_0(|\partial_r f|)(t,q+q_0,\omega),$$

where we have used again the properties of the derivatives of the absolute value. Let now $(\omega_1,\omega_2,\dots,\omega_{n-1})$ be a local coordinate patch in a neighbourhood of the point $\omega \in \mathbb{S}^{n-1}$. Using again a 1-dimensional inequality, we have

$$\begin{aligned}\big|\rho_0(|\partial_r^\alpha f|)(t,q+q_0,\omega)\big| \lesssim \int_{\omega_1} \big|\partial_{\omega_1}\rho_0(|\partial_r^\alpha f|)\big(t,q+q_0,\omega+(\omega_1,0,\dots,0)\big)\big|\,d\omega_1 \\ + \int_{\omega_1} \big|\rho_0(|\partial_r^\alpha f|)\big(t,q+q_0,\omega+(\omega_1,0,\dots,0)\big)\big|\,d\omega_1.\end{aligned}$$

Since ∂_{ω_1} can be rewritten in terms of the rotation vector fields, it follows from Lemma 2.13 that

$$\big|\partial_{\omega_1}\rho_0(|\partial_r^\alpha f|)\big| \lesssim \sum_{|\beta|\le 1} \rho_0\big(|\widehat{Z}^\beta \partial_r^\alpha f|\big).$$

Repeating until exhaustion of the number of variables on $\mathbb{S}^{n-1}$ and using that $\partial_r = (x^k/|x|)\partial_{x^k}$ and the commutation properties between $\widehat{Z}^\alpha$ and ∂_r, we obtain that

$$\big|\rho_0(|f|)(t,q+q_0,\omega)\big| \lesssim \int_{|q|\le\frac{1}{4}t} \int_{\eta\in\mathbb{S}^{n-1}} \sum_{|\alpha|\le n} \rho_0\big(|\widehat{Z}^\alpha(f)|\big)(t,q+q_0,\eta)\,dq\,d\sigma_{\mathbb{S}^{n-1}}.$$

Now since in the domain of integration $r = t+q+q_0 = q+|x| \sim t$, we have

$$\begin{aligned}t^{n-1}\big|\rho_0(|f|)(t,q+q_0,\omega)\big| &\lesssim \sum_{|\alpha|\le n} \int_{|q|\le\frac{1}{4}t} \int_{\eta\in\mathbb{S}^{n-1}} \rho_0\big(|\widehat{Z}^\alpha(f)|\big)(t,q+q_0,\eta) r^{n-1}\,dq\,d\sigma_{\mathbb{S}^{n-1}} \\ &\lesssim \sum_{|\alpha|\le n} \int_{\frac{1}{4}t\le r\le\frac{7}{4}t} \int_{\eta\in\mathbb{S}^{n-1}} \rho_0\big(|\widehat{Z}^\alpha(f)|\big)(t,r,\eta) r^{n-1}\,dr\,d\sigma_{\mathbb{S}^{n-1}} \\ &\lesssim \sum_{|\alpha|\le n} \int_{\frac{1}{4}t\le|y|\le\frac{7}{4}t} \rho_0\big(|\widehat{Z}^\alpha(f)|\big)(t,y)\,dy,\end{aligned}$$

which concludes the proof when $|q_0| \le 1$.

Assume now that $|q_0| > 1$. Let $\chi \in C_0^\infty\big(-\frac{1}{2},\frac{1}{2}\big)$ be a smooth cut-off function such that $\chi(0)=1$ and define $V_{q_0}(t,q,\omega) \equiv \chi((q-q_0)/q_0)v(t,q,\omega)$. To get the extra factor of $|q_0|$, we apply the method used above replacing the function v by the function $(s,\eta)\to V_{q_0}(t,q_0+q_0 s,\eta)$ and applying first a 1-dimensional Sobolev inequality in s on $|s|<\frac{1}{2}$. The extra powers of q_0 appearing are then absorbed since $|q_0+q_0 s|\sim|q_0|$ in the region of integration and since $(r-t)\partial_r$ can be expressed as a linear combination of commutation vector fields from Lemma 2.9 (with coefficients homogeneous of degree 0). The rest of the proof is similar to the one just given when $|q_0|\le 1$ and therefore omitted. □

Since the norm on the right-hand side is conserved for solutions of the homogeneous massless transport equations, we obtain in particular:

Theorem 7 (decay estimates for velocity averages of massless distribution functions). *Let f be a regular distribution function for the massless case, a solution to $\boldsymbol{T}_0(f)=0$ on $\mathbb{R}_t\times\mathbb{R}^n_x\times(\mathbb{R}^n_v\setminus\{0\})$. Then, for all $(t,x)\in\mathbb{R}_t\times\mathbb{R}^n_x$,*

$$\rho_0(|f|)(t,x)\lesssim\frac{1}{\big(1+\big|t-|x|\big|\big)\big(1+\big|t+|x|\big|\big)^{n-1}}\|f\|_{\mathbb{K},n}(0). \tag{40}$$

Finally, as for the wave equation, we have improved decay for derivatives of the solutions. More precisely, let $\partial=\partial_t$, let ∂_{x^i} be any translation, and let $\bar\partial$ be a derivative tangential to the cone $t=|x|$, such as $\partial_t+\partial_r$ or the projection on the angular derivatives of ∂_{x^i}, $\bar\partial_{x^i}=\partial_{x^i}-(x^i/r)\partial_r$. Then, we have the following proposition.

Proposition 3.2 (improved decay for derivatives of velocity averages of massless distribution functions). *Let f be a regular distribution function for the massless-case solution to $\boldsymbol{T}_0(f)=0$ on $\mathbb{R}_t\times\mathbb{R}^n_x\times(\mathbb{R}^n_v\setminus\{0\})$. Then, for all multi-indices l,k and for all $(t,x)\in\mathbb{R}_t\times\mathbb{R}^n_x$,*

$$|\partial^l\bar\partial^k\rho_0(f)(t,x)|\lesssim\frac{1}{\big(1+\big|t-|x|\big|\big)^{1+|l|}\big(1+\big|t+|x|\big|\big)^{n-1+|k|}}\|f\|_{\mathbb{K},n+k+l}(0). \tag{41}$$

Proof. This proof is similar to that of the improved decay estimates for the wave equation, and therefore omitted. □

Remark 3.3. Note that the improved decay estimates (41) only apply to velocity averages of f, because of the lack of regularity of velocity averages of $|f|$.

Finally, let us mention that we can obtain decay for other moments of the solutions, provided the corresponding moments for f and the $\widehat{Z}^\alpha(f)$ are finite initially. For instance, in Theorem 1 on page 1542, the decay estimate was written for the density of particles, while in Theorem 7, we considered the energy density. One can move freely from one to the other by considering $f|v|^q$ instead of f (provided the initial data can support it of course).

3C. *Klainerman–Sobolev inequalities and decay estimates: massive case.*

In the massive case $m>0$, we will prove:

Theorem 8 (Klainerman–Sobolev inequalities for velocity averages of massive distribution functions). *Let f be a regular distribution function for the massive case defined on $\bigcup_{1\le\rho<P}H_\rho\times\mathbb{R}^n_v$ for some*

$P \in [1, +\infty]$. *Then, for all* $(t, x) \in \bigcup_{1 \le \rho < P} H_\rho$,

$$\int_{v \in \mathbb{R}^n} |f|(t, x, v) \frac{dv}{v^0} \lesssim \frac{1}{(1+t)^n} \|f\|_{\mathbb{P},n}(\rho(t, x)), \tag{42}$$

where $\rho(t, x) = (t^2 - |x|^2)^{\frac{1}{2}}$ *and the norm* $\|f\|_{\mathbb{P},n}$ *is defined as in Section 3A.*

Proof. Recall from Remark 2.12 that

$$\chi_m(|f|)(t, x) \ge m^2 \frac{t}{2\rho} \int_v f \, d\mu_m = m^2 \frac{t}{2\rho} \int_{v \in \mathbb{R}^n} |f|(t, x, v) \frac{dv}{v^0}, \tag{43}$$

and thus,

$$\int_{H_\rho} \chi_m(|f|)(t, x) d\mu_{H_\rho} \ge \int_{H_\rho} m^2 \frac{t}{2\rho} \int_{v \in \mathbb{R}^n} |f|(t, x, v) \frac{dv}{v^0} \, d\mu_{H_\rho}. \tag{44}$$

Let (t, x) be fixed in $\bigcup_{1 \le \rho \le P} H_\rho$ and define the function ψ in the (y^α)-system of coordinates (see the end of Section 2C) as

$$\psi(y^0, y^j) \equiv \int_{v \in \mathbb{R}^n} |f|(y^0, x^j + t y^j) \, d\mu_m.$$

Similarly to the proof of the massless case, we apply first a 1-dimensional Sobolev inequality in the variable y^1

$$\int_{v \in \mathbb{R}^n} |f|(y^0, x^j) \, d\mu_m = |\psi(y^0, 0)| \lesssim \int_{|y^1| \le 1/(8n)^{\frac{1}{2}}} \big[|\partial_{y^1} \psi|(y^0, y^1, 0, \dots, 0) + |\psi|(y^0, y^1, 0, \dots, 0) \big] \, dy^1.$$

Now

$$\partial_{y^1} \psi = \frac{t}{t(y^0, x^1 + t y^1, x^2, \dots, x^n)} \Omega_{01},$$

where the t in the numerator is that of the point (t, x), while $t(y^0, x^j + t y^j) \equiv \big((y^0)^2 + (x^j + t y^j)^2\big)^{\frac{1}{2}}$ is the time of the point defined in the y^α-coordinates by $(y^0, x^j + t y^j)$. Now if $|x| \le \frac{1}{2} t$, then it follows from the condition $|y^1| \le 1/(8n^{\frac{1}{2}}) \le \frac{1}{8}$ that $(y^0)^2 \ge \frac{3}{4} t^2$ and thus that

$$\left| \frac{t}{t(y^0, x^1 + t y^1, x^2, \dots, x^n)} \right| \le C$$

for some uniform $C > 0$. On the other hand if $|x| \ge \frac{1}{2} t$, then it follows from the condition $|y^1| \le 1/(8n^{\frac{1}{2}}) \le \frac{1}{8}$ that $|x^j + t y^j| \ge \frac{3}{8} t$, where $y^j = (y^1, 0, \dots, 0)$. Thus, we have, for $|y^1| \le 1/(8n^{\frac{1}{2}})$,

$$|\partial_{y^1} \psi|(y^0, y^1, 0, \dots, 0) \lesssim \left| \int_v \Omega_{01} |f|(y^0, x^1 + t y^1, x^2, \dots, x^n, v) \, d\mu_m \right|.$$

The remainder of the proof is then similar to the massless case. We have

$$\left| \int_v \Omega_{01}(|f|)(y^0, x^1 + t y^1, x^2, \dots, x^n, v) \, d\mu_m \right| \lesssim \sum_{|\alpha| \le 1} \int_v |\hat{Z}^\alpha f|(y^0, x^1 + t y^1, x^2, \dots, x^n, v) \, d\mu_m.$$

Inserting in the Sobolev inequality and repeating up to exhaustion of all the variables (the fact that, for all j, $|y^j| \leq 1/(8n^{\frac{1}{2}})$ guarantees that $|y| = \left(\sum_{j=1}^n |y^j|^2\right)^{\frac{1}{2}} \leq \frac{1}{8}$ so that we still have $t/(t(y^0, x^j + ty^j)) \sim 1$), we obtain

$$\int_v |f|(y^0, x^1, x^2, \cdots, x^n, v)\, d\mu_m \lesssim \sum_{|\alpha| \leq n} \int_{|y| \leq \frac{1}{8}} \int_v |\hat{Z}^\alpha f|(y^0, x^j + ty^j, v)\, d\mu_m\, dy.$$

Recall that the volume form on each of the H_ρ is given in spherical coordinates by $(\rho/t) r^{n-1}\, dr\, d\sigma$, or in y^α-coordinates by $(y^0/t)\, dy$. Thus, we have

$$\begin{aligned}\int_v |f|(y^0, x^1, x^2, \ldots, x^n, v)\, d\mu_m &\lesssim \sum_{|\alpha| \leq n} \int_{|y| \leq \frac{1}{8}} \int_v |\hat{Z}^\alpha f|(y^0, x^j + ty^j, v)\, d\mu_m \frac{t(y^0, x^j + ty^j)}{y^0}\, d\mu_{H_\rho} \\ &\lesssim \frac{t(y^0, x^j)}{y^0} \sum_{|\alpha| \leq n} \int_{|y| \leq \frac{1}{8}} \int_v |\hat{Z}^\alpha f|(y^0, x^j + ty^j, v)\, d\mu_m\, d\mu_{H_\rho},\end{aligned}$$

where we have used again that $t(y^0, x^j + ty^j) \sim t(y^0, x^j)$ in the region of integration. Applying the change of coordinates $z^j = ty^j$ and noticing that the quantities on the right-hand side are controlled by the estimate (44) applied to $\hat{Z}^\alpha(f)$ completes the proof. □

Since the norm on the right-hand side of (42) is conserved if f is a solution to the massive transport equation, we obtain, as a corollary, the following pointwise decay estimate.

Theorem 9 (pointwise decay estimates for velocity averages of massive distribution functions). *Let f be a regular distribution function for the massive case satisfying the massive transport equation $\boldsymbol{T_m}(f) = 0$ on $\bigcup_{1 \leq \rho < +\infty} H_\rho \times \mathbb{R}^n_v$. Then, for all $(t, x) \in \bigcup_{1 \leq \rho < +\infty} H_\rho$,*

$$\int_{v \in \mathbb{R}^n} |f|(t, x, v) \frac{dv}{v^0} \lesssim \frac{1}{(1+t)^n} \|f\|_{\mathbb{P},n}.$$

Finally, let us mention the following improved decay for derivatives.

Proposition 3.4 (improved decay estimates for derivatives of velocity averages of massive distribution functions). *Let f be a regular distribution function for the massive case satisfying the massive transport equation $\boldsymbol{T_m}(f) = 0$ on $\bigcup_{1 \leq \rho < +\infty} H_\rho \times \mathbb{R}^n_v$. Then, for all $i \in \mathbb{N}$, for all multi-indices l and for all $(t, x) \in \bigcup_{1 \leq \rho < +\infty} H_\rho$,*

$$\left| \nu_\rho^i \partial_y^l \int_{v \in \mathbb{R}^n} f(t, x, v) \frac{dv}{v^0} \right| \lesssim \frac{1}{(1+t)^{n+|l|} \rho^i} \|f\|_{\mathbb{P},n+i+l},$$

where $\nu_\rho = x^\alpha \partial_{x^\alpha}/\rho$ is the future unit normal to H_ρ and ∂_y^l is a combination of $|l|$ vector fields among the ∂_{y^k}, $1 \leq k \leq n$, which are tangent to the H_ρ.

Proof. We have $\nu_\rho = S/\rho$ with S the scaling vector field. On the other hand, recall that S essentially commutes with the massive transport operator, so that in particular $\boldsymbol{T_m}(S(f)) = 0$ if $\boldsymbol{T_m}(f) = 0$. Thus, $\int_v S(f)/v^0\, dv = S\left(\int_v f/v^0\, dv\right)$ satisfies the same decay estimates as $\int_v f/v^0\, dv$, which shows the improved decay for $\nu_\rho(\int_v f/v^0\, dv)$. The higher-order derivatives follow similarly. Indeed, using that

$S(\rho) = \rho$, we have for instance $S^2(f) = \rho^2 \nu_\rho^2(f) + S(f)$. Applying the decay estimates for the velocity averages of $S^2(f)$ and $S(f)$ gives the correct improved decay for velocity averages of $\nu_\rho^2(f)$. Higher normal derivatives can be treated similarly. Finally, the improved decay for tangential derivatives of velocity averages is an easy consequence of the fact that $\partial_{y^k} = (1/t)\Omega_{0k}$. □

4. Applications to the Vlasov–Nordström system

4A. ***Generalities on the Vlasov–Nordström system.*** In 1913, Nordström introduced a gravitation theory based on the replacement of the Poisson equation by a scalar wave equation. The Vlasov–Nordström system describes the coupling of this gravitational theory with collisionless matter.[20]

It can be roughly obtained from the Einstein–Vlasov equations within the class of metrics conformal to the Minkowski metric by neglecting some of the nonlinear self-interactions of the conformal factor. In dimension $n = 3$, global existence for sufficiently regular massive distribution functions, with compact support in (x, v), has been proven in [Calogero 2006].

Following [Calogero 2003], it is possible to make a derivation of this system for arbitrary mass, as well as arbitrary dimension. Consider the metric

$$g = e^{2\phi}\eta$$

conformal to the Minkowski metric η, where ϕ is a function on $\mathbb{R}^{n+1}$. For this system, the mass shell is defined by the equation

$$e^{2\phi}\eta_{\alpha\beta}v^\alpha v^\beta = -m^2, \quad \text{which provides } v^0 = \sqrt{m^2 e^{-2\phi} + \eta_{ij}v^i v^j}.$$

We can introduce the coordinates

$$\hat{v}^i = e^\phi v^i,$$

which consistently also provides

$$\hat{v}^0 = \sqrt{m^2 + \eta_{ij}v^i v^j} = e^\phi v^0.$$

Considering distributions of particles which are conserved along the geodesic flow of g, we can define the associated transport operator as

$$\boldsymbol{T}_g \equiv v^\alpha\left(\frac{\partial}{\partial x^\alpha} - v^\beta \Gamma^i_{\alpha\beta}\frac{\partial}{\partial v^i}\right),$$

where $\Gamma^i_{\alpha\beta}$ are the Christoffel symbols of the metric g, which are given by

$$\Gamma^i_{\alpha\beta} = \delta^i_\alpha \frac{\partial\phi}{\partial x^\beta} + \delta^i_\beta \frac{\partial\phi}{\partial x^\alpha} - \eta_{\alpha\beta}\frac{\partial\phi}{\partial x^i},$$

so that

$$\boldsymbol{T}_g = v^\alpha \frac{\partial}{\partial x^\alpha} - (2v^\alpha \nabla_\alpha \phi v^i + m^2 e^{-2\phi}\nabla^i\phi)\frac{\partial}{\partial v^i}.$$

In the $(t, x, \hat{v})$-system of coordinates, we compute

$$\boldsymbol{T}_g = e^{-\phi}\left(\hat{v}^\alpha \frac{\partial}{\partial x^\alpha} - (\hat{v}^\alpha \nabla_\alpha\phi \hat{v}^i + m^2 \nabla^i\phi)\frac{\partial}{\partial \hat{v}^i}\right).$$

[20] See [Calogero 2003] for an introduction to the system.

To couple the Vlasov field and the scalar function ϕ, we follow [Calogero 2003] and require that[21]

$$\Box\phi = m^2 e^{(n+1)\phi} \int_v f \frac{dv}{v^0}. \tag{45}$$

Depending on the value of the mass m, we are thus faced with the following two systems:

- The massless Vlasov–Nordström system.

$$\Box\phi = 0, \tag{46}$$

$$\hat{v}^\alpha \frac{\partial f}{\partial x^\alpha} - \hat{v}^\alpha \nabla_\alpha \phi \hat{v}^i \frac{\partial f}{\partial \hat{v}^i} = 0. \tag{47}$$

  In this case, the equations decouple. We can of course solve the first equation and then think of the second equation as a linear transport equation for f.

- The massive Vlasov–Nordström system. In this case, we can perform yet another change of unknowns by considering

$$\tilde{f}(t, x, \hat{v}) \equiv e^{(n+1)\phi} f(t, x, \hat{v}),$$

  which has the advantage of removing the ϕ-dependence in the right-hand side of equation (45).

  We then obtain the usual expression of the (massive) Vlasov–Nordström system

$$\Box\phi = m^2 \int_{\hat{v}} \tilde{f} \frac{d\hat{v}}{\hat{v}^0}, \tag{48}$$

$$\hat{v}^\alpha \frac{\partial \tilde{f}}{\partial x^\alpha} - (\hat{v}^\alpha \nabla_\alpha \phi \hat{v}^i + m^2 \nabla^i \phi) \frac{\partial \tilde{f}}{\partial \hat{v}^i} = (n+1) \tilde{f} \hat{v}^\alpha \frac{\partial \phi}{\partial x^\alpha}. \tag{49}$$

  From now on, we will drop the ˆ and ˜ on all the variables to ease the notations.

4B. ***The massless Vlasov–Nordström system.*** We consider in this section, the system (46)–(47). We will denote by $\boldsymbol{T}_\phi$ the transport operator defined by

$$\boldsymbol{T}_\phi \equiv v^\alpha \frac{\partial}{\partial x^\alpha} - v^\alpha \nabla_\alpha \phi \cdot v^i \frac{\partial}{\partial v^i};$$

i.e.,

$$\boldsymbol{T}_\phi = \boldsymbol{T}_0 - \boldsymbol{T}_0(\phi) \cdot v^i \partial_{v^i}.$$

The massless Vlasov–Nordström system can then be rewritten as

$$\Box\phi = 0, \tag{50}$$

$$\boldsymbol{T}_\phi(f) = 0, \tag{51}$$

which we complement by the initial conditions

$$\phi(t=0) = \phi_0, \quad \partial_t \phi(t=0) = \phi_1, \tag{52}$$

$$f(t=0) = f_0, \tag{53}$$

[21]It should not be surprising that the right-hand side of this equation vanishes for massless distribution functions, as, up to an overall factor, it corresponds to the trace of the energy-momentum tensor, the latter being of course proportional to m for Vlasov fields.

where (ϕ_0, ϕ_1) are sufficiently regular functions defined on $\mathbb{R}^n_x$ and f_0 is a sufficiently regular function defined on $\mathbb{R}^n_x \times (\mathbb{R}^n_v \setminus \{0\})$.

By sufficiently regular, we mean that all the computations below make sense. We will eventually require that $\mathcal{E}_N[\phi_0, \phi_1] < +\infty$, where $\mathcal{E}_N$ is the energy norm defined by (57) and similarly, we will also require below that $\| f_0\|_{\mathbb{K},N} < +\infty$ (with some additional weights in the case of dimension 3). Provided N is large enough (depending only on n), these two regularity requirements are then enough so that all the computations below are justified. In the remaining, we will therefore omit any further mention of regularity issues.

4B1. *Commutation formula for $\boldsymbol{T}_\phi$.* Recall the algebra of commutation fields $\widehat{\mathbb{K}}_0 = \widehat{\mathbb{K}} \cup \{S\}$, where S is the usual scaling vector field, defined in (17). Similar to Lemma 2.3, we have:

Lemma 4.1. *For any $\widehat{Z} \in \widehat{\mathbb{K}}_0$,*

$$[\boldsymbol{T}_\phi, \widehat{Z}] = c_Z \boldsymbol{T}_0 + \big[-c_Z \boldsymbol{T}_0(\phi) + \boldsymbol{T}_0(Z(\phi))\big] v^i \partial_{v^i} = c_Z \boldsymbol{T}_\phi + \boldsymbol{T}_0(Z(\phi)) v^i \partial_{v^i},$$

where $c_Z = 0$ if $\widehat{Z} \in \widehat{\mathbb{K}}$ and $c_Z = 1$ if $\widehat{Z} = S$, and where Z is the nonlifted field corresponding to $\widehat{Z}$ if $\widehat{Z} \in \widehat{\mathbb{K}}$ and $Z = S$ if $\widehat{Z} = S$.

Proof. Note first that for all $\widehat{Z} \in \widehat{\mathbb{K}}_0$, we have $[\widehat{Z}, v^i \partial_{v^i}] = 0$. We then compute

$$\begin{aligned}
[\boldsymbol{T}_\phi, \widehat{Z}] &= [\boldsymbol{T}_0, \widehat{Z}] - [\boldsymbol{T}_0(\phi) v^i \partial_{v^i}, \widehat{Z}] \\
&= [\boldsymbol{T}_0, \widehat{Z}] + \widehat{Z}(\boldsymbol{T}_0(\phi)) v^i \partial_{v^i} + \boldsymbol{T}_0(\phi)[\widehat{Z}, v^i \partial_{v^i}] \\
&= [\boldsymbol{T}_0, \widehat{Z}] + \big([\widehat{Z}, \boldsymbol{T}_0]\phi + \boldsymbol{T}_0(\widehat{Z}\phi)\big) v^i \partial_{v^i}.
\end{aligned}$$

To conclude the proof, replace all the instances of $[\boldsymbol{T}_0, \widehat{Z}]$ by $c_Z \boldsymbol{T}_0$ according to Lemma 2.3. □

Iterating the above, one obtains:

Lemma 4.2. *Let f be a solution to (51). For any multi-index α, we have the commutator estimate*

$$\big|[\boldsymbol{T}_\phi, \widehat{Z}^\alpha] f\big| \le C \sum_{\substack{|\beta| \le |\alpha|, |\gamma| \le |\alpha| \\ |\beta| + |\gamma| \le |\alpha| + 1}} |\boldsymbol{T}_0(Z^\gamma \phi)| \cdot |\widehat{Z}^\beta f|, \tag{54}$$

where $Z^\gamma \in \mathbb{K}_0^{|\gamma|}$ and $\widehat{Z}^\beta \in \widehat{\mathbb{K}}_0^{|\beta|}$ and $C > 0$ is some constant depending only on $|\alpha|$.

4B2. *Approximate conservation law.* Similar to Lemma 2.17, we have:

Lemma 4.3. *Let h be a regular distribution function for the massless case in the sense of Section 2E. Let g be a regular solution to $\boldsymbol{T}_\phi(g) = v^0 h$, with $v^0 = |v|$, defined on $[0, T] \times \mathbb{R}^n_x \times (\mathbb{R}^n_v \setminus \{0\})$ for some $T > 0$. Then, for all $t \in [0, T]$,*

$$\begin{aligned}
\int_{\Sigma_t} \rho_0(|g|)(t, x)\, dx \le \int_{\Sigma_0} \rho_0(|g|)(0, x)\, dx + \int_0^t \int_{\Sigma_s} \rho_0(|h|)(s, x)\, dx\, ds& \\
+ (n+1) \int_0^t \int_{\Sigma_s} \int_{v \in \mathbb{R}^n_v \setminus \{0\}} |\boldsymbol{T}_0(\phi) f|\, dv\, dx\, ds.&
\end{aligned} \tag{55}$$

Proof. As for the proof of Lemma 2.17, this follows, after regularization of the absolute value, from integration by parts or an application of Stokes' theorem. The term $\boldsymbol{T}_0(\phi)v^i\,\partial_{v^i}|f|$, which appears in the computation, gives rise, after integration by parts in v, to the last term in (55) since $\partial_{v^i}(v^i\boldsymbol{T}_0(\phi)) = (n+1)\boldsymbol{T}_0(\phi)$. □

4B3. *Massless case in dimension $n \geq 4$.* In this section, we prove Theorem 3, found on page 1545. The $n = 3$ case requires slightly more refined techniques (which of course would work also for $n \geq 4$), but the estimates in the $n \geq 4$ case are slightly stronger and simpler; we therefore provide an independent proof.

If ϕ is a solution to the wave equation, let us consider the energy at time $t = 0$,

$$\mathcal{E}_N[\phi](t=0) \equiv \sum_{\substack{|\alpha|\leq N\\ |\alpha|\in\mathbb{K}^{|\alpha|}}} \left\|Z^\alpha(\partial\phi)(t=0)\right\|^2_{L^2(\mathbb{R}^n_x)}. \tag{56}$$

Now if $\phi(t=0) = \phi_0$ and $\partial_t\phi(t=0) = \phi_1$, for pairs of sufficiently regular functions (ϕ_0, ϕ_1) defined on $\mathbb{R}^n_x$, then the above quantity can be computed purely in terms of ϕ_0, ϕ_1, so we define[22]

$$\mathcal{E}_N[\phi_0, \phi_1] \equiv \mathcal{E}_N[\phi](t=0). \tag{57}$$

Similarly, if f is a solution to (51) arising from initial data f_0 at $t = 0$, then we define

$$E_N[f](t=0) \equiv \|f\|_{\mathbb{K},N}(t=0)\left(= \sum_{\substack{|\alpha|\leq N\\ \widehat{Z}^\alpha\in\widehat{\mathbb{K}}^\alpha}} \left\|\rho_0\big(\widehat{Z}^\alpha(f)(t=0)\big)\right\|_{L^1(\mathbb{R}^n_x)}\right), \tag{58}$$

and we remark that this quantity can be computed purely in terms of f_0, so we will set

$$E_N[f_0] \equiv E_N[f](t=0).$$

We will prove:

Theorem 10. *Let $n \geq 4$ and let $N \geq \frac{3}{2}n + 1$. Let (ϕ_0, ϕ_1, f_0) be an initial data set for the massless Vlasov–Nordström system such that $\mathcal{E}_N[\phi_0, \phi_1] + E_N[f_0] < +\infty$. Then, the unique solution (f, ϕ) to (50)–(51) satisfying the initial conditions (52)–(53) also satisfies the following estimates*:

(1) *Global bounds. For all $t \geq 0$,*

$$E_N[f](t) \leq e^{C\mathcal{E}_N^{1/2}[\phi_0,\phi_1]}E_N[f_0],$$

where $C > 0$ is a constant depending only on N, n.

(2) *Pointwise estimates for velocity averages. For all $(t, x) \in [0, +\infty) \times \mathbb{R}^n_x$ and all multi-indices α satisfying $|\alpha| \leq N - n$,*

$$\rho_0(|\widehat{Z}^\alpha f|)(t,x) \lesssim \frac{e^{C\mathcal{E}_N^{1/2}[\phi_0,\phi_1]}E_N[f_0]}{\big(1+\big|t-|x|\big|\big)\big(1+\big|t+|x|\big|\big)^{n-1}}.$$

[22] The alternative to the approach we use here is to assume that (ϕ_0, ϕ_1) are regular initial data with decay fast enough in x, for instance by assuming compact support, so that the resulting $\mathcal{E}_N[\phi(t=0)]$ is finite. What we want to emphasize here is that the quantity $\mathcal{E}_N[\phi(t=0)]$ can in fact be computed purely in terms of the initial data (using the equation to rewrite second and higher time derivatives of ϕ in terms of spatial derivatives), and that this is all that is needed in terms of decay in x.

Proof. Let N, n, ϕ_0, ϕ_1, f_0 be as in the statement of the theorem. From the conservation of energy and the commutation properties of the Z^α with the wave operator, we have, for all t,

$$\mathcal{E}_N[\phi](t) = \mathcal{E}_N[\phi_0, \phi_1] \equiv \mathcal{E}_N.$$

Applying the standard decay estimates obtained via the vector field method to ϕ, we have for all multi-indices α satisfying $|\alpha| \le N - \frac{1}{2}(n+2)$ and for all $(t,x) \in \mathbb{R}_t \times \mathbb{R}^n_x$,

$$|\partial Z^\alpha \phi(t,x)|^2 \lesssim \frac{\mathcal{E}_N[\phi](t)}{\big(1+\big|t-|x|\big|\big)\big(1+\big|t+|x|\big|\big)^{n-1}}. \tag{59}$$

It follows from a standard existence theory for regular data that for all t, we have $E_N[f(t)] < +\infty$.

Applying the Klainerman–Sobolev inequality (39), we obtain, for all multi-indices α satisfying $|\alpha| \le N-n$ and for all $(t,x) \in \mathbb{R}_t \times \mathbb{R}^n_x$,

$$\big|\rho_0(\widehat{Z}^\alpha(f))(t,x)\big| \lesssim \frac{E_N[f](t)}{\big(1+\big|t-|x|\big|\big)\big(1+\big|t+|x|\big|\big)^{n-1}}.$$

From Lemma 4.3 and the commutator estimate (54), we have for all $t \ge 0$ and all multi-indices α,

$$\int_{\Sigma_t} \rho_0(|\widehat{Z}^\alpha f|)(t,x)\,dx \le \int_{\Sigma_0} \rho_0(|\widehat{Z}^\alpha f|)(0,x)\,dx + \int_0^t \int_{\Sigma_s} \rho_0(|h^\alpha|)(s,x)\,dx\,ds, \tag{60}$$

where[23]

$$|h^\alpha| \lesssim \frac{1}{v^0} \sum_{\substack{|\beta|\le|\alpha|,|\gamma|\le|\alpha| \\ |\beta|+|\gamma|\le|\alpha|+1}} |\boldsymbol{T}_0(Z^\gamma\phi)| \cdot |\widehat{Z}^\beta f| \lesssim \sum_{\substack{|\beta|\le|\alpha|,|\gamma|\le|\alpha| \\ |\beta|+|\gamma|\le|\alpha|+1}} |\partial(Z^\gamma\phi)| \cdot |\widehat{Z}^\beta f|,$$

so that

$$\rho_0(|h^\alpha|) \lesssim \sum_{\substack{|\beta|\le|\alpha|,|\gamma|\le|\alpha| \\ |\beta|+|\gamma|\le|\alpha|+1}} |\partial(Z^\gamma\phi)| \rho_0(|\widehat{Z}^\beta f|),$$

since ϕ is independent of v. Integrating over x, we obtain, for all $s \in [0,t]$,

$$\int_{\Sigma_s} \rho_0(|h^\alpha|)(s,x)\,dx \lesssim \sum_{\substack{|\beta|\le|\alpha|,|\gamma|\le|\alpha| \\ |\beta|+|\gamma|\le|\alpha|+1}} \int_{\Sigma_s} |\partial(Z^\gamma\phi)| \rho_0(|\widehat{Z}^\beta f|)(s,x)\,dx.$$

We now estimate each term in the above sum depending on the values of $|\gamma|$ and $|\beta|$.

If $|\beta| \le N-n$, we then apply the pointwise estimates on $\rho_0(\widehat{Z}^\beta(f))$ to obtain

$$\int_{\Sigma_s} |\partial(Z^\gamma\phi)| \rho_0(|\widehat{Z}^\beta f|)(s,x)\,dx \lesssim \int_{\Sigma_s} |\partial(Z^\gamma\phi)| \frac{E_N[f](s)}{\big(1+\big|s-|x|\big|\big)\big(1+\big|s+|x|\big|\big)^{n-1}}\,dx.$$

[23]Note that the last term in the right-hand side of Lemma 4.3 is similar to the error terms arising from the commutator estimate of Lemma 4.2 and is therefore accounted for in the h^α error term in equation (60).

Applying the Cauchy–Schwarz inequality and using that[24]

$$\left\| \frac{1}{\left(1+\big|s-|x|\big|\right)\left(1+\big|s+|x|\big|\right)^{n-1}} \right\|_{L^2_x} \lesssim \frac{1}{(1+s)^{\frac{n-1}{2}}}, \tag{61}$$

we obtain

$$\int_{\Sigma_s} |\partial(Z^\gamma \phi)| \rho_0(|\widehat{Z}^\beta f|)(s,x)\, dx \lesssim \mathcal{E}_N^{\frac{1}{2}}[\phi](s) \frac{E_N[f](s)}{(1+s)^{\frac{n-1}{2}}}. \tag{62}$$

If now $|\beta| > N-n$, then $|\gamma| \le |\alpha| + 1 - |\beta| \le N - \frac{1}{2}(n+2)$ and thus, we also have (62), using this time the pointwise estimates on $\partial(Z^\gamma\phi)$ given by (59). Applying Grönwall's inequality finishes the proof of the theorem. □

4B4. *Massless case in dimension $n = 3$.* We now turn to the case of dimension 3, where the slower pointwise decay of solutions to the wave equations leads to a slightly harder analysis. First, let us strengthen our norms for the Vlasov field.

For this, recall the algebra of weights $\Bbbk_0$ introduced in Section 2H and define a rescaled version κ_0 by

$$\kappa_0 \equiv (v^0)^{-1}\Bbbk_0 = \left\{ \frac{\mathfrak{z}}{v^0} \,\middle|\, \mathfrak{z} \in \Bbbk_0 \right\},$$

where we recall that $v^0 = |v|$ in the massless case. If α is a multi-index, we will write $[\mathfrak{z}/v^0]^\alpha \in \kappa_0^{|\alpha|}$ to denote a product $|\alpha|$ elements of κ_0 and $[|\mathfrak{z}|/v^0]^\alpha$ in the case we take the product of the absolute values of these elements.

Let us now define, for any regular distribution function f, the weighted norm

$$\begin{aligned} E_{N,q}[f] &\equiv \sum_{\substack{|\alpha|\le N \\ |\beta|\le q}} \sum_{\widehat{Z}^\alpha \in \widehat{\mathbb{K}}^{|\alpha|}} \int_{\Sigma_t} \rho_0\left(|\widehat{Z}^\alpha f| \left[\frac{|\mathfrak{z}|}{v^0}\right]^\beta \right)(x)\, dx \\ &\left(= \sum_{\substack{|\alpha|\le N \\ |\beta|\le q}} \sum_{\widehat{Z}^\alpha \in \widehat{\mathbb{K}}^{|\alpha|}} \int_{\Sigma_t} \int_{v\in\mathbb{R}^n\setminus\{0\}} \left(|\widehat{Z}^\alpha f|(x,v) \left[\frac{|\mathfrak{z}|}{v^0}\right]^\beta \right) v^0\, dv\, dx \right), \end{aligned} \tag{63}$$

where the weights $\mathfrak{z}/v^0$ lie in κ_0.

Theorem 11 (asymptotic behaviour in dimension $n = 3$). *Consider the dimension $n = 3$. Let $N \ge 7$ and $q \ge 1$. Let (ϕ_0, ϕ_1, f_0) be an initial data set for the massless Vlasov–Nordström system such that $\mathcal{E}_N[\phi_0, \phi_1] + E_N[f_0]_{N,q} < +\infty$. Then, the unique solution (f, ϕ) to* (50)–(51) *satisfying the initial conditions* (52)–(53) *also satisfies the following estimates*:

(1) *Global bounds with growth for the top-order norms. For all $t \in \mathbb{R}_t$,*

$$E_{N,q}[f](t) \le (1+t)^{C\mathcal{E}_N^{1/2}[\phi_0,\phi_1]} E_{N,q}[f_0], \tag{64}$$

where $C > 0$ is a constant depending only on N, n and q.

[24] For the convenience of the reader, we have added in Appendix B certain integral estimates which include (61).

(2) *Small data improvement for the low-order norms. There exists an* ε_0 *(depending only on* n, N, q*) such that if* $\mathcal{E}_N[\phi_0, \phi_1] \le \varepsilon_0$*, then for all* $t \in \mathbb{R}_t$,

$$E_{N-\frac{n+4}{2},q-1}[f](t) \le e^{C\mathcal{E}_N^{1/2}[\phi_0,\phi_1]} E_{N,q}[f_0]. \tag{65}$$

(3) *Under the above smallness assumption, we also have the optimal pointwise estimates for velocity averages. For all* $(t,x) \in \mathbb{R}_t \times \mathbb{R}^n_x$ *and all multi-indices* α *satisfying* $|\alpha| \le N - \frac{1}{2}(3n+4)$ *and all* $|\beta| \le q-1$,

$$\rho_0\left(\left|\widehat{Z}^\alpha(f)\left[\frac{\mathfrak{z}}{v^0}\right]^\beta\right|\right)(t,x) \lesssim \frac{e^{C\mathcal{E}_N^{1/2}[\phi_0,\phi_1]} E_{N,q}[f_0]}{\big(1+\big|t-|x|\big|\big)\big(1+\big|t+|x|\big|\big)^{n-1}}.$$

Proof. First, let us note that for all $\mathfrak{z} \in \Bbbk_0$, we have

$$v^i \partial_{v^i}\left(\frac{\mathfrak{z}}{v^0} f\right) = \frac{\mathfrak{z}}{v^0}\, v^i \partial_{v^i} f,$$

from which it follows that for all regular distribution functions g, we have $[\boldsymbol{T}_\phi, \mathfrak{z}/v^0]g = 0$. Thus, we can upgrade Lemma 4.2 to

$$\left|\left[\boldsymbol{T}_\phi, \left[\frac{\mathfrak{z}}{v^0}\right]^\sigma \widehat{Z}^\alpha\right] f\right| \le C \sum_{\substack{|\beta|\le|\alpha|,|\gamma|\le|\alpha| \\ |\beta|+|\gamma|\le|\alpha|+1}} |\boldsymbol{T}_0(Z^\gamma\phi)| \cdot \left[\frac{|\mathfrak{z}|}{v^0}\right]^\sigma |\widehat{Z}^\beta f|, \tag{66}$$

where $Z^\gamma \in \mathbb{K}_0^{|\gamma|}$, $\widehat{Z}^\beta \in \widehat{\mathbb{K}}_0^{|\beta|}$, $[\mathfrak{z}/v^0]^\sigma \in \kappa_0^{|\sigma|}$ and $C > 0$ is some constant depending only on $|\alpha|$. Applying arguments similar to those used in the $n \ge 4$ case yields

$$\begin{aligned} E_{N,q}[f](t) &\le E_{N,q}[f_0] + C\int_0^t \sum_{\substack{|\beta|\le|\alpha|,|\gamma|\le|\alpha| \\ |\beta|+|\gamma|\le|\alpha|+1}} \sum_{|\sigma|\le q} \int_{\Sigma_s} |\partial(Z^\gamma\phi)|\rho_0\left(\left[\frac{|\mathfrak{z}|}{v^0}\right]^\sigma |\widehat{Z}^\beta f|\right)(s,x)\,dx\,ds \\ &\le E_{N,q}[f_0] + C\int_0^t \mathcal{E}_N^{\frac{1}{2}} \frac{E_{N,q}[f](s)}{(1+s)}\,ds. \end{aligned} \tag{67}$$

Applying Grönwall's inequality, we obtain (64).

Now assume that $\mathcal{E}_N \le \varepsilon_0$ with ε_0 small enough that

$$E_{N,q}[f](t) \le (1+t)^\delta E_{N,q}[f_0],$$

with $\delta = C\mathcal{E}_N^{\frac{1}{2}} < \frac{1}{2}$.

The key to the improved estimates is the following decomposition of the transport operator $\boldsymbol{T}_0$:

$$\begin{aligned} \boldsymbol{T}_0 = v^0\partial_t + v^i\partial_{x^i} &= v^0\left(\partial_t + \frac{x^i}{|x|}\partial_{x^i}\right) - v^0\frac{x^i}{|x|}\partial_{x^i} + v^i\partial_{x^i} \\ &= v^0\left(\partial_t + \frac{x^i}{|x|}\partial_{x^i}\right) + \frac{v^0 x^i}{t|x|}(|x|-t)\partial_{x^i} + \frac{v^i t - x^i v^0}{t}\partial_{x^i} \end{aligned}$$

$$= v^0 \underbrace{\left(\partial_t + \frac{x^i}{|x|}\partial_{x^i}\right)}_{\text{outgoing derivatives}} - \frac{v^0}{t} \underbrace{\frac{x^i}{|x|}}_{\text{bounded}} \underbrace{\left(\frac{-x^j\Omega_{ij} + t\Omega_{0i} - x_i S}{t+r}\right)}_{\leq C(|\Omega_{ij}|+|\Omega_{0i}|+|S|)} + v^0 \frac{\mathfrak{z}}{v^0 t}\partial_{x^i},$$

where the weight $\mathfrak{z}$ in the last term is $v^i t - x^i v^0 \in \Bbbk_0$. Recall[25] now the following improved decay for outgoing derivatives of solutions to the wave equations: for all multi-indices α such that $|\alpha| \leq N - \frac{1}{2}(n+2) - 1$,

$$\left|\left(\partial_t + \frac{x^i}{r}\partial_{x^i}\right) Z^\alpha(\phi)\right| \lesssim \frac{\mathcal{E}_N}{(1+t)^{\frac{3}{2}}}.$$

To estimate the second term, we need to obtain decay for $Z\phi$ as solution to the wave equation. This is done by integrating the decay of $\partial Z\phi$ coming from the Klainerman–Sobolev inequality along ingoing null rays. We do not perform the proof of this fact here, but the reader can refer to the proof of Lemma 4.14, where a similar result is proven (for a Cauchy problem with initial data on a hyperboloid). One then obtains:

$$\left|\left(\frac{-x^j\Omega_{ij} + t\Omega_{0i} - x_i S}{t+r}\right) Z^\alpha \phi\right| \lesssim \frac{\mathcal{E}_N}{(1+t)^{\frac{1}{2}}}.$$

As a consequence, it follows that for all multi-indices $|\alpha| \leq N - \frac{1}{2}(n+2) - 1$,

$$|\boldsymbol{T}_0(Z^\alpha \phi)| \lesssim \mathcal{E}_N v^0 \left(\frac{1}{(1+t)^{\frac{3}{2}}} + \sum_{\mathfrak{z}\in\Bbbk_0} \frac{|\mathfrak{z}|}{v^0} \frac{1}{t(1+t)}\right).$$

Repeating the previous ingredients then gives (65). The pointwise estimates then follow from the Klainerman–Sobolev inequality (42). □

4C. *The massive Vlasov–Nordström system.* We now turn to the massive case, that is to say the system

$$\Box\phi = m^2 \int_v f \frac{dv}{v^0} = m^2 \rho_1\left(\frac{f}{(v^0)^2}\right)$$

$$\boldsymbol{T}_m(f) - \left(\boldsymbol{T}_m(\phi)v^i + m^2\nabla^i\phi\right)\frac{\partial f}{\partial v^i} = (n+1)\boldsymbol{T}_m(\phi) f.$$

As in the massless case, we introduce the notation $\boldsymbol{T}_\phi \equiv \boldsymbol{T}_m - (\boldsymbol{T}_m(\phi)v^i + m^2\nabla^i\phi)(\partial/\partial v^i)$ for the transport operator that appears on the left-hand side of the last equation. With this notation, we will seek solutions of the massive Vlasov–Nordström system

$$\Box\phi = m^2 \int_v f \frac{dv}{v^0}, \tag{68}$$

$$\boldsymbol{T}_\phi(f) = (n+1) f\, \boldsymbol{T}_m(\phi) \tag{69}$$

completed by the initial conditions

$$\phi_{|H_1} = \phi_0, \quad \partial_t\phi_{|H_1} = \phi_1, \tag{70}$$

$$f_{|H_1\times\mathbb{R}^n_v} = f_0. \tag{71}$$

[25]This can be obtained from the usual Klainerman–Sobolev inequality and the formula for ∂_s in (24) by integration along the constant $t = |x|$ null lines. See, for instance, [Wang 2015b] for details.

As for the massless case, the lower the dimension, the harder it is to close the estimates. We consider here only the dimensions $n \geq 4$. As already explained, to treat the case $n = 3$, we need a refinement of our method, for instance, the use of modified vector fields in the spirit of [Smulevici 2016], and we postpone this to future work. The proof that we shall give below will be enough to close the estimates for $n = 4$ with some ε-growth in the norms, and without any growth if $n > 4$.

In the following, we will set the mass m equal to 1.

4D. *The norms.* In the context of the massive Vlasov–Nordström system, we define the following energies, similar to the energies defined in (56) and (58):

- For the field ϕ, satisfying a wave equation,

$$\mathcal{E}_N[\phi](\rho) \equiv \sum_{\substack{|\alpha| \leq N \\ Z^\alpha \in \mathbb{P}^{|\alpha|}}} \int_{H_\rho} T[Z^\alpha \phi](\partial_t, \nu_\rho)\, d\mu_{H_\rho}, \tag{72}$$

  where, for any scalar function ψ we denote by $T[\psi] = d\psi \otimes d\psi - \frac{1}{2}\eta(\nabla\psi, \nabla\psi)\eta$ its energy-momentum tensor.

- For the field f, satisfying a transport equation,

$$E_N[f](\rho) \equiv \sum_{\substack{|\alpha| \leq \lfloor N/2 \rfloor \\ Z^\alpha \in \widehat{\mathbb{P}}^\alpha}} \big\| \chi_1\big((v^0)^2 |\widehat{Z}^\alpha(f)|\big) \big\|_{L^1(H_\rho)} + \sum_{\substack{\lfloor N/2 \rfloor + 1 \leq |\alpha| \leq N \\ Z^\alpha \in \widehat{\mathbb{P}}^\alpha}} \big\| \chi_1(|\widehat{Z}^\alpha(f)|) \big\|_{L^1(H_\rho)},$$

  where for any regular distribution function g, the energy density $\chi_1(g)$ is defined as in Section 2K.

Remark 4.4. The weight on the lower-order derivatives contained in the norm of f ensures that pointwise estimates can be performed on terms of the form

$$\int_v \big| v^0 \widehat{Z}^\alpha f(t,x,v) \big|\, dv \lesssim \frac{E_N[f](\rho)}{t^n},$$

according to Theorem 8, given on page 1564, provided that $|\alpha| \leq \lfloor \frac{N}{2} \rfloor - n$. It should furthermore be noticed that the "unweighted" standard estimates coming from Theorem 8 are still true for $|\alpha| \leq N - n$:

$$\int_v \big| \widehat{Z}^\alpha f(t,x,v) \big| \frac{dv}{v^0} \lesssim \frac{E_N[f](\rho)}{t^n}.$$

They will nonetheless not be used in the following.

4D1. *The main result.* Our main result for the massive Vlasov–Nordström system is contained in the following theorem.

Theorem 12. *Let $n \geq 4$ and $N \geq 3n + 4$. Let (f_0, ϕ_0, ϕ_1) be an initial data set for the system (68)–(71). Then, there exists an $\varepsilon_0 > 0$ such that, for all $0 \leq \varepsilon < \varepsilon_0$, if*

- $\mathcal{E}_N[\phi_0, \phi_1] \leq \varepsilon$ *(initial regularity of ϕ),*
- $E_{N+n}[f_0] \leq \varepsilon$ *(initial regularity of f),*

then, the unique classical solution (f,ϕ) of (68)–(71) exists in the whole of the future unit hyperboloid and satisfies the following estimates:

(1) *Energy bounds for ϕ. For all $\rho \geq 1$,*

$$\mathcal{E}_N[\phi](\rho) \leq 2\varepsilon.$$

(2) *Global bounds for f at order less than N. For all $\rho \geq 1$,*

$$E_N[f](\rho) \leq \rho^{C\varepsilon^{1/4}} 2\varepsilon,$$

where $C = 1$ when $n = 4$, and $C = 0$ when $n > 4$.

(3) *Pointwise decay for $\partial Z^\alpha \phi$. For all multi-indices $|\alpha|$ such that $|\alpha| \leq N - \frac{n+2}{2}$ and all (t,x) with $t \geq \sqrt{1+|x|^2}$, we have*

$$|\partial Z^\alpha \phi| \lesssim \frac{\varepsilon}{(1+t)^{\frac{n-1}{2}}(1+(t-|x|))^{\frac{1}{2}}}.$$

(4) *Pointwise decay for $\rho(|\partial Z^\alpha f|)$. For all multi-indices α and β such that $|\alpha| \leq N-n$ and $|\beta| \leq \lfloor \frac{N}{2} \rfloor - n$ and all (t,x) with $t \geq \sqrt{1+|x|^2}$, we have*

$$\int_v |\widehat{Z}^\alpha f| \frac{dv}{v^0} \lesssim \frac{\varepsilon}{(1+t)^{n-C\varepsilon^{1/4}}},$$

$$\int_v v^0 |\widehat{Z}^\beta f|\, dv \lesssim \frac{\varepsilon}{(1+t)^{n-C\varepsilon^{1/4}}},$$

where $C = 1$ when $n = 4$ and $C = 0$ when $n > 4$.

(5) *Finally, the following L^2 estimates on f hold. For all multi-indices α with $\lfloor \frac{N}{2} \rfloor - n + 1 \leq |\alpha| \leq N$, and all (t,x) with $t \geq \sqrt{1+|x|^2}$, we have*

$$\int_{H_\rho} \frac{t}{\rho} \left(\int_v |\widehat{Z}^\alpha f| \frac{dv}{v^0} \right)^2 d\mu_{H_\rho} \lesssim \varepsilon^2 \rho^{C\varepsilon^{1/4}-n},$$

where $C = 2$ when $n = 4$ and $C = 0$ when $n > 4$.

4E. *Proof of Theorem 12.*

4E1. *Structure of the proof and the bootstrap assumptions.* From now on, we consider a solution (f,ϕ) to (68)–(71) arising from initial data satisfying the requirements of Theorem 12. Let P be the largest (hyperboloidal) time so that the following bootstrap assumptions hold on $[1, P]$: assume that there exist an ε small enough and $\delta \in \left[0, \frac{1}{2}\right)$ such that, for all (ρ, r, ω) in $[1, P] \times \mathbb{R}^3$, we have

- energy bounds for ϕ,
$$\mathcal{E}_N[\phi](\rho) \leq 2\varepsilon; \tag{73}$$
- global bounds for f,
$$E_N[f](\rho) \leq \rho^\delta 2\varepsilon. \tag{74}$$

It follows from a continuity argument[26] that $P > 1$ and the remainder of the proof will be devoted to the improvement of each of the above inequalities, establishing the validity of Theorem 12. The proof is organized as follows:

- We first prove the necessary commutation formulae with the transport operator $\boldsymbol{T}_\phi$ in Section 4E2. The fundamental commutator is given in Lemma 4.11.

- A second step consists in rewriting the well-known standard Klainerman–Sobolev estimates for scalar fields using the hyperboloidal foliation (Proposition 4.12), in Section 4E3. These decay estimates for derivatives of scalar fields also provide estimates on the fields themselves after integration along null lines (Lemma 4.14).

- In Section 4E5, the bootstrap assumption (73) is improved, assuming weighted L^2_x decay estimates for the higher-order derivatives of the solution to the transport equation (see Lemma 4.18). The proof is based on energy estimates for which we need the source terms to have sufficient decay. When only low derivatives are involved, our Klainerman–Sobolev inequalities for f are sufficient to close the energy estimates for ϕ, so that the L^2_x decay estimates are only required to handle the high derivatives case (see Lemma 4.18).

- In Section 4E6, the bootstrap assumption (74) is improved. The proof relies on the conservation law for the massive transport equation (Lemma 4.20). Unfortunately, some of the source terms arising from the commutation relations are a priori not space-time integrable. To handle this lack of decay, we use *renormalized* variables by incorporating part of the source term in the original variables; see equation (98). Here we use pointwise estimates for $\partial Z^\alpha \phi$ but also the pointwise estimates on $Z^\alpha \phi$ provided by Lemma 4.14. The improvement of the bootstrap assumption is obtained after returning back to the original variables, provided that the initial data are small enough (Proposition 4.24).

- One finally proves in Section 4E7 the L^2-estimates for the transport equation, which are required in Section 4E5 to improve the bootstrap assumption on the solution of the wave equation. To this end, the equations for the renormalized variables introduced in equation (98) in Section 4E6 are rewritten as a system (Lemma 4.28) of inhomogeneous transport equations. Using the fact that we have control on the initial data for $N + n$ derivatives, it is possible to prove strong pointwise estimates for the homogeneous part of the solution to this system carrying the initial data (Lemma 4.29). The inhomogeneous part of the solution to this system (with no initial data) can be decomposed into an L^2 integrable function and a pointwise decaying function; see equations (114) and (115). This decomposition can then be exploited to prove the decay of weighted L^2-norms of higher-order derivatives of f (see Proposition 4.31). The later decay estimate is then used to improve the bootstrap assumption for the wave equation (see Lemma 4.18).

Note that:

- An estimate for the size of δ in (74) is obtained in Section 4E7 (Lemma 4.30).

- Finally, the maximal regularity is required in Lemma 4.22, when pointwise estimates have to be performed on f.

[26] Note that the methods of this paper show in particular that the system is well-posed in the spaces corresponding to the norms $\mathcal{E}_N^{1/2}[\phi]$ and $E_N[f]$ for N sufficiently large. See also [Calogero and Rein 2004] for another local existence statement.

In the sequel, we will heavily use the following pointwise estimates, which hold under the bootstrap assumptions (73) and (74):

- As a consequence of Proposition 4.12, if $|\gamma| \leq N - \lfloor \frac{n}{2} \rfloor - 1$, then

$$|\partial Z^\gamma \phi| \lesssim \frac{\sqrt{\varepsilon}}{(t-|x|)^{\frac{1}{2}}(1+t)^{\frac{n-1}{2}}} = \frac{\sqrt{\varepsilon}}{\rho(1+t)^{\frac{n}{2}-1}}.$$

- As a consequence of Lemma 4.14, if $|\gamma| \leq N - \lfloor \frac{n}{2} \rfloor - 1$, then

$$|Z^\gamma \phi| \lesssim \frac{\sqrt{\varepsilon}(t-|x|)^{\frac{1}{2}}}{(1+t)^{\frac{n-1}{2}}} = \frac{\sqrt{\varepsilon}\rho}{(1+t)^{\frac{n}{2}}}.$$

- As a consequence of Theorem 8, given on page 1564, if $|\beta| \leq N-n$, then

$$\int_v |\widehat{Z}^\beta f| \frac{dv}{v^0} \lesssim \frac{\varepsilon \rho^\delta}{(1+t)^n}.$$

- Finally, as a consequence of Theorem 8, if $|\beta| \leq \lfloor \frac{N}{2} \rfloor - n$, then

$$\int_v |v^0 \widehat{Z}^\beta f| \, dv \lesssim \frac{\varepsilon \rho^\delta}{(1+t)^n}.$$

4E2. *Commutators in the massive case.* Let us start with the following commutation relations.

Lemma 4.5.

$$[\partial_t, \partial_{v^i}] = 0, \tag{75}$$

$$[\partial_{x^i}, \partial_{v^i}] = 0, \tag{76}$$

$$[t\partial_{x^j} + x^j \partial_t + v^0 \partial_{v^j}, \partial_{v^i}] = -\frac{v^i}{v^0}\partial_{v^j}, \tag{77}$$

$$[x^i \partial_j - x^j \partial_i + v^i \partial_{v^j} - v^j \partial_{v^i}, \partial_{v^k}] = -\delta^i_k \partial_{v^j} + \delta^j_k \partial_{v^i}, \tag{78}$$

$$[t\partial_{x^j} + x^j \partial_t + v^0 \partial_{v^j}, v^i \partial_{v^i}] = \frac{1}{v^0}\partial_{v^j}, \tag{79}$$

$$[x^i \partial_j - x^j \partial_i + v^i \partial_{v^j} - v^j \partial_{v^i}, v^k \partial_{v^k}] = 0. \tag{80}$$

We now evaluate the commutators $[\boldsymbol{T}_\phi, \widehat{Z}]$ for $\widehat{Z} \in \widehat{\mathbb{P}}$. We have

$$\begin{aligned}
[\boldsymbol{T}_\phi, \widehat{Z}] f &= \underbrace{[\boldsymbol{T}_1, \widehat{Z}] f}_{=0 \text{ if } \widehat{Z} \in \widehat{\mathbb{P}}} - [\boldsymbol{T}_1(\phi) v^i \partial_{v^i}, \widehat{Z}] f - [\nabla^i \phi \cdot \partial_{v^i}, \widehat{Z}] f \\
&= \widehat{Z}[\boldsymbol{T}_1(\phi)] v^i \partial_{v^i} f + \boldsymbol{T}_1(\phi)[\widehat{Z}, v^i \partial_{v^i}] f - [\nabla^i \phi \cdot \partial_{v^i}, \widehat{Z}] f \\
&= \big([\widehat{Z}, \boldsymbol{T}_1]\phi + \boldsymbol{T}_1(\widehat{Z}\phi)\big) v^i \partial_{v^i} f - [\nabla^i \phi \partial_{v^i}, \widehat{Z}] f + \begin{cases} \boldsymbol{T}_1(\phi) \frac{1}{v^0} \partial_{v^j} f & \text{if } \widehat{Z} = t\partial_{x^j} + x^j \partial_t + v^0 \partial_{v^j}, \\ 0 & \text{otherwise} \end{cases} \\
&= \boldsymbol{T}_1(Z\phi) v^i \partial_{v^i} f - [\nabla^i \phi \cdot \partial_{v^i}, \widehat{Z}] f + \begin{cases} \boldsymbol{T}_1(\phi) \frac{1}{v^0} \partial_{v^j} f & \text{if } \widehat{Z} = t\partial_{x^j} + x^j \partial_t + v^0 \partial_{v^j}, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned} \tag{81}$$

We have used that $\widehat{Z}(\phi) = Z(\phi)$ since ϕ is independent of v. To estimate the second term on the right-hand side of the last equation, we need:

Lemma 4.6. *For any $Z \in \mathbb{P}$:*

- *If Z is a translation, then*
$$[\nabla^i \phi \cdot \partial_{v^i}, \widehat{Z}] = -\nabla^i (Z\phi)\partial_{v^i}.$$
- *If $Z = \Omega_{jk}$ is a rotation such that $\widehat{Z} = \Omega_{jk} + v^j \partial_{v^k} - v^k \partial_{v^j}$, then*
$$[\nabla^i \phi \cdot \partial_{v^i}, \widehat{Z}] = -\nabla^i (Z\phi)\partial_{v^i}.$$
- *If $Z = \Omega_{0j}$ is a Lorentz boost such that $\widehat{Z} = \Omega_{0j} + v^0 \partial_{v^j}$, then*
$$[\nabla^i \phi \cdot \partial_{v^i}, \widehat{Z}] = -\nabla^i (Z\phi)\partial_{v^i} + \nabla_i \phi \frac{v^i}{v^0} \partial_{v^j} + \partial_t(\phi)\, \partial_{v^j} f.$$

We now summarize these computations:

Lemma 4.7. *Let $\widehat{Z} \in \widehat{\mathbb{P}}$. Then*
$$[\boldsymbol{T}_\phi, \widehat{Z}] f = \boldsymbol{T}_1(Z\phi) v^i\, \partial_{v^i} f + \sum_{\substack{|\alpha| \le 1 \\ 1 \le j \le n \\ 0 \le \beta \le n}} p^{j\beta}\left(\frac{v}{v^0}\right) \partial_{x^\beta} Z^\alpha(\phi) \cdot \partial_{v^j} f,$$
where the $p^{i\beta}(v/v^0)$ are polynomial of degree at most 1 in the variables v^k/v^0, $1 \le k \le n$.

The terms containing derivatives of v in the above formulae are problematic, since the ∂_v are not part of the algebra $\widehat{\mathbb{P}}$. We use the following decomposition for all $1 \le i \le n$:
$$\partial_{v^i} = \frac{1}{v^0}(t\partial_{x^i} + x^i \partial_t + v^0 \partial_{v^i}) - \frac{1}{v^0}(t\partial_{x^i} + x^i \partial_t) = \frac{1}{v^0}\widehat{\Omega}_{0i} - \frac{1}{v^0}(t\partial_{x^i} + x^i \partial_t). \tag{82}$$

Remark 4.8. Note that
$$\frac{1}{v^0}|t\, \partial_{x^i} f + x^i\, \partial_t f| \le \frac{t}{v^0}\big(|\partial_t f| + |\partial_{x^i} f|\big) \tag{83}$$
for (t,x) in the future of the unit hyperboloid. Now ∂_t and ∂_{x^i} belong to $\widehat{\mathbb{P}}$, but the price to pay is the extra t-factor. It is precisely this extra t-growth which forbids us to close the estimate in dimension 3. A similar obstacle was identified for the Vlasov–Poisson system in dimension 3 and solved by means of *modified vector fields* in [Smulevici 2016]. We hope to treat the 3-dimensional massive Vlasov–Nordström system in future work.

This leads to the commutation formula:

Lemma 4.9. *Let $\widehat{Z} \in \widehat{\mathbb{P}}$. Then*
$$[\boldsymbol{T}_\phi, \widehat{Z}] f = \boldsymbol{T}_1(Z\phi) \sum_{|\alpha|=1} q_\alpha\left(\frac{v}{v^0}, t, x\right)\widehat{Z}^\alpha(f) + \sum_{\substack{|\alpha| \le 1, |\beta|=1 \\ 0 \le \gamma \le n}} \frac{p^\gamma_{\alpha\beta}(v/v^0, t, x)}{v^0} \partial_{x^\gamma} Z^\alpha(\phi) \cdot \widehat{Z}^\beta(f),$$
where the $q_\alpha(v/v^0, t, x)$ and $p^\gamma_{\alpha\beta}(v/v^0, t, x)$ are polynomial of degree at most 1 in the variables
$$\frac{v^k}{v^0}, \quad \frac{v^k}{v^0} t, \quad \frac{v^k}{v^0} x^i, \quad 1 \le i, k \le n.$$

Iterating the above formula, we obtain:

Lemma 4.10. *Let α be a multi-index and $\widehat{Z}^\alpha \in \widehat{\mathbb{P}}^{|\alpha|}$. Then*

$$[\boldsymbol{T}_\phi, \widehat{Z}^\alpha] f = \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\ 1\le|\gamma|\le|\alpha|\\ 1\le|\beta|\le|\alpha|}} \boldsymbol{T}_1(Z^\gamma\phi) q_{\beta\gamma}\left(\frac{v}{v^0}, t, x\right)\widehat{Z}^\beta(f) + \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\ 0\le\sigma\le n\\ 1\le|\beta|\le|\alpha|\\ 0\le|\gamma|\le|\alpha|}} \frac{1}{v^0} p^\sigma_{\gamma\beta}\left(\frac{v}{v^0}, t, x\right)\partial_{x^\sigma} Z^\gamma(\phi)\cdot\widehat{Z}^\beta f,$$

where

- *the $q_{\beta\gamma}(v/v^0, t, x)$ are linear combinations of the terms*

$$q\left(\frac{v^k}{v^0}\right), \quad q'\left(\frac{v^k}{v^0}\right)t, \quad q''\left(\frac{v^k}{v^0}\right)x^i, \quad 1\le i,k\le n,$$

  where q, q', q'' are polynomials of degree at most $|\alpha|$,

- *the $p^\gamma_{\gamma\beta}(v/v^0, t, x)$ are linear combinations with constant coefficient of the terms*

$$p\left(\frac{v^k}{v^0}\right), \quad p'\left(\frac{v^k}{v^0}\right)t, \quad p''\left(\frac{v^k}{v^0}\right)x^i, \quad 1\le i,k\le n,$$

  where p, p', p'' are polynomials of degree at most $|\alpha|$.

Proof. This follows by an induction argument on the length of the multi-index α and we therefore only provide some details here. Assume the lemma is true for $|\alpha|$. Recall that, for any $\widehat{Z} \in \widehat{\mathbb{P}}_0$,

$$[\boldsymbol{T}_\phi, \widehat{Z}\widehat{Z}^\alpha](f) = [\boldsymbol{T}_\phi, \widehat{Z}]\widehat{Z}^\alpha(f) + \widehat{Z}[\boldsymbol{T}_\phi, \widehat{Z}^\alpha](f) = I_1 + I_2,$$

with

$$I_1 = [\boldsymbol{T}_\phi, \widehat{Z}]\widehat{Z}^\alpha(f), \quad I_2 = \widehat{Z}[\boldsymbol{T}_\phi, \widehat{Z}^\alpha](f).$$

Using Lemma 4.9, we have for I_1,

$$I_1 = \boldsymbol{T}_1(Z\phi)\sum_{|\gamma|=1} q_\gamma\left(\frac{v}{v^0}, t, x\right)\widehat{Z}^\gamma(\widehat{Z}^\alpha f) + \sum_{\substack{|\gamma|\le1, |\beta|=1\\ 0\le\sigma\le n}} \frac{1}{v^0} p^\sigma_{\gamma\beta}\left(\frac{v}{v^0}, t, x\right)\partial_{x^\sigma} Z^\gamma(\phi)\cdot\widehat{Z}^\beta(\widehat{Z}^\alpha f), \tag{84}$$

with q_γ and $p^\sigma_{\gamma\beta}$ as in the statement of Lemma 4.9. Since all the terms in (84) clearly have the desired form, we turn to I_2. Applying the induction hypothesis, we have

$$\begin{aligned}
I_2 &= \widehat{Z}\Bigg[\sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\ 1\le|\gamma|\le|\alpha|\\ 1\le|\beta|\le|\alpha|}} \boldsymbol{T}_1(Z^\gamma\phi) q_{\beta\gamma}\left(\frac{v}{v^0}, t, x\right)\widehat{Z}^\beta(f) + \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\ 0\le\sigma\le n\\ 1\le|\beta|\le|\alpha|}} \frac{1}{v^0} p^\sigma_{\gamma\beta}\left(\frac{v}{v^0}, t, x\right)\partial_{x^\sigma} Z^\gamma(\phi)\cdot\widehat{Z}^\beta f\Bigg]\\
&= \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\ 1\le|\gamma|\le|\alpha|\\ 1\le|\beta|\le|\alpha|}} \widehat{Z}\left[\boldsymbol{T}_1(Z^\gamma\phi) q_{\beta\gamma}\left(\frac{v}{v^0}, t, x\right)\widehat{Z}^\beta(f)\right] + \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\ 0\le\sigma\le n\\ 1\le|\beta|\le|\alpha|}} \widehat{Z}\left[\frac{1}{v^0} p^\sigma_{\gamma\beta}\left(\frac{v}{v^0}, t, x\right)\partial_{x^\sigma} Z^\gamma(\phi)\cdot\widehat{Z}^\beta f\right]\\
&= J_1 + J_2,
\end{aligned}$$

where

$$J_1 = \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\ 1\le|\gamma|\le|\alpha|\\ 1\le|\beta|\le|\alpha|}} \widehat{Z}\left[\boldsymbol{T}_1(Z^\gamma\phi) q_{\beta\gamma}\left(\frac{v}{v^0}, t, x\right)\widehat{Z}^\beta(f)\right]$$

and

$$J_2 = \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1 \\ 0\le\sigma\le n \\ 1\le|\beta|\le|\alpha|}} \widehat{Z}\left[\frac{1}{v^0} p^{\sigma}_{\gamma\beta}\left(\frac{v}{v^0}, t, x\right)\partial_{x^\sigma} Z^\gamma(\phi)\cdot \widehat{Z}^\beta f\right].$$

To see that J_1 has the correct form, we distribute $\widehat{Z}$, which gives rise to three types of terms. The terms arising when $\widehat{Z}$ hits $\boldsymbol{T}_1(Z^\gamma\phi)$ or $\widehat{Z}^\beta(f)$ are easily seen to have the right form. It remains to look at the case when $\widehat{Z}$ hits $q_{\beta\gamma}$. If $\widehat{Z}$ is a translation, $\widehat{Z}=\partial_{x^\beta}$ and one easily sees that $\widehat{Z}(q_{\beta\gamma})$ has the correct form. If $\widehat{Z}$ is the lift of a rotation or a Lorentz boost, then we write schematically $\widehat{Z} = {}^xZ + {}^vZ$, where ${}^xZ$ is a homogeneous differential operator of order 1 in (t,x) and ${}^vZ$ is a homogeneous differential operator of order 1 in v. It is then easy to check that ${}^xZ$ applied to a polynomial in the variables v^i/v^0 of degree $\le|\alpha|$, possibly multiplied by the variables t, x^i will produce a polynomial, of the same degree $\le|\alpha|$, in the variables v^i/v^0, possibly multiplied by the variables t, x^i. Similarly, ${}^vZ$ applied to a polynomial in the variable v^i/v^0 of degree $\le|\alpha|$ will produce a polynomial in the variables v^i/v^0 of degree $\le|\alpha|+1$. As a consequence, $\widehat{Z}^\alpha(q_{\beta\gamma})$ is a linear combination of polynomials of degree $|\alpha|+1$, possibly multiplied by t, x^i. The term J_2 can be treated similarly. □

The full expression for $\boldsymbol{T}_\phi(\widehat{Z}^\alpha f)$ can now be computed using the transport equation (69) satisfied by f.

Lemma 4.11. *Let Z^α be in $\mathbb{P}^{|\alpha|}$. Then the following equation holds*:

$$\boldsymbol{T}_\phi(\widehat{Z}^\alpha(f)) = \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1 \\ |\gamma|\ge1, |\beta|\ge1}} \boldsymbol{T}_1(Z^\gamma\phi) q_{\beta\gamma}\left(\frac{v^i}{v^0}, t, x\right)\widehat{Z}^\beta(f)$$
$$+ \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1 \\ 0\le\sigma\le n \\ 1\le|\beta|\le|\alpha|}} \frac{1}{v^0} p^\sigma_{\gamma\beta}\left(\frac{v}{v^0}, t, x\right)\partial_\sigma Z^\gamma(\phi)\widehat{Z}^\beta f + \sum_{|\gamma|+|\beta|=|\alpha|} r_{\gamma\beta}\boldsymbol{T}_1(Z^\gamma\phi)\widehat{Z}^\beta(f), \quad (85)$$

where the $q_{\beta\gamma}$ and $p^\sigma_{\gamma\beta}$ are as in Lemma 4.10 and the $r_{\gamma\beta}$ are constants.

Proof. We have

$$\boldsymbol{T}_\phi(\widehat{Z}^\alpha f) = [\boldsymbol{T}_\phi, \widehat{Z}^\alpha]f + \widehat{Z}^\alpha \boldsymbol{T}_\phi f = [\boldsymbol{T}_\phi, \widehat{Z}^\alpha]f + \widehat{Z}^\alpha((n+1)\boldsymbol{T}_1(\phi)f). \quad (86)$$

The lemma thus follows from Lemma 4.10 and the fact that

$$\widehat{Z}^\alpha(\boldsymbol{T}_1(\phi)f) = \sum_{|\gamma|+|\beta|=|\alpha|} r_{\gamma\beta}\boldsymbol{T}_1(Z^\gamma\phi)\widehat{Z}^\beta(f), \quad (87)$$

where the $r_{\gamma\beta}$ are constants. □

4E3. *The H_ρ foliation and the wave equation.* The aim of this section is to provide a Klainerman–Sobolev-type inequality, applicable to solutions of the inhomogeneous wave equation if the inhomogeneities decay sufficiently fast, using only energies on the H_ρ foliation. This question was addressed by Klainerman

[1993] for the Klein–Gordon operator and we show here how a similar proof can also be applied to the wave operator. We thus consider a function ψ and its energy-momentum tensor

$$T[\psi] = d\psi \otimes d\psi - \tfrac{1}{2}(\eta(\nabla\psi, \nabla\psi))\eta. \tag{88}$$

If we want to perform energy estimates on H_ρ, we need to multiply $T[\psi]$ by ∂_t and the normal to H_ρ, ν_ρ, and integrate on H_ρ. Let us thus compute the quantity $T(\partial_t, \nu_\rho)$.

We find

$$T[\psi](\partial_t, \nu_\rho) = \frac{t}{2\rho}(\psi_t^2 + \psi_r^2 + |\nabla\!\!\!/\psi|^2) + \frac{r}{\rho}\psi_t\psi_r. \tag{89}$$

Recall also that the volume form on H_ρ is given by $(\rho/t)r^{n-1}\,dr\,d\sigma$. Since we are looking only at the region $\rho \geq 1$, we have that $t > r$ and $T(\partial_t, \nu_\rho)$ is clearly positive definite, with some degeneration as $r \to t$. More precisely, fix (t, x) in the future of the unit hyperboloid. Assume first that $r = |x| \leq \frac{1}{2}t$; then $(\rho/t)T[\psi](\partial_t, \nu_\rho) \geq |\partial\psi|^2$.

Let (Y^0, Y^i) be the coordinates of (t, x) in the (y^α)-system of coordinates adapted to the H_ρ foliation as introduced in Section 2C. Let $\Phi(y) = \partial\psi(Y^0, Y^j + ty^j)$. Then, a classical Sobolev inequality yields

$$\begin{aligned} |\partial\psi(Y^0, Y^j)|^2 = |\Phi(0)|^2 &\lesssim \sum_{k\leq\frac{n+2}{2}} \int_{|y|\leq\delta} |\partial_y^k\Phi(y)|^2\,dy \\ &\lesssim \sum_{k\leq\frac{n+2}{2}} \int_{|y|\leq\delta} \bigl|Z^k(\partial\psi)(Y^0, Y^j + ty^j)\bigr|^2\,dy, \end{aligned} \tag{90}$$

using that $\partial_{y^i} = (1/t)\Omega_{0i}$ and the fact that $\partial_{y^i}\Phi = t\,\partial_{y^i}\psi$ together with estimates on $t/(t(Y^0, Y^j + ty^j))$ similar to those of Section 3C. Applying the change of coordinates $y^j \to ty^j$ yields

$$|\partial\psi(Y^0, Y^j)|^2 \lesssim \frac{1}{t^n} \sum_{k\leq\frac{n+2}{2}} \int_{|y|\leq t\delta} \bigl|Z^k(\partial\psi)(Y^0, Y^j + y^j)\bigr|^2\,dy. \tag{91}$$

Finally, $|Y^j + y^j| = |x^j + y^j| \lesssim \bigl(\frac{1}{2} + \delta\bigr)t$ so that, if $\delta > 0$, we are still away from the light cone. Thus, the right-hand side of the previous equations can be controlled by the energies of $Z^k(\partial\psi)$ on H_ρ. On the other hand, if $\frac{1}{2}t \leq r < t$, we first remark that

$$\frac{\rho}{t}T[\psi](\partial_t, \nu_\rho) \geq \left(1 - \frac{r}{t}\right)|\partial\psi|^2. \tag{92}$$

Thus, we may repeat the previous arguments, losing the factor $(1 - r/t)$ in the process, as follows:

$$\begin{aligned} |\partial\psi(Y^0, Y^j)|^2 &\lesssim \frac{1}{t^n} \sum_{k\leq\frac{n+2}{2}} \int_{|y|\leq t\delta} \bigl|Z^k(\partial\psi)(Y^0, Y^j + y^j)\bigr|^2\,dy \\ &\lesssim \frac{1}{t^n} \sum_{k\leq\frac{n+2}{2}} \int_{|y|\leq t\delta} \left(1 - \frac{r}{t}\right)^{-1}\left(1 - \frac{r}{t}\right)\bigl|Z^k(\partial\psi)(Y^0, Y^j + y^j)\bigr|^2\,dy. \end{aligned} \tag{93}$$

Since
$$\frac{1}{1-r/t} = \frac{t(t+r)}{\rho^2} \lesssim \frac{t^2}{\rho^2}$$
and since again, we can replace $t(Y^0, Y^j + ty^j)$ by t in all the above computations, we have shown that
$$|\partial\psi(Y^0, Y^j)|^2 \lesssim \frac{1}{t^{n-2}\rho^2}.$$
Since $\rho^2 = (t+r)(t-r)$ and since $t \geq r$ in the future of the unit hyperboloid, this is exactly the decay estimate predicted by the usual Klainerman–Sobolev inequality using the Σ_t foliation. We summarize this in the following proposition.

Proposition 4.12 (Klainerman–Sobolev inequality for the wave equation using the hyperboloidal foliation). *For any sufficiently regular function ψ of (t,x) defined on the future of the unit hyperboloid, let $\mathcal{E}_{\frac{1}{2}(n+2)}[\psi](\rho)$ denote the energy*
$$\mathcal{E}_{\frac{n+2}{2}}[\psi](\rho) = \sum_{|\alpha|\leq\frac{n+2}{2}} \int_{H_\rho} T[Z^\alpha[\psi]](\partial_t, \nu_\rho)\, d\mu_\rho.$$
Then, for all (t,x) in the future of the unit hyperboloid,
$$|\partial\psi|(t,x) \lesssim \frac{1}{t^{\frac{n-1}{2}}(t-|x|)^{\frac{1}{2}}} \mathcal{E}_N^{\frac{1}{2}}[\psi](\rho(t,x)). \tag{94}$$

It is interesting to note that the above proof does not make use of the scaling vector field.

Remark 4.13. We will use in the following the inequality
$$\frac{\rho}{2t} T[\psi](\partial_t, \nu_\rho) = |\nabla\!\!\!/\psi|^2 + \left(\psi_t^2 + \psi_r^2 + \frac{2r}{t}\psi_t\psi_r\right) = |\nabla\!\!\!/\psi|^2 + \left(1 - \frac{r}{t}\right)(\psi_t^2 + \psi_r^2) + \frac{r}{t}(\psi_t + \psi_r)^2;$$
that is to say,
$$|\partial\psi|^2 \lesssim \frac{\rho}{t-r} T[\psi](\partial_t, \nu_\rho) = \frac{t+r}{\rho} T[\psi](\partial_t, \nu_\rho).$$

The inequality (94) provides decay for $\partial\psi$ but not for ψ. By integration along null lines, one can obtain the following decay for ψ.

Lemma 4.14. *Let ψ be such that $\mathcal{E}_{\frac{1}{2}(n+2)}[\psi](\rho)$ is uniformly bounded on $[1, P]$ for some $P > 1$. Assume moreover that $\psi_{|\rho=1}$ vanishes at ∞.*

Then ψ satisfies, for all (t,x) in the future of H_1,
$$|\psi(t,x)| \lesssim \sup_{[1,P]} \left[\mathcal{E}_N^{\frac{1}{2}}[\Psi]\right] \frac{(1+u)^{\frac{1}{2}}}{t^{\frac{n-1}{2}}},$$
where $u = t - |x|$.

Proof. The Klainerman–Sobolev estimates provide
$$|\partial\psi(t,x)| \lesssim \frac{\sup_{[1,P]}\left[\mathcal{E}_N^{\frac{1}{2}}\right]}{\rho t^{\frac{n}{2}-1}}.$$

Let $(t,x)=(t,r,\omega\in\mathbb{S}^n)$ be a point in the future of H_1, and consider the point on the hyperboloid lying at the intersection of the past light cone from (t,x), the hyperboloid H_1, and the direction ω. Writing $u=t-|x|$ and $v=t+|x|$, we have

$$(t_1,x_1)=\Big(t_1=\sqrt{1+r_1^2},r_1=\frac{1}{2}\Big(v-\frac{1}{v}\Big),\omega\Big)\in H_1.$$

Note that

$$\frac{1}{\langle r_1\rangle}=\frac{2}{v+1/v}\lesssim\frac{1}{\langle v\rangle}\quad\text{since } v\geq t\geq 1.$$

Integrating along the direction ω along the past light cone from (t,x) from (t_1,r_1,ω) to (t,r,ω), one obtains

$$\psi(t,x)=\psi(t_1,x_1)+\int_{t_1-r_1}^{t-r}(\partial_u\psi)\,du,$$

so that

$$\begin{aligned}|\psi(t,x)|&\lesssim\frac{\mathcal{E}_N^{\frac12}(\rho=1)}{\langle r_1\rangle^{\frac{n-1}{2}}}+\int_{t_1-r_1}^{t-r}\frac{\sup_{[1,P]}[\mathcal{E}_N^{\frac12}]}{u^{\frac12}v^{\frac{n-1}{2}}}\,du\\&\lesssim\frac{\mathcal{E}_N^{\frac12}(\rho=1)}{\langle v\rangle^{\frac{n-1}{2}}}+\frac{\sup_{[1,P]}[\mathcal{E}_N^{\frac12}]u^{\frac12}}{v^{\frac{n-1}{2}}}\lesssim\sup_{[1,P]}[\mathcal{E}_N^{\frac12}]\frac{(1+u)^{\frac12}}{v^{\frac{n-1}{2}}},\end{aligned}$$

which concludes the proof since $1/v\leq 2/t$. Here we have used that

$$|\psi(t_1,x_1)|\lesssim\frac{\mathcal{E}_N^{\frac12}(\rho=1)}{\langle r_1\rangle^{\frac{n-1}{2}}},$$

which follows from usual weighted Sobolev inequalities on $\mathbb{R}^n$ applied to $\partial\psi$ and the assumption that ψ restricted to $\rho=1$ vanishes at infinity. □

4E4. *Commutation of the wave equation.* The commutation of the wave equation with our set of vector fields is straightforward and leads to:

Lemma 4.15. *For any multi-index α,*

$$\Box Z^\alpha(\phi)=\int_{v\in\mathbb{R}^n_v}\widehat{Z}^\alpha(f)\frac{dv}{v^0}.\tag{95}$$

Proof. First, recall that the vector fields in the algebra $\mathbb{P}$ commute exactly with $\Box$. The lemma then follows from Lemma 2.14 and Remark 2.15 in the case Z^α contains some combinations of Lorentz boosts. □

Remark 4.16. The following inequality will be used later on:

$$\left|\int_{v\in\mathbb{R}^n_v}\widehat{Z}^\alpha(f)\frac{dv}{v^0}\right|\leq\chi_1\big(|\widehat{Z}^\alpha(f)|\big),$$

which is a direct consequence of Remark 2.12.

4E5. *Energy estimates for the wave equation on hyperboloids.* Consider ψ defined in hyperboloidal time for all $\rho \in [1, P]$ and assume that ψ solves $\Box\psi = h$. Recalling the expression for $T[\psi](\partial_t, \nu_\rho)$ given by (89), we have:

Lemma 4.17. *Let $\rho \in [1, P]$. Then,*

$$\int_{H_\rho} T[\psi](\partial_t, \nu_\rho)\, d\mu_{H_\rho} = \int_{H_1} T[\psi](\partial_t, \nu_\rho)\, d\mu_{H_1} + \int_1^\rho \int_{H_{\rho'}} (\partial_t \psi)(\rho')\, h(\rho')\, d\mu_{H_{\rho'}}\, d\rho'.$$

Proof. The proof of this fact is only sketched, since classical. The reader can refer to [Klainerman 1993]. Remember that the divergence of the stress-energy tensor $T[\psi]$ is given by

$$\partial^\alpha T_{\alpha\beta}[\psi] = h\, \partial_\beta \psi$$

when ψ satisfies the equation $\Box\psi = h$. The lemma then follows by integration between the two hyperboloids H_1 and H_ρ and an application of Stokes' theorem. ☐

To close the energy estimates for $Z^\alpha(\phi)$, we need the right-hand side of (95) to decay. Since for $|\alpha| \le N - n$, the required decay follows from our Klainerman–Sobolev inequality (42) as well as the bootstrap assumption (74), we have the following lemma:

Lemma 4.18. *Assume that $\delta < 1$. Assume moreover that for all multi-indices α such that $N - n + 1 \le |\alpha| \le N$, the following L^2_x decay estimate holds*:

$$\int_{H_\rho} \frac{t}{\rho} \left(\int_v |\widehat{Z}^\alpha f| \frac{dv}{v^0} \right)^2 d\mu_{H_\rho} \lesssim \varepsilon^2 \rho^{\delta - n}. \tag{96}$$

Then, the following inequality holds for all $\rho \in [1, P]$:

$$\mathcal{E}_N[\phi](\rho) \le \varepsilon(1 + C\varepsilon^{\frac{1}{2}}),$$

where C is a constant depending solely on the dimension n and the regularity N. In particular, for ε small enough, for all $\rho \in [1, P]$,

$$\mathcal{E}_N[\phi](\rho) \le \tfrac{3}{2}\varepsilon.$$

Remark 4.19.
- The weighted L^2-estimates (96) will in fact be proven in Section 4E7 for the wider range of multi-indices α with $\lfloor \frac{N}{2} \rfloor - n + 1 \le |\alpha| \le N$.
- Note that the L^2-estimates are needed only for $|\alpha| > N - n$: for lower order, the pointwise decay estimates for the velocity averages are sufficient to conclude.

Proof. The proof of this lemma relies on Lemma 4.17. Applying first Lemmata 4.17 and 4.15 to $Z^\alpha(\phi)$ for all multi-indices of length $|\alpha| \le N$, one obtains immediately, for $\rho \le P$,

$$\begin{aligned}
\mathcal{E}_N[\phi](\rho) - \mathcal{E}_N[\phi](1) &\lesssim \sum_{|\alpha| \le N} \sum_{Z^\alpha \in \mathbb{P}^{|\alpha|}} \int_1^\rho \int_{H_{\rho'}} |\partial_t Z^\alpha \phi| \left(\int_v \frac{|\widehat{Z}^\alpha f|}{v^0}\, dv \right) d\mu_{H_{\rho'}}\, d\rho' \\
&\lesssim \sum_{|\alpha| \le N} \sum_{Z^\alpha \in \mathbb{P}^{|\alpha|}} \int_1^\rho \left(\int_{H_{\rho'}} \left| \left(\frac{\rho'}{t} \right)^{\frac{1}{2}} \partial_t Z^\alpha \phi \right| \cdot \left(\frac{t}{\rho'} \right)^{\frac{1}{2}} \int_v \frac{|\widehat{Z}^\alpha f|}{v^0}\, dv \right) d\mu_{H_{\rho'}}\, d\rho' \Bigg|
\end{aligned}$$

$$\lesssim \int_1^\rho \mathcal{E}_N^{\frac{1}{2}}[\phi](\rho')\Biggl(\sum_{|\alpha|\le N}\sum_{Z^\alpha\in\mathbb{P}^{|\alpha|}}\biggl(\int_{H_{\rho'}}\Bigl(\frac{t}{\rho'}\Bigr)\Bigl(\int_v \frac{|\widehat{Z}^\alpha f|}{v^0}\,dv\Bigr)^2 d\mu_{H_{\rho'}}\biggr)^{\frac{1}{2}}\Biggr)\,d\rho'.$$

We now apply, for the low derivatives of f, Theorem 8 on page 1564 (in conjunction with Remark 2.12): for $|\alpha| \le N-n$,

$$\int_{H_{\rho'}}\Bigl(\frac{t}{\rho'}\Bigr)\Bigl(\int_v \frac{|\widehat{Z}^\alpha f|}{v^0}\,dv\Bigr)^2 d\mu_{H_{\rho'}} \lesssim \int_0^\infty \frac{t}{\rho'}\cdot\frac{\varepsilon^2\rho'^{2\delta}}{t^{2n}}r^{n-1}\frac{\rho'}{t}\,dr \lesssim \varepsilon^2\rho'^{2\delta-n}\int_0^\infty \frac{y^{n-1}}{\langle y\rangle^{2n}}\,dy.$$

Since the last integral is convergent, we obtain

$$\biggl[\int_{H_{\rho'}}\Bigl(\frac{t}{\rho'}\Bigr)\Bigl(\int_v \frac{|\widehat{Z}^\alpha f|}{v^0}\,dv\Bigr)^2 d\mu_{H_{\rho'}}\biggr]^{\frac{1}{2}} \lesssim \varepsilon\rho'^{\delta-\frac{n}{2}},$$

which, assuming $\delta < 1$, is integrable in ρ. For the higher derivatives of f, i.e., for $|\alpha| > N-n$, one uses the assumption of the lemma to obtain

$$\sum_{N-n<|\alpha|\le N}\sum_{Z^\alpha\in\mathbb{P}^{|\alpha|}}\biggl(\int_{H_\rho}\frac{t}{\rho}\Bigl(\int_v |\widehat{Z}^\alpha f|\frac{dv}{v^0}\Bigr)^2 d\mu_{H_\rho}\biggr)^{\frac{1}{2}} \lesssim \varepsilon\rho^{\frac{\delta}{2}-\frac{n}{2}}.$$

We obtain finally

$$\mathcal{E}_N[\phi](\rho) \le \varepsilon + C\varepsilon\Bigl(\int_1^\rho (\rho'^{\frac{\delta-n}{2}})\mathcal{E}_N^{\frac{1}{2}}[\phi](\rho')\,d\rho'\Bigr),$$

where C is a constant depending only on the regularity and the dimension. We remark that

$$\frac{\delta-n}{2} \le \frac{\delta}{2}-\frac{3}{2} < -1$$

for $n > 3$ and $\delta < 1$. The result then follows using the bootstrap assumptions (73) and integrating in ρ. □

4E6. *L^1-estimates for the transport equation.* In the remainder of the article, we will use the notation

$$\mathbb{E}[g](\rho) = \int_{H_\rho}\chi_1(g)\,d\mu_\rho$$

for any regular distribution function g.

Lemma 4.20. *Let h be a regular distribution function for the massive case in the sense of Section 2E. Let g be a regular solution to $\boldsymbol{T}_\phi g = v^0 h$, with $v^0 = (1+|v|^2)^{\frac{1}{2}}$, defined on $\bigcup_{\rho\in[1,P]} H_\rho\times\mathbb{R}^n_v$, for some $P > 1$. Then, for all $\rho\in[1,P]$,*

$$\int_{H_\rho}\chi_1(|g|)\,d\mu_{H_\rho} - \int_{H_1}\chi_1(|g|)\,d\mu_{H_1} \lesssim \int_1^\rho\int_{H_{\rho'}}\int_v\Bigl(v^0|h| + \frac{1}{v^0}|\partial_{x^0}\phi g| + |\boldsymbol{T}_1(\phi)g|\Bigr)\,dv\,d\mu_{H'_\rho}\,d\rho'. \tag{97}$$

Proof. One proves first that

$$\int_{H_\rho}\chi_1(g)\,d\mu_{H_\rho} = \int_{H_1}\chi_1(g)\,d\mu_{H_1} + \int_1^\rho\int_{H_{\rho'}}\int_v\Bigl(v^0h + \Bigl(\frac{1}{v^0}\partial_{x^0}\phi - (n+1)\boldsymbol{T}_1(\phi)\Bigr)g\Bigr)\,dv\,d\mu_{H'_\rho}\,d\rho',$$

since by integration by parts

$$\int_v (\boldsymbol{T}_1(\phi)v^i + \nabla^i\phi)\,\partial_{v^i} g\,dv = -\int_v g\left((n+1)\boldsymbol{T}_1(\phi) - \frac{\partial_{x^0}\phi}{v^0}\right)dv.$$

This establishes the lemma in the case $g \geq 0$. As in the proof of Lemma 2.18, the conclusion in the general case follows after regularization of the absolute value. $\square$

For any multi-index α, let us now introduce the auxiliary function g^α:

$$g^\alpha = \widehat{Z}^\alpha(f) - \sum_{\substack{|\gamma|+|\beta|\leq|\alpha|+1\\ 1\leq|\beta|\\ 1\leq|\gamma|\leq N-\frac{n+2}{2}}} q_{\beta\gamma} Z^\gamma(\phi)\widehat{Z}^\beta(f), \tag{98}$$

where the $q_{\beta\gamma}$ are as in the statement of Lemma 4.11. One can view g^α as a renormalization of $\widehat{Z}^\alpha(f)$. The extra terms in the definition of g^α will allow us to absorb certain source terms in the equation satisfied by $\boldsymbol{T}_\phi[\widehat{Z}^\alpha(f)]$ (see Lemma 4.11) which cannot be estimated adequately because they carry too much v^0-weight, leading either to a t-loss (see Remark 4.16) or to a v^0-loss.

To perform L^1-estimates on $\widehat{Z}^\alpha(f)$, we therefore proceed as follows:

- We derive the equations satisfied by the g^α and then use them to obtain L^1-estimates for the g^α.
- We then prove the same L^1-estimates for $(v^0)^2 g^\alpha$ with $|\alpha| \leq \left\lfloor \frac{N}{2} \right\rfloor$ to take into account that the lower derivatives of f are weighted by $(v^0)^2$ in the $E_N[f]$ norm (see the definition of the norm in Section 4D).
- Finally, the L^1 estimates on g^α are then transformed into L^1-estimates on $\widehat{Z}^\alpha(f)$ using pointwise estimates on $Z^\gamma(\phi)$ for γ sufficiently small.

We start by deriving the equations for the g^α.

Lemma 4.21. *For any multi-index α, g^α satisfies the equation*

$$\begin{aligned}
\boldsymbol{T}_\phi g^\alpha = &\sum_{\substack{|\gamma|+|\beta|\leq|\alpha|+1\\ |\gamma|\geq 1,\,|\beta|\geq 1\\ |\gamma|>N-\frac{n+2}{2}}} q_{\beta\gamma}\boldsymbol{T}_1(Z^\gamma\phi)\widehat{Z}^\beta(f) + \sum_{\substack{|\gamma|+|\beta|\leq|\alpha|+1\\ 0\leq\sigma\leq n,\,1\leq|\beta|\leq|\alpha|}} \frac{1}{v^0} p^\sigma_{\gamma\beta}\,\partial_{x^\sigma} Z^\gamma(\phi)\widehat{Z}^\beta f\\
&+ \sum_{|\gamma|+|\beta|=|\alpha|} r_{\gamma\beta}\boldsymbol{T}_1(Z^\gamma\phi)\widehat{Z}^\beta f - \sum_{\substack{|\gamma|+|\beta|\leq|\alpha|+1\\ 1\leq|\beta|\\ 1\leq|\gamma|\leq N-\frac{n+2}{2}}} \boldsymbol{T}_\phi(q_{\beta\gamma})Z^\gamma(\phi)\widehat{Z}^\beta f\\
&- \sum_{\substack{|\gamma|+|\beta|\leq|\alpha|+1\\ 1\leq|\beta|\\ 1\leq|\gamma|\leq N-\frac{n+2}{2}}} q_{\beta\gamma}Z^\gamma(\phi)\Bigg(\sum_{\substack{|\kappa|+|\sigma|\leq|\beta|+1\\ 1\leq|\kappa|,\,1\leq|\sigma|\leq|\beta|}} \boldsymbol{T}_1(Z^\kappa\phi)q_{\sigma\kappa}\widehat{Z}^\sigma(f)\\
&\qquad\qquad + \sum_{\substack{|\kappa|+|\sigma|\leq|\beta|+1\\ 0\leq\omega\leq n,\,1\leq|\sigma|\leq|\beta|}} \frac{1}{v^0}p^\omega_{\kappa\sigma}\,\partial_{x^\omega}Z^\kappa(\phi)\widehat{Z}^\sigma f\\
&\qquad\qquad + \sum_{|\kappa|+|\sigma|=|\beta|} r_{\kappa\sigma}\boldsymbol{T}_1(Z^\kappa\phi)\widehat{Z}^\sigma f\Bigg).
\end{aligned}$$

Proof. This formula is a direct consequence of the product rule and a double application of Lemma 4.11. □

Based on Lemma 4.21, we now proceed to the estimates on g:

Lemma 4.22. *Assume that $\delta = \varepsilon^{\frac{1}{4}}$, and let α be a multi-index such that $|\alpha| \geq \lfloor \frac{N}{2} \rfloor + 1$. Then, g^α satisfies, for all $\rho \in [1, P]$,*

$$\mathbb{E}[|g^\alpha|](\rho) \leq (1 + C\varepsilon^{\frac{1}{4}}) e^{C\varepsilon^{1/2}} \varepsilon \rho^{\varepsilon^{1/4}},$$

where C is a constant depending only on the regularity and the dimension. If, furthermore $n > 4$, then δ can be vanishing.

Proof. Applying Lemmata 4.20 and 4.21, we obtain L^1 estimates for g^α provided we can control the source terms. The worse terms that have to be estimated are integrals in ρ of quantities of the forms

$$\int_{H_\rho}\int_v v^0 q_{\beta\gamma} |\partial Z^\gamma \phi| |\widehat{Z}^\beta f| \, dv \, d\mu_{H_\rho} \qquad \text{for } |\gamma| > N - \frac{n+2}{2}, \ |\beta| \leq N+1-|\gamma|, \tag{99}$$

$$\int_{H_\rho}\int_v |\boldsymbol{T}_\phi(q_{\beta\gamma})| |Z^\gamma \phi| |\widehat{Z}^\beta f| \, dv \, d\mu_{H_\rho} \qquad \text{for } |\gamma| \leq N - \frac{n+2}{2}, \ |\beta| \leq N+1-|\gamma|, \tag{100}$$

$$\int_{H_\rho}\int_v v^0 |q_{\beta\gamma} q_{\kappa\sigma}| |Z^\gamma \phi| |\partial Z^\kappa \phi| |\widehat{Z}^\sigma f| \, dv \, d\mu_{H_\rho} \quad \text{for } |\gamma|, |\kappa| \leq N - \frac{n+2}{2}, \ |\sigma| \leq N+1-|\kappa|, \tag{101}$$

$$\int_{H_\rho}\int_v v^0 |q_{\beta\gamma} q_{\kappa\sigma}| |Z^\gamma \phi| |\partial Z^\kappa \phi| |\widehat{Z}^\sigma f| \, dv \, d\mu_{H_\rho} \quad \text{for } |\gamma|, |\kappa| > N - \frac{n+2}{2}, \ |\sigma| \leq N+1-|\kappa|. \tag{102}$$

The other error terms are easier to handle, so, as an illustration, we will only give details below for the extra error terms

$$\int_{H_\rho}\int_v \frac{1}{v^0} |\partial_t \phi| |g^\alpha| \, dv \, d\mu_{H_\rho}, \tag{103}$$

$$\int_{H_\rho}\int_v |\boldsymbol{T}_1 \phi| |g^\alpha| \, dv \, d\mu_{H_\rho}. \tag{104}$$

We deal first with equation (99). To this end, recall that

$$|q_{\beta\gamma}| \lesssim t.$$

Furthermore, since

$$|\beta| \leq \frac{n+2}{2} \leq \left\lfloor \frac{N}{2} \right\rfloor - n \quad \text{since } N \geq 3n+4,$$

the Klainerman–Sobolev estimates of Theorem 8 on page 1564 can be applied, because the bootstrap assumption (74) is satisfied. Note that the theorem is applied estimating by the low-order part of the energy, which allows for the absorption of the additional v^0, as pointed out in Remark 4.4:

$$\begin{aligned}
\int_{H_\rho}\int_v v^0 q_{\beta\gamma} |\partial Z^\gamma \phi| |\widehat{Z}^\beta f| \, dv \, d\mu_{H_\rho} &\lesssim \int_{H_\rho}\int_v v^0 t |\partial Z^\gamma \phi| |\widehat{Z}^\beta f| \, dv \, d\mu_{H_\rho} \\
&\lesssim \int_{H_\rho} \frac{\varepsilon \rho^\delta}{t^{n-1}} \left(\frac{t}{\rho}\right)^{\frac{1}{2}} \left(\frac{\rho}{t}\right)^{\frac{1}{2}} |\partial Z^\gamma \phi| \, d\mu_{H_\rho} \\
&\lesssim \varepsilon^{\frac{3}{2}} \rho^{\delta - \frac{1}{2}} \left(\int_0^\infty t^{3-2n} r^{n-1} \, dr \right)^{\frac{1}{2}} \lesssim \varepsilon^{\frac{3}{2}} \rho^{\delta + 1 - \frac{n}{2}}.
\end{aligned}$$

We now deal with (100). To this end, we notice that since $|\partial\phi|$ decays faster than $1/t$, we have

$$|\boldsymbol{T}_\phi(q_{\beta\gamma})| \lesssim v^0.$$

Using Lemma 4.14 and the bootstrap assumption (73), we obtain

$$\begin{aligned}\int_{H_\rho}\int_v |\boldsymbol{T}_\phi(q_{\beta\gamma})||Z^\gamma\phi||\widehat{Z}^\beta f|\,dv\,d\mu_{H_\rho} &\lesssim \int_{H_\rho}\int_v |Z^\gamma\phi||v^0\widehat{Z}^\beta f|\,dv\,d\mu_{H_\rho}\\ &\lesssim \int_{H_\rho}\frac{\varepsilon^{\frac{1}{2}}\rho}{t^{\frac{n}{2}}}\cdot\frac{t}{\rho}\chi_1(|\widehat{Z}^\beta f|)\,d\mu_{H_\rho} \lesssim \varepsilon^{\frac{3}{2}}\rho^{\delta+1-\frac{n}{2}}.\end{aligned}$$

We have in the course of the estimate used Remark 2.12.

Consider now the term (101) and recall that

$$|q_{\beta\gamma}q_{\kappa\sigma}| \lesssim t^2.$$

Assume first that $|\kappa| \le N-\frac{1}{2}(n+2)$. Using Lemma 4.14 and the bootstrap assumption (73), we obtain

$$\begin{aligned}\int_{H_\rho}\int_v v^0|q_{\beta\gamma}q_{\kappa\sigma}||Z^\gamma\phi||\partial Z^\kappa\phi||\widehat{Z}^\sigma f|\,dv\,d\mu_{H_\rho} &\lesssim \int_{H_\rho} t^2\cdot\frac{\varepsilon^{\frac{1}{2}}\rho}{t^{\frac{n}{2}}}\cdot\frac{\varepsilon^{\frac{1}{2}}}{\rho t^{\frac{n}{2}-1}}\cdot\frac{t}{\rho}\chi_1(|\widehat{Z}^\sigma f|)\,d\mu_{H_\rho}\\ &\lesssim \frac{\varepsilon}{\rho}\int_{H_\rho} t^{4-n}\chi_1(|\widehat{Z}^\sigma f|)\,d\mu_{H_\rho}\\ &\lesssim \varepsilon^2\rho^{\delta+3-n}.\end{aligned}$$

We have in the course of the estimate used Remark 2.12. Now, if $|\kappa| > N-\frac{1}{2}(n+2)$, then $|\sigma| \le \frac{1}{2}(n+2) \le \lfloor\frac{N}{2}\rfloor - n$ since $N \ge 3n+4$.

Thus,

$$\begin{aligned}\int_{H_\rho}\int_v v^0|q_{\beta\gamma}q_{\kappa\sigma}||Z^\gamma\phi||\partial Z^\kappa\phi||\widehat{Z}^\sigma f|\,dv\,d\mu_{H_\rho} &\lesssim \int_{H_\rho} t^2\cdot\frac{\varepsilon^{\frac{1}{2}}\rho}{t^{\frac{n}{2}}}\frac{\varepsilon\rho^\delta}{t^n}\left(\frac{t}{\rho}\right)^{\frac{1}{2}}\left(\frac{\rho}{t}\right)^{\frac{1}{2}}|\partial Z^\kappa(\phi)|\,d\mu_\rho\\ &\lesssim \varepsilon^{\frac{3}{2}}\rho^{\delta+\frac{1}{2}}\left(\int_{H_\rho}\frac{\rho}{t}|\partial Z^\kappa(\phi)|^2\,d\mu_\rho\right)^{\frac{1}{2}}\left(\int_{H_\rho}\frac{t^5}{t^{3n}}\,d\mu_\rho\right)^{\frac{1}{2}}\\ &\lesssim \varepsilon^2\rho^{\delta+3-n}.\end{aligned}$$

The term (102) can be estimated similarly since

$$|\sigma| \le \frac{n+2}{2} \le \left\lfloor\frac{N}{2}\right\rfloor - n \quad \text{since } N \ge 3n+4.$$

Finally, for the error terms (103) and (104), we apply Proposition 4.12 and Remark 2.12:

$$\int_{H_\rho}\int_v \frac{1}{v^0}|\partial_t\phi||g^\alpha|\,dv\,d\mu_{H_\rho} \lesssim \frac{\varepsilon^{\frac{1}{2}}}{\rho^{\frac{n}{2}}}\mathbb{E}[|g^\alpha|](\rho),$$

$$\int_{H_\rho}\int_v |\boldsymbol{T}_1(\phi)g^\alpha|\,dv\,d\mu_{H_\rho} \lesssim \int_{H_\rho}\frac{\varepsilon^{\frac{1}{2}}}{\rho t^{\frac{n}{2}-1}}\frac{t}{\rho}\chi_1(|g^\alpha|)(\rho)\,d\rho \lesssim \frac{\varepsilon^{\frac{1}{2}}}{\rho^{\frac{n}{2}}}\mathbb{E}[|g^\alpha|](\rho).$$

After integration in ρ, we then obtain that g^α satisfies the integral inequality

$$\begin{aligned}\mathbb{E}[|g^\alpha|](\rho)-\mathbb{E}[|g^\alpha|](1) &\lesssim \int_1^\rho \varepsilon^{\frac{3}{2}}\delta^{-1}\rho'^{\delta}\,d\rho' + \int_1^\rho \frac{\varepsilon^{\frac{1}{2}}}{\rho'^{\frac{n}{2}}}\mathbb{E}[|g^\alpha|](\rho')\,d\rho' \\ &\le \varepsilon + C\left(\frac{\varepsilon^{\frac{3}{2}}}{\delta}\rho^\delta + \int_1^\rho \frac{\varepsilon^{\frac{1}{2}}}{\rho'^{\frac{n}{2}}}\mathbb{E}[|g^\alpha|](\rho')\,d\rho'\right),\end{aligned}$$

and thus,

$$\mathbb{E}[|g^\alpha|](\rho) \le \varepsilon\rho^\delta\left(1+C\frac{\varepsilon^{\frac{1}{2}}}{\delta}\right) + C\int_1^\rho \frac{\varepsilon^{\frac{1}{2}}}{\rho'^{\frac{n}{2}}}\mathbb{E}[|g^\alpha|](\rho')\,d\rho'$$

for some constant C depending only on the dimension and the regularity. Grönwall's lemma provides

$$\mathbb{E}[|g^\alpha|](\rho) \le \varepsilon\rho^\delta\left(1+C\frac{\varepsilon^{\frac{1}{2}}}{\delta}\right)\cdot\exp\left(C\varepsilon^{\frac{1}{2}}\int_1^\rho \frac{d\rho'}{\rho'^{\frac{n}{2}}}\right);$$

that is, if $\delta = \varepsilon^{\frac{1}{4}}$, there exists a constant $\widetilde{C}$ such that

$$\mathbb{E}[|g^\alpha|](\rho) \le (1+\widetilde{C}\varepsilon^{\frac{1}{4}})e^{\widetilde{C}\varepsilon^{1/2}}\varepsilon\rho^{\varepsilon^{1/4}}. \qquad \square$$

We then consider the lower-order derivatives of f, particularly since these low derivatives of f are weighted in v^0 in the energy:

Lemma 4.23. *Assume $\delta = \varepsilon^{\frac{1}{4}}$, and let α be a multi-index such that $|\alpha| \le \lfloor \frac{N}{2} \rfloor$. Then, $(v^0)^2 g^\alpha$ satisfies, for all $\rho \in [1, P]$,*

$$\mathbb{E}\big[|(v^0)^2 g^\alpha|\big](\rho) \le (1+C\varepsilon^{\frac{1}{4}})e^{C\varepsilon^{1/2}}\varepsilon\rho^\delta,$$

where C is a constant depending only on the regularity and the dimension. If, furthermore $n > 4$, then δ can be vanishing.

Proof. Let us compute first $\boldsymbol{T}_\phi((v^0)^2)$:

$$|\boldsymbol{T}_\phi((v^0)^2)| = |2v^0\boldsymbol{T}_\phi(v^0)| = \big|2v^0(\nabla^i\phi + \boldsymbol{T}_1(\phi)v^i)\,\partial_{v^i}v^0\big| = \big|2\big((v^0)^2\boldsymbol{T}_1(\phi) - v^0\partial_{x^0}\phi\big)\big| \lesssim |(v^0)^3\,\partial\phi|.$$

Using Proposition 4.12, one consequently obtains

$$|\boldsymbol{T}_\phi((v^0)^2)| \lesssim (v^0)^3\frac{\varepsilon^{\frac{1}{2}}}{\rho t^{\frac{n}{2}-1}},$$

with $(v^0)^2 g^\alpha$ satisfying the equation

$$\boldsymbol{T}_\phi((v^0)^2 g^\alpha) = (v^0)^2\boldsymbol{T}_\phi(g^\alpha) + \boldsymbol{T}_\phi((v^0)^2)g^\alpha.$$

Furthermore, since $|\alpha| \le \frac{1}{2}N$, the source terms of the equation satisfied by g^α (see Lemma 4.21) are only of the forms ∂Z^γ or $Z^\gamma\phi$ with

$$|\gamma| \le |\alpha| \le \left\lfloor \frac{N}{2} \right\rfloor \le N - \frac{n+2}{2} \quad \text{since } N \ge n+2.$$

As a consequence, for low derivatives acting on f, all the terms containing ϕ can be estimated pointwise.

The terms to be considered separately in this context are the same as in the proof of Lemma 4.22. The only difference is the term of (99), which is absent, since only low derivatives of ϕ appear in the expression of g^α. The estimates which one obtains are listed below. The arguments to perform the estimates are the same as above:

$$\int_{H_\rho}\int_v |p^i_{\gamma\beta}||\partial_i Z^\gamma\phi|(v^0)^2|\widehat{Z}^\beta f|\,dv\,d\mu_{H_\rho} \lesssim \int_{H_\rho} t\cdot\frac{\varepsilon^{\frac12}}{\rho t^{\frac n2-1}}\chi_1((v^0)^2|\widehat{Z}^\beta f|)\,d\mu_{H_\rho} \lesssim \varepsilon^{\frac32}\rho^{\delta+1-\frac n2},$$

$$\int_{H_\rho}\int_v |\boldsymbol{T}_\phi(q_{\beta\gamma})||Z^\gamma\phi|(v^0)^2|\widehat{Z}^\beta f|\,dv\,d\mu_{H_\rho} \lesssim \int_{H_\rho} \frac{\varepsilon^{\frac12}\rho}{t^{\frac n2}}\cdot\frac{t}{\rho}\chi_1((v^0)^2|\widehat{Z}^\beta f|)\,d\mu_{H_\rho} \lesssim \varepsilon^{\frac32}\rho^{\delta+1-\frac n2},$$

$$\begin{aligned}\int_{H_\rho}\int_v v^0|q_{\beta\gamma}q_{\kappa\sigma}||Z^\gamma\phi||\partial Z^\kappa\phi|(v^0)^2|\widehat{Z}^\sigma f|\,dv\,d\mu_{H_\rho}&\\ \lesssim \int_{H_\rho} t^2\cdot\frac{\varepsilon^{\frac12}\rho}{t^{\frac n2}}\cdot\frac{\varepsilon^{\frac12}}{\rho t^{\frac n2-1}}\cdot\frac{t}{\rho}\chi_1((v^0)^2|\widehat{Z}^\beta f|)\,d\mu_{H_\rho} &\lesssim \varepsilon^2\rho^{\delta+3-n},\end{aligned}$$

$$\int_{H_\rho}\int_v \frac{1}{v^0}|\partial_t\phi|(v^0)^2|g^\alpha|\,dv\,d\mu_{H_\rho} \lesssim \frac{\varepsilon^{\frac12}}{\rho^{\frac n2}}\mathbb{E}\big[(v^0)^2|g^\alpha|\big](\rho),$$

$$\int_{H_\rho}\int_v \boldsymbol{T}_\phi((v^0)^2)|g^\alpha|\,dv\,d\mu_{H_\rho} \lesssim \int_{H_\rho} \frac{\varepsilon^{\frac12}}{\rho t^{\frac n2-1}}\cdot\frac{t}{\rho}\chi_1((v^0)^2|g^\alpha|)\,d\mu_{H_\rho} \lesssim \frac{\varepsilon^{\frac12}}{\rho^{\frac n2}}\mathbb{E}\big[(v^0)^2|g^\alpha|\big](\rho).$$

Altogether $(v^0)^2g^\alpha$ satisfies the integral inequality

$$\mathbb{E}\big[|(v^0)^2g^\alpha|\big](\rho)-\mathbb{E}\big[|(v^0)^2g^\alpha|\big](1) \lesssim \int_1^\rho \varepsilon^{\frac32}\rho'^{\delta+1-\frac n2}\,d\rho' + \int_1^\rho \frac{\varepsilon^{\frac12}}{\rho'^{\frac n2}}\mathbb{E}\big[(v^0)^2|g^\alpha|\big](\rho')\,d\rho'.$$

The conclusion is obtained in a similar fashion as in the end of the proof of Lemma 4.22. □

Proposition 4.24. *Assume $\delta=\varepsilon^{\frac14}$. For all $\rho\in[1,P]$, we have*

- *if $n>4$,*

$$E_N[f](\rho) \le \frac{(1+C\varepsilon^{\frac14})e^{C\varepsilon^{1/2}}\varepsilon\rho^{\widetilde{C}\varepsilon^{1/4}}}{1-C\varepsilon^{\frac12}},$$

- *and if $n=4$,*

$$E_N[f](\rho) \le \frac{(1+C\varepsilon^{\frac14})e^{C\varepsilon^{1/2}}\varepsilon\rho^{\widetilde{C}\varepsilon^{1/4}}}{1-C\varepsilon^{\frac12}},$$

where C is a constant depending on the dimension and the regularity, and $\widetilde{C}=0$ when $n>4$ and $\widetilde{C}=1$ when $n=4$. In particular, for ε small enough, in dimension $n>4$, for all $\rho\in[1,P]$,

$$E_N[f](\rho)\le\tfrac32\varepsilon,$$

and, in dimension 4,

$$E_N[f](\rho)\le\tfrac32\varepsilon\rho^{\varepsilon^{1/4}}.$$

Proof. Recall the definition of g^α:

$$g^\alpha = Z^\alpha f - \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\ 1\le|\beta|\\ 1\le|\gamma|\le N-\frac{n+2}{2}}} q_{\beta\gamma}(v^i/v^0, t, x) Z^\gamma\phi \widehat{Z}^\beta f.$$

By the second triangular inequality, we immediately have

$$\mathbb{E}\big[|(v^0)^p g^\alpha|\big] \ge \left|\mathbb{E}\big[|(v^0)^p \widehat{Z}^\alpha f|\big] - \mathbb{E}\Bigg[\Bigg|\sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\ 1\le|\beta|\\ 1\le|\gamma|\le N-\frac{n+2}{2}}} q_{\beta\gamma}(v^i/v^0, t, x) Z^\gamma\phi (v^0)^p \widehat{Z}^\beta f\Bigg|\Bigg]\right|,$$

so that, using Lemma 4.14,

$$\mathbb{E}\big[|(v^0)^p \widehat{Z}^\alpha f|\big] - \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\ 1\le|\beta|\\ 1\le|\gamma|\le N-\frac{n+2}{2}}} \mathbb{E}\Bigg[\bigg|t\cdot\frac{C\varepsilon\rho}{t^{\frac{n}{2}}}(v^0)^p \widehat{Z}^\beta f\bigg|\Bigg] \le \mathbb{E}\big[|(v^0)^p g^\alpha|\big],$$

$$\mathbb{E}\big[|(v^0)^p \widehat{Z}^\alpha f|\big] - C\varepsilon^{\frac{1}{2}}\rho^{2-\frac{n}{2}} \sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\ 1\le|\beta|\\ 1\le|\gamma|\le N-\frac{n+2}{2}}} \mathbb{E}\big[|(v^0)^p \widehat{Z}^\beta f|\big] \le \mathbb{E}\big[|(v^0)^p g^\alpha|\big]$$

for some constants C.

We now split the sum above between $|\beta| = |\alpha|$ and $|\beta| < |\alpha|$ and sum over the multi-indices $|\alpha| \le N$, taking $p = 2$ for $|\alpha| \le \lfloor \frac{N}{2} \rfloor$ and $p = 0$ otherwise to build the energy $E_N[f]$. One gets, using the bootstrap assumptions (74), as well as Lemmata 4.22 and 4.23, for all ρ in $[1, P]$,

$$(1 - C\varepsilon^{\frac{1}{2}})E_N[f](\rho) \le (1 + \widetilde{C}\varepsilon^{\frac{1}{4}})e^{\widetilde{C}\varepsilon^{\frac{1}{2}}}\varepsilon\rho^\delta + C\varepsilon^{\frac{3}{2}}\rho^{\delta+2-\frac{n}{2}},$$

where C is a constant (possibly different from the one above) depending only on the dimension and the regularity. Note finally that the ρ-loss is present only in dimension 4.

As a consequence, for all ρ in $[1, P]$, if $n > 4$,

$$E_N[f](\rho) \le \frac{\le (1 + \widetilde{C}\varepsilon^{\frac{1}{4}})e^{\widetilde{C}\varepsilon^{\frac{1}{2}}}\varepsilon\rho^\delta}{1 - C\varepsilon^{\frac{1}{2}}},$$

and, if $n = 4$,

$$E_N[f](\rho) \le \frac{\le (1 + \widetilde{C}\varepsilon^{\frac{1}{4}})e^{\widetilde{C}\varepsilon^{\frac{1}{2}}}\varepsilon\rho^\delta}{1 - C\varepsilon^{\frac{1}{2}}}.$$

□

4E7. *L^2-estimates for the transport equation.* Consider here the vector X defined by

$$X = (\widehat{Z}^{\alpha_1} f, \dots, \widehat{Z}^{\alpha_q} f) \quad \text{with } |\alpha_1| \ge \left\lfloor \frac{N}{2} \right\rfloor - n + 1 \text{ and } |\alpha_q| = N, \tag{105}$$

where the multi-index α goes over all the multi-indices of length larger than $\lfloor \frac{N}{2} \rfloor - n + 1$. Using the same notation, we introduce the vector G^h:

$$G^h = (g^{\alpha_1}, \dots, g^{\alpha_q}) \quad \text{with } |\alpha_1| \geq \left\lfloor \frac{N}{2} \right\rfloor - n + 1 \text{ and } |\alpha_q| = N,$$

where g^α has been defined by equation (98). Consider finally the vector H defined by

$$H = (v^0 \widehat{Z}^{\alpha_1} f, \dots, v^0 \widehat{Z}^{\alpha_q} f) \quad \text{with } |\alpha_1| = 0 \text{ and } |\alpha_q| \leq \left\lfloor \frac{N}{2} \right\rfloor - n.$$

In a similar fashion as above, let us now consider the vector G^l defined by

$$G^l = (v^0 g^{\alpha_1}, \dots, v^0 g^{\alpha_q}) \quad \text{with } |\alpha_1| = 0 \text{ and } |\alpha_q| \leq \left\lfloor \frac{N}{2} \right\rfloor - n.$$

In the following, we will denote by $|A|_\infty$ the supremum over all the components of a vector or matrix A. Recall that, if A and B are two matrices, then,

$$|AB|_\infty \lesssim |A|_\infty |B|_\infty.$$

Throughout this section, an inequality of the form

$$|A| \lesssim \frac{1}{1 - \varepsilon^{\frac{1}{2}}}$$

appears often for some quantity A. Since we have assumed that

$$\varepsilon^{\frac{1}{2}} \leq \tfrac{1}{2},$$

$|A|$ can be bounded by 2, and we can ignore the dependency on the upper bound when $1 - \varepsilon^{\frac{1}{2}}$ appears in the denominator. We can then write

$$|A| \lesssim 1.$$

In the same spirit, $e^{C\sqrt{\varepsilon}}$ is treated as a constant.

The relation between the vectors X, H, G^h, G^l is now stated in the following lemma:

Lemma 4.25. *Assume that ε is sufficiently small. Then, the following relations between G^h, G^l, H, X hold*:

(1) *There exists a square matrix A^l such that*
- $G^l = H - A^l H$,
- *A satisfies*

$$|A^l|_\infty \lesssim \frac{\varepsilon^{\frac{1}{2}} \rho}{t^{\frac{n}{2} - 1}},$$

- *if ε is small enough, then $1 - A^l$ is invertible, and*

$$|(1 - A^l)^{-1}|_\infty \lesssim 1.$$

(2) *There exist a square matrix A' and a rectangular matrix A'' such that*
- $G^h = X - A'X - A''H$,

- (A') *and* (A'') *satisfy*

$$|A'_{ij}|_\infty \lesssim \frac{\varepsilon^{\frac{1}{2}}\rho}{t^{\frac{n}{2}-1}} \quad \textit{and} \quad |A''_{ij}|_\infty \lesssim \frac{\varepsilon^{\frac{1}{2}}\rho}{v^0 t^{\frac{n}{2}-1}},$$

- *if* ε *is small enough, then* $1-A'$ *is invertible, and*

$$|(1-A')^{-1}|_\infty \lesssim 1.$$

Proof. The proof of the algebraic relations between G^l, G^h, X, and H is a direct consequence of the definition of g^α, as stated in equation (98). The components of A are of the form $q_{\beta\gamma}Z^\gamma\phi$ with $|\gamma| \le N - \lfloor \frac{n}{2} \rfloor - 1$ (see Lemma 4.11 for the definition of $q_{\beta\gamma}$). Since $q_{\beta\gamma} \lesssim t$, the decay estimates for $Z^\gamma\phi$ (see Lemma 4.14), as well as the bootstrap assumptions (73), provide the estimates on the components of A^l, A', and A''. Standard algebraic manipulations ensure the invertibility of the square matrices $1-A^l$ and $1-A'$, and

$$|A^l|_\infty \lesssim \frac{1}{1-\varepsilon^{\frac{1}{2}}}, \quad |A'|_\infty \lesssim \frac{1}{1-\varepsilon^{\frac{1}{2}}}.$$

□

Lemma 4.26. *The commutator relation of Lemma 4.21 can be rewritten as*

$$\boldsymbol{T}_\phi G^h + \bar{\boldsymbol{A}} X = \bar{\boldsymbol{B}} H,$$

where:

- $\bar{\boldsymbol{A}} = (\bar{A}_{ij})$ *is square matrix, depending on* (t,x,v), *whose components can be bounded by*

$$|\bar{A}_{ij}| \lesssim \frac{\varepsilon^{\frac{1}{2}}v^0}{\rho t^{\frac{n}{2}-1}} \quad \textit{or} \quad |\bar{A}_{ij}| \lesssim \frac{\varepsilon^{\frac{1}{2}}}{v^0\rho t^{\frac{n}{2}-2}} \quad \textit{or} \quad |\bar{A}_{ij}| \lesssim \frac{\varepsilon^{\frac{1}{2}}v^0\rho}{t^{\frac{n}{2}}} \quad \textit{or} \quad |\bar{A}_{ij}| \lesssim \frac{\varepsilon v^0}{t^{n-3}};$$

in particular, for any regular distribution function g,

$$\int_{H_\rho}\int_v |\bar{A}_{ij} g|\, dv\, d\mu_{H_\rho} \lesssim \frac{\varepsilon^{\frac{1}{2}}}{\rho^{\frac{n}{2}-1}} \int_{H_\rho} \chi_1(|g|)\, d\mu_{H_\rho}.$$

- $\bar{\boldsymbol{B}}$ *is rectangular matrix, depending on* (t,x,v), *whose components can be bounded by*

$$|\bar{B}_{ij}| \lesssim t|f_{ij}| \quad \textit{with } \|f_{ij}\|_{L^2(H_\rho)} \lesssim \varepsilon^{\frac{1}{2}}.$$

Proof. The proof of this lemma consists essentially in rearranging the terms of the commutator formula stated in Lemma 4.21 and in the relation

$$\boldsymbol{T}_\phi(v^0) = v^0(\boldsymbol{T}_1(\phi) - \partial_t\phi),$$

which can be bounded by means of Lemma 4.14 by

$$|\boldsymbol{T}_\phi(v^0)| \lesssim \frac{(v^0)^2\varepsilon^{\frac{1}{2}}}{t^{\frac{n}{2}-1}\rho}.$$

Recall furthermore that, if $\widehat{Z}^\beta f$ can be estimated pointwise, then $\partial Z^\gamma\phi$ should be estimated in energy, as explained in the proof of Lemma 4.22. The estimates which then follow from splitting are obtained much as in the proof of Lemma 4.22. To understand the components of the matrices $\bar{\boldsymbol{A}}$ and $\bar{\boldsymbol{B}}$, and the related estimates, we finally provide the following examples:

- The matrix $\boldsymbol{A}$ contains the terms (here in absolute value)

$$|r_{\beta\gamma}\boldsymbol{T}_1(Z^\gamma\phi)| \lesssim \frac{\varepsilon^{\frac{1}{2}}v^0}{\rho t^{\frac{n}{2}-2}} \qquad \text{for } |\gamma| \le N - \frac{n+2}{2},$$

$$\left|\frac{1}{v^0}p^\delta_{\gamma\beta}\partial_\delta(Z^\gamma\phi)\right| \lesssim \frac{\varepsilon^{\frac{1}{2}}}{v^0\rho t^{\frac{n}{2}-2}} \qquad \text{for } |\gamma| \le N - \frac{n+2}{2},$$

$$|\boldsymbol{T}_\phi(q_{\beta\gamma}Z^\gamma\phi)| \lesssim \frac{\varepsilon^{\frac{1}{2}}v^0\rho}{t^{\frac{n}{2}}} \qquad \text{for } |\gamma|, |\kappa| \le N - \frac{n+2}{2},$$

$$|q_{\beta\gamma}q_{\kappa\sigma}Z^\gamma\phi\boldsymbol{T}_1(Z^\kappa\phi)| \lesssim \frac{\varepsilon v^0}{t^{n-3}} \qquad \text{for } |\gamma|, |\kappa| \le N - \frac{n+2}{2}.$$

- The matrix $\boldsymbol{B}$ contains the terms (here in absolute value)

$$\frac{1}{v^0}q_{\beta\gamma}\boldsymbol{T}_1(Z^\gamma\phi) \quad \text{with } |\gamma| > N - \frac{n+2}{2},$$

where $|q_{\beta\gamma}| \lesssim t$, and $\|(v^0)^{-1}\boldsymbol{T}_1(Z^\gamma\phi)\|^2_{L^2(H_\rho)} \lesssim \varepsilon$. □

Consider now the case when the operator $\boldsymbol{T}_\phi$ acts on the vector G^l:

Lemma 4.27. *There exists a square matrix $\hat{\boldsymbol{A}}$ such that*

$$\boldsymbol{T}_\phi G^l = \hat{\boldsymbol{A}}G^l.$$

The components $(\hat{A}_{ij})$ of the matrix $\hat{\boldsymbol{A}}$ satisfy

$$|\hat{A}_{ij}| \lesssim \frac{\varepsilon^{\frac{1}{2}}v^0}{\rho t^{\frac{n}{2}-1}} \quad \text{or} \quad |\hat{A}_{ij}| \lesssim \frac{\varepsilon^{\frac{1}{2}}}{v^0\rho t^{\frac{n}{2}-2}} \quad \text{or} \quad |\hat{A}_{ij}| \lesssim \frac{\varepsilon^{\frac{1}{2}}v^0\rho}{t^{\frac{n}{2}}} \quad \text{or} \quad |\hat{A}_{ij}| \lesssim \frac{\varepsilon^{\frac{1}{2}}v^0}{t^{n-3}};$$

in particular, for any regular distribution function g,

$$\int_{H_\rho}\int_v |\hat{A}_{ij}g|\,dv\,d\mu_{H_\rho} \lesssim \frac{\varepsilon^{\frac{1}{2}}}{\rho^{\frac{n}{2}-1}}\int_{H_\rho}\chi_1(|g|)\,d\mu_{H_\rho}.$$

Proof. This equation essentially relies on the commutator formula of Lemma 4.21. Note that for this formula, since α is very low, the first term of the right-hand side of the commutator formula

$$\sum_{\substack{|\gamma|+|\beta|\le|\alpha|+1\\|\gamma|\ge1,\,|\beta|\ge1\\|\gamma|\ge N-\lfloor n/2\rfloor}} q_{\beta\gamma}\boldsymbol{T}_1(Z^\gamma\phi)\widehat{Z}^\beta(f)$$

does not appear in the formula. Following the arguments of Lemma 4.23, in the situation when the number of derivatives is low $\big(\text{smaller than } \lfloor\frac{N}{2}\rfloor - n\big)$, the derivatives of the wave equation can all be estimated pointwise. Proposition 4.12 and Lemma 4.14 provide a relation of the form

$$\boldsymbol{T}_\phi G^l = \hat{\boldsymbol{A}}' H,$$

where $\hat{\boldsymbol{A}}'$ satisfies the same properties as $\bar{\boldsymbol{A}}$ in Lemma 4.26. Finally, we use the relation between G^l and H stated in Lemma 4.25. □

Lemma 4.28. *There exists a square matrix $\boldsymbol{A}$ and a rectangular matrix $\boldsymbol{B}$ satisfying*

- *$\boldsymbol{A} = (A_{ij})$ is square matrix, depending on (t, x, v) whose components can be bounded by*

$$|A_{ij}| \lesssim \varepsilon^{\frac{1}{2}} \cdot \frac{v^0}{\rho t^{\frac{n}{2}-1}} \quad \textit{or} \quad |A_{ij}| \lesssim \varepsilon^{\frac{1}{2}} \cdot \frac{1}{v^0 \rho t^{\frac{n}{2}-2}} \quad \textit{or} \quad |A_{ij}| \lesssim \frac{\varepsilon^{\frac{1}{2}} v^0 \rho}{t^{\frac{n}{2}}} \quad \textit{or} \quad |A_{ij}| \lesssim \varepsilon \cdot \frac{v^0}{t^{n-3}};$$

in particular, for any regular distribution function g,

$$\int_{H_\rho} \int_v |A_{ij} g| \, dv \, d\mu_{H_\rho} \lesssim \frac{1}{\rho^{\frac{n}{2}-1}} \varepsilon^{\frac{1}{2}} \int_{H_\rho} \chi_1(|g|) \, d\mu_{H_\rho},$$

- *$\boldsymbol{B}$ is a rectangular matrix, depending on (t, x, v) whose components can be bounded by*

$$|B_{ij}| \lesssim t |f_{ij}| \quad \textit{with } \|f_{ij}\|_{L^2(H_\rho)} \lesssim \varepsilon^{\frac{1}{2}}$$

such that the vector G^h satisfies the equation

$$\boldsymbol{T}_\phi G^h + \boldsymbol{A} G^h = \boldsymbol{B} G^l.$$

Proof. The proof relies on the combination of Lemmata 4.25, 4.26, and 4.27. Assuming ε small enough that the matrices $1 - A$ and $1 - A'$ are invertible, and substituting the expressions of X and H into the functions G^h and G^l in the equation satisfied by G^h stated in Lemma 4.26, one obtains the equation

$$\boldsymbol{T}_\phi G^h + \boldsymbol{A} \cdot (1 - A')^{-1} \cdot G^h = \boldsymbol{B} \cdot (1 - A)^{-1} \cdot G^l + \boldsymbol{A} \cdot (1 - A')^{-1} \cdot A'' \cdot G^l.$$

Since the components of the matrices $(1 - A')^{-1}$ and $(1 - A)^{-1}$ are both bounded by $1/(1 - \varepsilon)$, the components of the matrices $\boldsymbol{A} \cdot (1 - A')^{-1}$ and $\boldsymbol{B} \cdot (1 - A)^{-1}$ satisfy the same properties as the $\boldsymbol{A}$ and $\boldsymbol{B}$, up to constant $1/(1 - \varepsilon)$. We finally consider $C = \boldsymbol{A} \cdot (1 - A')^{-1} \cdot A''$, whose components can be bounded by means of Lemma 4.14, as follows (we remind the reader that A'' contains a $(v^0)^{-1}$):

$$|C_{ij}| \lesssim \frac{\varepsilon^{\frac{1}{2}}}{1 - \varepsilon^{\frac{1}{2}}} \cdot \frac{\varepsilon^{\frac{1}{2}}}{\rho t^{\frac{n}{2}-1}} \cdot \frac{\varepsilon^{\frac{1}{2}} \rho}{t^{\frac{n}{2}}} \quad \text{or} \quad |C_{ij}| \lesssim \frac{\varepsilon^{\frac{1}{2}}}{1 - \varepsilon^{\frac{1}{2}}} \cdot \frac{\varepsilon^{\frac{1}{2}}}{\rho t^{\frac{n}{2}-2}} \cdot \frac{\varepsilon^{\frac{1}{2}} \rho}{t^{\frac{n}{2}}} \quad \text{or} \quad |C_{ij}| \lesssim \frac{\varepsilon^{\frac{1}{2}}}{1 - \varepsilon^{\frac{1}{2}}} \cdot \frac{\varepsilon^{\frac{1}{2}}}{t^{n-3}} \cdot \frac{\varepsilon^{\frac{1}{2}} \rho}{t^{\frac{n}{2}}}.$$

One now easily notices that they can all be bounded by terms of the form

$$|C_{ij}| \lesssim t |f_{ij}| \quad \text{with } \|f_{ij}\|_{L^2(H_\rho)} \lesssim \frac{\varepsilon^{\frac{1}{2}}}{1 - \varepsilon^{\frac{1}{2}}}.$$

□

Decomposition of the solution. We consider now the set of equations of the form

$$\boldsymbol{T}_\phi G^h + \boldsymbol{A} G^h = \boldsymbol{B} G^l, \quad \text{with initial data } G^h(\rho{=}1), \tag{106}$$

where all matrices are as above. To obtain estimates on the solution G^h, we split it into two parts

$$G^h = G + G_{\mathsf{hom}}, \tag{107}$$

where G_{hom} and G are respective solutions to the Cauchy problems with initial data on the hyperboloid H_1:

$$\boldsymbol{T}_\phi G_{\mathsf{hom}} + \boldsymbol{A} G_{\mathsf{hom}} = 0 \qquad \text{with } G_{\mathsf{hom}} = G^h, \tag{108}$$

$$\boldsymbol{T}_\phi G + \boldsymbol{A} G = \boldsymbol{B} G^l \quad \text{with } G = 0. \tag{109}$$

We proceed by evaluating both components individually.

Homogeneous part. Commuting equation (108) n-times with the vector field $\widehat{Z}$ yields

$$\boldsymbol{T}_\phi(\widehat{Z}^\alpha G_{\mathsf{hom}}) = -\widehat{Z}^\alpha(\boldsymbol{A} G_{\mathsf{hom}}) + [\boldsymbol{T}_\phi, \widehat{Z}^\alpha] G_{\mathsf{hom}} \tag{110}$$

for $|\alpha| \le n$. Applying estimate (97) directly to this equation would yield problematic terms in the estimate for the same reason as discussed after (98). As before, we introduce an auxiliary function g^α_{hom} analogous to that defined in (97). Applying estimate (97) to the transport equation of g^α_{hom} yields

$$\mathbb{E}[|g^\alpha_{\mathsf{hom}}|](\rho) \le \varepsilon(1 + C\varepsilon)e^{C\sqrt{\varepsilon}} \tag{111}$$

for $|\alpha| \le n$. In turn, for G_{hom} we obtain, for $n > 4$,

$$\sum_{|\alpha| \le n} \mathbb{E}[|\widehat{Z}^\alpha G_{\mathsf{hom}}|] \le \frac{\varepsilon(1 + C\varepsilon^{\frac{1}{2}})e^{C\varepsilon^{\frac{1}{2}}}}{1 - \varepsilon^{\frac{1}{2}}},$$

and, if $n = 4$,

$$\sum_{|\alpha| \le n} \mathbb{E}[|\widehat{Z}^\alpha G_{\mathsf{hom}}|] \le \frac{\varepsilon(1 + C\varepsilon^{\frac{1}{2}})e^{C\varepsilon^{\frac{1}{2}}}\rho^\delta}{1 - \varepsilon^{\frac{1}{2}}},$$

where the constant C does depend only on the dimension and the regularity N.

This yields

$$\|G_{\mathsf{hom}}\|_{\mathbb{P},n}(\rho) \le \begin{cases} \dfrac{\varepsilon(1 + C\varepsilon)e^{C\varepsilon}}{1 - \varepsilon^{\frac{1}{2}}} & n > 4, \\[2ex] \dfrac{\varepsilon(1 + C\varepsilon)e^{C\varepsilon}\rho^\delta}{1 - \varepsilon^{\frac{1}{2}}} & n = 4. \end{cases} \tag{112}$$

In combination with the Klainerman–Sobolev estimate (42) this implies the following lemma.

Lemma 4.29. *The following estimate holds*:

$$\int_{v\in\mathbb{R}^n} |G_{\mathsf{hom}}| \frac{dv}{v^0} \leq \begin{cases} \dfrac{\varepsilon(1+C\varepsilon^{\frac{1}{2}})e^{C\varepsilon^{\frac{1}{2}}}}{(1+t)^n} & n > 4, \\ \dfrac{\varepsilon(1+C\varepsilon^{\frac{1}{2}})e^{C\varepsilon^{\frac{1}{2}}}\rho^\delta}{(1+t)^n} & n = 4, \end{cases} \tag{113}$$

where the constant C does depend only on the dimension and the regularity N.

Inhomogeneous part. Before giving the full proof of the actual L^2-decay estimates for the inhomogeneous part of G^h, let us explain the main ideas on a simple model problem. Assume that T is a transport operator such as the relativistic transport operator or even just the classical one and that f is a function of (t,x,v) satisfying

$$\boldsymbol{T}(f) = hg,$$

where $h = h(t,x)$ is uniformly bounded in L^2_x and such that g is itself a solution to the free transport equation $\boldsymbol{T}(g) = 0$ with g regular enough that $L^1_{x,v}$-bounds hold for g and decay estimates similar to our Klainerman–Sobolev inequality can be applied for the velocity averages of g. The aim is to prove L^2_x-decay estimates on $\int_v |f|\,dv$, the difficulty being that h has very little regularity so that we cannot commute the equation. Instead, note that, by uniqueness, $f = gH$, where H is the solution to the inhomogeneous transport equation $\boldsymbol{T}(H) = h$ with zero data. Indeed,

$$\boldsymbol{T}(gH) = \boldsymbol{T}(g)H + g\boldsymbol{T}(H) = gh,$$

since $\boldsymbol{T}(g) = 0$. Now,

$$\left\|\int_v gH\,dv\right\|_{L^2_x} \lesssim \left\|\left(\int_v |g|\,dv\right)^{\frac{1}{2}}\left(\int_v |g|H^2\,dv\right)^{\frac{1}{2}}\right\|_{L^2_x} \lesssim \left\|\left(\int_v |g|\,dv\right)^{\frac{1}{2}}\right\|_{L^\infty_x}\left\|\int_v |g|H^2\,dv\right\|^2_{L^1_x}.$$

Since we have assumed g to solve the free transport equations and to be as regular as needed, we know that we have some decay for $\left\|\left(\int_v |g|\,dv\right)^{\frac{1}{2}}\right\|_{L^\infty_x}$. Thus, it remains only to prove boundedness for $\left\|\int_v |g|H^2\,dv\right\|_{L^1_x}$. This can be obtained using again the transport equation for gH and the associated approximate conservation laws. Indeed, we have

$$\boldsymbol{T}(gH^2) = 2ghH,$$

and thus, we need to estimate an integral of the form $\int_{t,x,v} |ghH|\,dt\,dx\,dv$. This is done as follows. First,

$$\begin{aligned}\int_{t,x,v} |ghH|\,dt\,dx\,dv &= \int_t \int_{x,v} |g|^{\frac{1}{2}}|h||g|^{\frac{1}{2}}H\,dx\,dv\,dt \\ &\lesssim \int_t \left(\int_{x,v} |g||h|^2\,dx\,dv\right)^{\frac{1}{2}}\left(\int_{x,v} |g|H^2\,dx\,dv\right)^{\frac{1}{2}}dt.\end{aligned}$$

It follows that, if one can obtain enough decay for $\left(\int_{x,v} |g|(x,v)|h|^2(x)\,dx\,dv\right)^{\frac{1}{2}}$, then the estimate can close via a Grönwall-type inequality. For the decay estimate, simply note again that

$$\left|\int_{x,v} |g|(t,x,v)|h|^2(t,x)\,dx\,dv\right| \lesssim \left\|\int_v g\,dv\right\|_{L^\infty_x} \|h(t,x)\|^2_{L^2_x}.$$

This concludes the discussion of the estimates for the model problem. To estimate the inhomogeneous part of G^h, we will essentially follow this strategy except that

- we need to work with systems;
- the operator $\boldsymbol{T}$ needs to be replaced by $\boldsymbol{T}_\phi$ (or rather $\boldsymbol{T}_\phi + A$);
- the matrix $\boldsymbol{B}$ replacing h is not uniformly bounded in L^2_x (there is a t-loss);
- the vector replacing g does not quite satisfy a homogeneous transport equation;
- and finally, in all steps, we need to keep track of the exact decay rates in ρ to make sure the time integrals converge.

To perform the estimates on G, we first notice the following useful decomposition. Let $\boldsymbol{K}$ be the matrix solution to the equation

$$\boldsymbol{T}_\phi \boldsymbol{K} + \boldsymbol{A}\boldsymbol{K} + \boldsymbol{K}\hat{\boldsymbol{A}} = \boldsymbol{B}, \quad \text{with } \boldsymbol{K}|_{H_1} = 0. \tag{114}$$

Then, an immediate calculation proves that the vector $\boldsymbol{K}G^l$ satisfies the equation

$$\boldsymbol{T}_\phi(\boldsymbol{K}G^l) + \boldsymbol{A}(\boldsymbol{K}G^l) = \boldsymbol{B}G^l, \quad \text{with } (\boldsymbol{K}G^l)|_{H_1} = 0.$$

By uniqueness of solutions to the Cauchy problem and (109), we obtain

$$G = \boldsymbol{K}G^l. \tag{115}$$

Before performing the estimates on G, let us remind that

- using the bootstrap assumption (74) and Theorem 8 on page 1564, the elements of the vector G^l can be estimated pointwise, and, by Lemma 4.27,
$$\int_v |G^l|_\infty\,dv \lesssim \varepsilon t^{\delta-n},$$
- the components of the source terms in the equation satisfied by $\boldsymbol{K}$ can be estimated in $L^2(H_\rho)$ (see Lemma 4.28).

Following the strategy described in the Introduction, we introduce the scalar function

$$|\boldsymbol{K}\boldsymbol{K}\boldsymbol{G}|_\infty = \sum_{\alpha,\beta,\gamma,\kappa,\mu} |K^\beta_\alpha K^\kappa_\gamma G^l_\mu|,$$

where K^β_α, K^κ_γ are the components of the matrix $\boldsymbol{K}$ and G^l_μ are the components of the vector G^l, and where the sum is taken over all possible combinations of two elements of $\boldsymbol{K}$ and G^l. One furthermore easily checks that each element of this sum satisfies the equation

$$\boldsymbol{T}_\phi(K^\beta_\alpha K^\kappa_\gamma G^l_\mu) = (A^\sigma_\alpha K^\beta_\sigma + K^\sigma_\beta \hat{A}^\alpha_\sigma) K^\kappa_\gamma G^l_\mu + K^\beta_\alpha (A^\sigma_\gamma K^\kappa_\sigma + K^\sigma_\gamma \hat{A}^\kappa_\sigma) G^l_\mu + K^\beta_\alpha K^\kappa_\gamma \hat{A}^\sigma_\mu G^l_\sigma \tag{116}$$

$$- (B^\alpha_\beta K^\kappa_\gamma + K^\alpha_\beta B^\kappa_\gamma) G^l_\mu. \tag{117}$$

As a consequence, we have:

Lemma 4.30. *Assume $\delta = \varepsilon^{\frac{1}{4}}$. The function $|\boldsymbol{KKG}|_\infty$ satisfies*

- *for $n > 4$, for all ρ in $[1, P]$,*
$$\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho) \lesssim \varepsilon^{\frac{3}{2}};$$
- *for $n = 4$, for all ρ in $[1, P]$,*
$$\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho) \lesssim \varepsilon \rho^{\varepsilon^{\frac{1}{4}}}.$$

Proof. The right-hand side of (116) can easily be estimated, using the properties of the matrices $\boldsymbol{A}$ and $\hat{\boldsymbol{A}}$ stated in Lemmata 4.28 and 4.27, by

$$\int_{H_\rho} \chi_1\big(\big|(A^\sigma_\alpha K^\beta_\sigma + K^\sigma_\beta \hat{A}^\alpha_\sigma) K^\kappa_\gamma G^l_\mu + K^\beta_\alpha (A^\sigma_\gamma K^\kappa_\sigma + K^\sigma_\gamma \hat{A}^\kappa_\sigma) G^l_\mu + K^\beta_\alpha K^\kappa_\gamma \hat{A}^\sigma_\mu G^l_\sigma\big|\big)\, d\mu_{H_\rho} \lesssim \frac{\sqrt{\varepsilon}}{\rho^{\frac{n}{2}-1}} \mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho).$$

Furthermore, by the Cauchy–Schwarz inequality, as well as the property of the matrix $\boldsymbol{B}$ stated in Lemma 4.28, one gets, when estimating (117):

$$\begin{aligned}\int_v \big(\big|(B^\alpha_\beta K^\kappa_\gamma + K^\alpha_\beta B^\kappa_\gamma) G^l_\mu\big|\big)\, dv &\lesssim |\boldsymbol{B}|_\infty \chi_1(|G^l|)^{\frac{1}{2}} \chi_1(|\boldsymbol{KKG}|_\infty)^{\frac{1}{2}} \\ &\lesssim |\boldsymbol{B}|_\infty \frac{\varepsilon^{\frac{1}{2}}}{t^{\frac{n}{2}-\frac{\delta}{2}}} \chi_1(|\boldsymbol{KKG}|_\infty)^{\frac{1}{2}},\end{aligned}$$

$$\begin{aligned}\int_{H_\rho}\int_v \big(\big|(B^\alpha_\beta K^\kappa_\gamma + K^\alpha_\beta B^\kappa_\gamma) G^l_\mu\big|\big)\, dv\, d\mu_{H_\rho} &\lesssim \varepsilon^{\frac{1}{2}} \rho^{\frac{\delta+2-n}{2}} \|t^{-1}\boldsymbol{B}\|_{L^2(H_\rho)} \big(\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho)\big)^{\frac{1}{2}} \\ &\lesssim \varepsilon \rho^{\frac{\delta+2-n}{2}} \big(\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho)\big)^{\frac{1}{2}}.\end{aligned}$$

As a consequence $|\boldsymbol{KKG}|_\infty$ satisfies, for all $\rho \in [1, P]$, the integral inequality

$$\begin{aligned}\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho) &\lesssim \int_1^\rho \frac{\sqrt{\varepsilon}}{\rho'^{\frac{n}{2}-1}} \mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho') + \varepsilon \rho'^{\frac{\delta}{2}+1-\frac{n}{2}} \big(\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho')\big)^{\frac{1}{2}}\, d\rho' \\ &\lesssim \int_1^\rho \frac{\sqrt{\varepsilon}}{\rho'^{\frac{n}{2}-1}} \mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho') + \varepsilon^{\frac{3}{4}} \rho^{\frac{\delta}{2}+\frac{1}{2}-\frac{n}{4}} \cdot \varepsilon^{\frac{1}{4}} \rho^{\frac{1}{2}-\frac{n}{4}} \big(\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho')\big)^{\frac{1}{2}}\, d\rho' \\ &\lesssim \frac{\varepsilon^{\frac{3}{2}}}{\frac{n}{2}-2-\delta}(1-\rho^{\delta+2-\frac{n}{2}}) + \int_1^\rho \frac{\sqrt{\varepsilon}}{\rho'^{\frac{n}{2}-1}} \mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho')\, d\rho'.\end{aligned}$$

Assume now $n > 4$. Then, Grönwall's inequality implies immediately, for all ρ in $[1, P]$,

$$\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho) \lesssim e^{C\varepsilon^{1/2}} \varepsilon^{\frac{3}{2}} \lesssim \varepsilon^{\frac{3}{2}}$$

for some constant C, depending only on n and N.

In the case $n = 4$, the integral inequality becomes

$$\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho) \lesssim \frac{\varepsilon^{\frac{3}{2}}}{\delta}\rho^\delta + \int_1^\rho \frac{\sqrt{\varepsilon}}{\rho'}\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho')\, d\rho'. \tag{118}$$

To perform the estimates in this case, we make the bootstrap assumption: let $P' < P$ be the maximal P such that, for all ρ in $[1, P']$,

$$\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho) \leq C\varepsilon\rho^{\varepsilon^{1/4}}, \tag{119}$$

where C depends only on the dimension n and the regularity N. Inserting the bootstrap assumption (119) in (118), one gets, for $\delta = \varepsilon^{\frac{1}{4}}$,

$$\mathbb{E}[|\boldsymbol{KKG}|_\infty](\rho) \lesssim \varepsilon^{1+\frac{1}{4}}\rho^{\varepsilon^{1/4}}$$

for all $\rho < P'$. The bootstrap assumption (119) can then be improved for ε small enough. □

We can finally state the L^2-estimates for f:

Proposition 4.31. *Assume that ε is sufficiently small. Under the bootstrap assumptions (73) and (74), the following estimate holds for all multi-indices α such that $\lfloor \frac{N}{2} \rfloor - n + 1 \leq |\alpha| \leq N$:*

$$\int_{H_\rho} \frac{t}{\rho}\left(\int_v |\widehat{Z}^\alpha f|\,\frac{dv}{v^0}\right)^2 d\mu_{H_\rho} \lesssim \rho^{2\varepsilon^{1/4}-n}\varepsilon^2.$$

Proof. We first notice that, by Lemma 4.25, for α such that $\frac{N}{2} + 1 \leq |\alpha| \leq N$,

$$\begin{aligned}
&\left(\int_{H_\rho} \frac{t}{\rho}\left(\int_v |\widehat{Z}^\alpha f|\,\frac{dv}{v^0}\right)^2 d\mu_{H_\rho}\right)^{\frac{1}{2}}\\
&\qquad\lesssim \frac{1}{1-\varepsilon^{\frac{1}{2}}}\left(\int_{H_\rho} \frac{t}{\rho}\left(\int_v |g^\alpha|\,\frac{dv}{v^0}\right)^2 d\mu_{H_\rho}\right)^{\frac{1}{2}}\\
&\qquad\lesssim \frac{1}{1-\varepsilon^{\frac{1}{2}}}\left\{\left(\int_{H_\rho} \frac{t}{\rho}\left(\int_v |G^\alpha_{\mathrm{hom}}|\,\frac{dv}{v^0}\right)^2 d\mu_{H_\rho}\right)^{\frac{1}{2}} + \left(\int_{H_\rho} \frac{t}{\rho}\left(\int_v |G^\alpha|\,\frac{dv}{v^0}\right)^2 d\mu_{H_\rho}\right)^{\frac{1}{2}}\right\},
\end{aligned}$$

where G^α_{hom} and G^α are the components of G_{hom} and G respectively. The first term of this sum is estimated by means of Lemma 4.29:

$$\left(\int_{H_\rho} \frac{t}{\rho}\left(\int_v |G^\alpha_{\mathrm{hom}}|\,\frac{dv}{v^0}\right)^2 d\mu_{H_\rho}\right)^{\frac{1}{2}} \lesssim \varepsilon\rho^{\frac{\delta-n}{2}}\left(\int_0^\infty \frac{y^{n-1}}{\langle y\rangle^{2n}}\,dy\right)^{\frac{1}{2}}. \tag{120}$$

The second term of the sum is estimated as follows: Let us denote by G^α the components of the vector G; we have

$$\begin{aligned}
\left(\int_{H_\rho} \frac{t}{\rho}\left(\int_v |G^\alpha|\,\frac{dv}{v^0}\right)^2 d\mu_{H_\rho}\right)^{\frac{1}{2}} &\lesssim \left(\int_{H_\rho} \frac{t}{\rho}\left(\int_v |\boldsymbol{K}^k_\alpha G^l_k|\,\frac{dv}{v^0}\right)^2 d\mu_{H_\rho}\right)^{\frac{1}{2}}\\
&\lesssim \sum_k \int_{H_\rho} \frac{t}{\rho}\left(\int_v |G^l_k|\,\frac{dv}{v^0}\right)\left(\int_v |(\boldsymbol{K}^k_\alpha)^2 G^l_k|\,\frac{dv}{v^0}\right) d\mu_{H_\rho},
\end{aligned} \tag{121}$$

where the last sum over k is taken of all the components of $\boldsymbol{K}$ and G^l and is consequently finite. In combination with the pointwise decay for G^l and the estimate in Lemma 4.30, this implies

$$\int_{H_\rho} \frac{t}{\rho}\left(\int_v |G^\alpha| \frac{dv}{v^0}\right)^2 d\mu_{H_\rho} \lesssim \rho^{2\varepsilon^{1/4}-n}\varepsilon^2. \tag{122}$$

In combination with the bound on the homogeneous part above, this yields the claim. □

Appendix A: Distribution functions for massive particles with compact support in x

Theorem 2, on page 1543, and Theorem 5, on page 1547, require that the initial data be given on the initial hyperboloid H_1 instead of a more traditional $t = \text{const}$ hypersurface. In this appendix, we explain how we can go from the $t = 0$ hypersurface to H_1, provided the initial data on $t = 0$ has sufficient decay in x. For simplicity, consider the homogeneous massive transport equation with initial data f_0 given at $t = 0$. Assume that the support of f_0 is contained in the ball of radius R. Without loss of generality, we may translate the problem in time, so that we now consider the problem with data at time $t = \sqrt{R^2+1}$:

$$\boldsymbol{T}_m(f) = 0, \tag{123}$$

$$f(t = \sqrt{R^2+1}) = f_0. \tag{124}$$

Now, by the finite speed of propagation, the solution to this problem vanishes outside of the cone

$$\mathcal{C}(R) \equiv \left\{(t,r,\omega)\;\middle|\; t-r = \sqrt{R^2+1}-R,\ \omega\in\mathbb{S}^{n-1},\ t \geq \sqrt{R^2+1}\right\} \cup \left\{(t,r,\omega)\;\middle|\; t+r = \sqrt{R^2+1}+R,\ \omega\in\mathbb{S}^{n-1},\ t \leq \sqrt{R^2+1}\right\}$$

depicted in Figure 3.

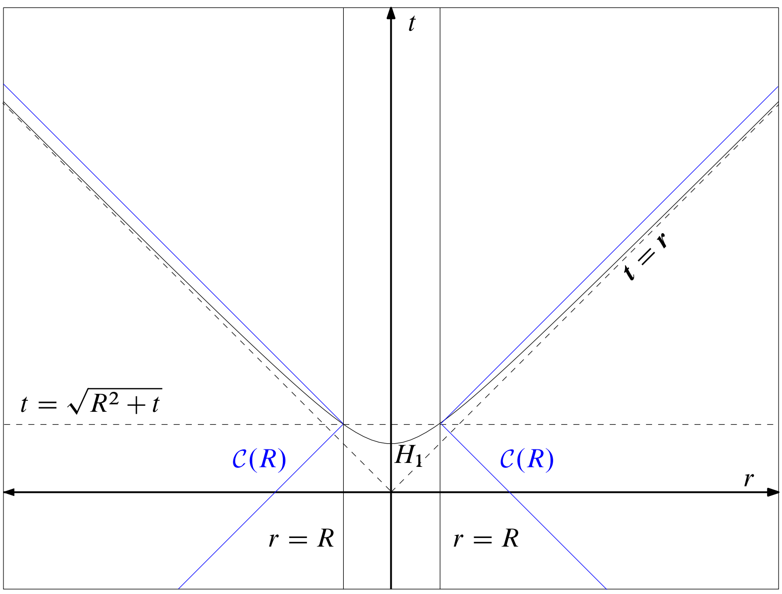


Figure 3. The trace of a distribution function with compact support on H_1.

Thus, the trace of f on H_1 is compactly supported and as a consequence, the norm appearing on the right-hand side of Theorem 2 is finite. Recall also that Theorem 2 gives pointwise estimates for $t \geq \sqrt{1+|x|^2}$. On the other hand, the region $t < \sqrt{1+|x|^2}$ lies in the exterior of $\mathcal{C}(R)$ and hence $f(t,x) = 0$ for $t < \sqrt{1+|x|^2}$. Thus, for compactly supported initial data given on some $t = \text{const}$ hypersurfaces, we can apply Theorem 2 and obtain a $1/t^n$ decay uniformly in x.

Finally, let us mention that the above arguments can be easily adapted to the nonlinear massive Vlasov–Nordström system for small initial data. Thus, once again, the use of hyperboloids in Theorem 5 is merely technical.

Appendix B: Integral estimate

Lemma B.1. *Let n be a positive integer. Consider α, β such that*

$$\alpha + \beta > n.$$

There exists a constant $C_{\alpha,\beta,n}$ such that the following estimate is true: for all $t > 0$, if $\beta \neq 1$, then

$$\int_0^\infty \frac{r^{n-1}\,dr}{(1+t+r)^\alpha(1+|t-r|)^\beta} \leq \frac{C_{\alpha,\beta,n}}{t^{\alpha+\beta-n}}(1+t^{\beta-1}).$$

If $\beta = 1$, then

$$\int_0^\infty \frac{r^{n-1}\,dr}{(1+t+r)^\alpha(1+|t-r|)} \leq \frac{C_{\alpha,n}}{t^{\alpha+1-n}}(1+\log(t+1)).$$

Proof. Let

$$A = \int_0^\infty \frac{r^{n-1}\,dr}{(1+t+r)^\alpha(1+|t-r|)^\beta}.$$

First, let us make the change of variable

$$r = ty.$$

This gives

$$A = \frac{1}{t^{\alpha+\beta-n}} \int_0^\infty \frac{y^{n-1}\,dy}{(1/t+1+y)^\alpha(1/t+|1-y|)^\beta}.$$

The first part of the denominator is bounded below by $(1+y)^\alpha$, so that

$$A \leq \frac{1}{t^{\alpha+\beta-n}} \int_0^\infty \frac{y^{n-1}\,dy}{(1+y)^\alpha(1/t+|1-y|)^\beta}.$$

We then cut the integral in two at the value $r = 2$. Let us thus introduce the constants $K_{\alpha,\beta,n}$ by

$$K_{\alpha,\beta,n} = \int_2^\infty \frac{y^{n-1}\,dy}{(1+y)^\alpha(y-1)^\beta}.$$

A can then be bounded by

$$A \leq \frac{1}{t^{\alpha+\beta-n}}\left(K_{\alpha,\beta,n} + 2^n \int_0^1 \frac{dy}{(1/t+1-y)^\beta}\right).$$

The remaining integral can be computed: for $\beta \neq 1$,

$$\int_0^1 \frac{dy}{(1/t+1-y)^\beta} = \frac{1}{\beta-1}\left(t^{\beta-1} - \frac{1}{(1/t+1)^{\beta-1}}\right) = \frac{t^{\beta-1}}{1-\beta}\left(1-\frac{1}{1+t^{\beta-1}}\right) \leq C_\beta t^{\beta-1}.$$

If $\beta = 1$, we get

$$\int_0^1 \frac{dy}{(1/t+1-y)} = \log(1+t).$$

We finally get the announced result: if $\beta \neq 1$, then

$$\int_0^\infty \frac{r^{n-1}\,dr}{(1+t+r)^\alpha(1+|t-r|)^\beta} \leq \frac{C_{\alpha,\beta,n}}{t^{\alpha+\beta-n}}(1+t^{\beta-1}).$$

If $\beta = 1$, then

$$\int_0^\infty \frac{r^{n-1}\,dr}{(1+t+r)^\alpha(1+|t-r|)^1} \leq \frac{C_{\alpha,\beta,n}}{t^{\alpha+1-n}}(1+\log(t+1)).$$ □

Appendix C: Geometry of Vlasov fields

In this section we present the necessary elements to understand the underlying geometry of Vlasov fields on an arbitrary curved manifold. In particular, we will present, with some amount of detail,

- the geometry of the tangent bundle;
- the notion of complete lift, which is an essential tool to understand the commutators with the Vlasov field;
- how the ambient geometry of the tangent bundle can be reduced to the mass shell.

Most of the calculations will be left to the reader.

The reader who wishes to know more about the geometry of the tangent bundle can refer to the book by Crampin and Pirani [1986]. This section has also been greatly inspired by the work of Sarbach et al. [2014a; 2014b].

Throughout this section, let M be an $(n+1)$-dimensional smooth, oriented manifold, endowed with a Lorentzian metric g, of signature $(-,+,+,+)$. The Levi-Civita connection is denoted by ∇. The tangent bundle of M is denoted by TM. We furthermore assume that M is time oriented: there exists a uniformly timelike vector field T chosen, by convention, to be future pointing.

Geometry of the tangent bundle. This section is a reminder of some elementary geometric facts.

Definition C.1. The tangent bundle of M is the disjoint union of the tangent plane to M:

$$TM = \bigsqcup_{x\in M} T_x M.$$

TM is a vector bundle of dimension $2n+2$ over M, with fibre $\mathbb{R}^{n+1}$, and projection given by

$$\pi : V = (x,v) \in TM \longmapsto x \in M.$$

Consider a chart (U, x^α). One defines on the open set TU the system of coordinates

$$(x^\alpha, v^\alpha = dx^\alpha).$$

This system of coordinates provides a local trivialization of TM by

$$V \in TM \longrightarrow (\pi(V), v^\alpha(V)) \in U \times \mathbb{R}^{n+1}.$$

In these coordinates, the metric reads

$$g = g_{\alpha\beta} v^\alpha \otimes v^\beta.$$

We now consider the tangent space to the tangent space of M, denoted by TTM. If $V = (x, v)$ is a point of TM, the tangent space to TM at the point (x, v), denoted by $T_{(x,v)}TM$, is generated by the vectors

$$\left\{ \frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial v^\alpha} \right\}.$$

Let (x, v) be in TM. If $t \mapsto \sigma(t)$ is a curve on M, with $\sigma(0) = x, \dot{\sigma}(0) = v$, the *natural lift* of σ is the curve of TM defined by

$$\sigma^h = t \mapsto (\sigma(t), \dot{\sigma}(t)).$$

As a consequence, any curve in M can be obtained by projection on M of a curve in TM. We consequently define:

Definition C.2 (vertical space). The push forward $\pi_\star$ of the mapping π defines, for $(x, v) \in TM$, a surjective mapping $T_{(x,v)}TM$ into $T_x M$. The kernel of $\pi_\star : T_{(x,v)}TM \to M$ is the horizontal space $V_{(x,v)}M$ at (x, v). This is a subspace of dimension $n + 1$ of $T_{(x,v)}TM$. In a system of coordinates (x^α, v^α), it is generated by

$$\left\{ \frac{\partial}{\partial v^\alpha} \right\}.$$

If $t \mapsto \sigma(t)$ is a curve on M, with $\sigma(0) = x$ and $\dot{\sigma}(0) = v$, the *horizontal lift* of σ is the curve of TM defined by

$$\sigma^h = t \mapsto (\sigma(t), V(t)),$$

where $V(t)$ is the vector field along the curve σ obtained by the parallel transport of v along the curve σ. In the coordinates (x^α, v^α), V obeys the differential equation

$$\nabla^\alpha_{\dot{\sigma}} V = \dot{V}^\alpha + \Gamma^\alpha_{\beta\gamma} \dot{\sigma}^\beta V^\gamma = 0 \quad \text{with } V(0) = v,$$

where the $\Gamma^\alpha_{\beta\gamma}$ are the Christoffel symbols of the connection. The tangent vector to the curve σ is given, in the coordinates (x^α, v^α), at $t = 0$, by

$$\dot{\sigma}(0)^\alpha \frac{\partial}{\partial x^\alpha} + \dot{V}^\alpha(0) \frac{\partial}{\partial v^\alpha} = \dot{\sigma}(0)^\alpha \left(\frac{\partial}{\partial x^\alpha} - \Gamma^\beta_{\alpha\gamma} v^\gamma \frac{\partial}{\partial v^\beta} \right).$$

The vector

$$\dot{\sigma}(0)^\alpha \left(\frac{\partial}{\partial x^\alpha} - \Gamma^\beta_{\alpha\gamma} v^\gamma \frac{\partial}{\partial v^\beta} \right)$$

is the *horizontal lift* of the vector $\dot{\sigma}(0)$. This definition depends only on the vector v in $T_x M$.

Definition C.3. The horizontal space $H_{(x,v)}M$ at (x, v) is the subspace of $T_{(x,v)}TM$ generated by the horizontal vectors

$$e_\alpha = \frac{\partial}{\partial x^\alpha} - \Gamma^\beta_{\alpha\gamma} v^\gamma \frac{\partial}{\partial v^\beta}.$$

It is independent of the chosen system of coordinates, and has trivial intersection with the vertical subspace of $T_{(x,v)}TM$.

Finally, the tangent space is endowed with a metric:

Definition C.4. The Sasaki metric on the tangent bundle TM is the metric of signature $(2n, 2)$ defined, in coordinates, by

$$g_s(e_\alpha, e_\beta) = g_s\left(\frac{\partial}{\partial v^\alpha}, \frac{\partial}{\partial v^\beta}\right) = g_{\alpha\beta},$$

and

$$g_s\left(e_\alpha, \frac{\partial}{\partial v^\beta}\right) = 0.$$

Geodesic spray, and its commutators. We now turn our attention to the lift of geodesics to the tangent bundle.

Definition C.5. Let γ be a geodesic with $\gamma(0) = x$ and $\dot\gamma(0) = v$. The vector of $H_{(x,v)}M$ obtained by performing the horizontal lift of v, denoted by T, is given by

$$T = v^\alpha e_\alpha = v^\alpha\left(\frac{\partial}{\partial x^\alpha} - \Gamma^\beta_{\alpha\gamma} v^\gamma \frac{\partial}{\partial v^\beta}\right).$$

This defines globally a vector field on TM, called the geodesic spray.

Contrary to the geodesic flow, this vector field is defined globally on the manifold. Furthermore, since its integral curves are the natural lift of geodesics, it naturally models the behaviour of freely falling particles in the context of general relativity.

As we have seen earlier, one key aspect of the this work relies on the commutators with the transport operator T (see Section 2G2). The right tool to understand this is the notion of *complete lift* (see Section 2G1). It can be introduced as follows. Consider a vector field X on M. Assume that (locally) this vector field arises from a flow ϕ^t:

$$\frac{d\phi^t}{dt} = X(\phi^t).$$

The mapping ϕ^t can naturally be lifted into a mapping of TM by the formula

$$\phi^t_\star = (\phi^t, d\phi^t).$$

This immediately defines a vector field $\widehat{X}$ on TM by the formula

$$\frac{d\phi^t_\star}{dt} = \widehat{X}(\phi^t_\star).$$

It is also possible to have a definition relying on Lie transport along curves.

Definition C.6. Let X be a vector field on M. The complete lift of the vector field X on M into a vector field $\widehat{X}$ on TM is done as follows: Let $p \in M$ and consider $X(p)$. Let γ be an integral curve of X with initial data

$$\begin{cases} \gamma(0) = p, \\ \dfrac{d\gamma}{ds}(s) = X(\gamma(s)). \end{cases}$$

Let $v \in T_pM$, and consider the vector field Y defined on γ by Lie transporting the vector v along γ. It obeys the equation

$$\mathcal{L}_X Y = [X, Y] = 0 \quad \text{with } Y(p) = v.$$

This defines a curve $\Gamma = (\gamma, Y(\gamma))$ on TM. The mapping

$$\widehat{X} : TM \to TTM,$$
$$(x, v) \mapsto \frac{d\Gamma}{ds}(0),$$

defines a vector field on TM, defined as being the complete lift of X on TM.

An expression of the complete lift of the vector $W = W^\alpha \partial_{x^\alpha}$ in adapted coordinates is given in [Crampin and Pirani 1986, page 330] (and on page 288 of that work for affine transformations) by

$$\widehat{W} = W^\alpha \frac{\partial}{\partial x^\alpha} + v^\beta \frac{\partial W^\alpha}{\partial x^\beta} \frac{\partial}{\partial v^\alpha}. \tag{125}$$

This expression can also be written as

$$\widehat{W} = W^\alpha e_\alpha + v^\beta \nabla_\beta W^\alpha \frac{\partial}{\partial v^\alpha}. \tag{126}$$

One of the main interests of the complete lift is its relation with the commutators of the geodesic spray. It is possible to give a precise characterization of the commutators with the geodesic spray which arise from vector fields on the base manifold; see [Crampin and Pirani 1986, Chapter 13, Section 6].

Theorem 13. *A complete lift $\widehat{X}$ of a vector field is a symmetry of the geodesic spray Θ, i.e., commutes with the geodesic spray*

$$[\widehat{X}, T] = 0,$$

if, and only if, the vector field X is an infinitesimal affine transformation of the corresponding affine connection, and satisfies the equation, for all vector fields V, W,

$$\mathcal{L}_X \nabla_V W = \nabla_{[X,V]} W + \nabla_V \mathcal{L}_X W.$$

In the presence of a metric, the commutator of a complete lift $\widehat{X}$, of a vector field X, can be written explicitly; see [Sarbach and Zannias 2014b, Formula (74)].

Lemma C.7. *Let X be a vector field on M. The complete lift $\widehat{X}$ of X commutes with the geodesic spray Θ if, and only if,*

$$[T, \widehat{X}] = v^\alpha v^\beta [\nabla_\alpha \nabla_\beta X^\mu - R^\mu{}_{\beta\alpha\nu} X^\nu] \frac{\partial}{\partial v^\mu} = 0.$$

Remark C.8. The equation

$$\nabla_\alpha \nabla_\beta X^\mu - R^\mu{}_{\beta\alpha\nu} X^\nu = 0$$

is the equation for Jacobi fields; see [Crampin and Pirani 1986, page 340].

Geometry of the mass shell. If one considers a set of freely falling particles of given mass m, the 4-velocity of such a particle satisfies

$$g(v, v) = -m^2.$$

It is consequently natural to consider the subset of the tangent bundle TM defined by

$$P_m = \{(x, v) \in TM \mid g_x(v, v) = -m^2,\ v \text{ is future oriented}\},$$

called the *mass shell*. This set is the phase space of the considered set of particles. When the mass m is positive, P_m is a smooth submanifold of TM. When the mass m is vanishing, P_m is no longer a smooth submanifold because of the singularity at the tip of the vertex. If we ignore this fact, P_m is a fibre bundle over M. The projection over M is obtained by the restriction of the canonical projection of the bundle TM over M. The fibre at a point x is the subset of the tangent plane $T_x M$ given by

$$\{v \in T_x M \mid g_x(v, v) = -m^2,\ v \text{ is future oriented}\}.$$

Consider now a local chart (U, x^α) on M. We have seen that this local system of coordinates gives rise to a local chart on TM given by $(TU, x^\alpha, v^\alpha = dx^\alpha)$. This system of coordinates gives rise to a system of coordinates $(\bar{x}^\alpha x^\alpha, \bar{v}^i = v^i)$ on the mass shell by eliminating v^0 in the equation

$$g_{\alpha\beta} v^\alpha v^\beta = -m^2. \tag{127}$$

After one has chosen this system of coordinates, it is necessary to derive the relations between the partial derivatives in the variables (x^α, v^α), and the partial derivatives in the variables $(\bar{x}^\alpha = x^\alpha, \bar{v}^i = v^i)$. This is done by a simple application of the chain rule. Since v^0 does depend on the metric, it is first necessary to derive the following relations first: differentiating (127) gives

$$\frac{\partial v^0}{\partial x^\alpha} = -\frac{1}{2v_0} \frac{\partial g_{\beta\gamma}}{\partial x^\alpha} v^\beta v^\gamma, \tag{128}$$

$$\frac{\partial v^0}{\partial \bar{v}^i} = -\frac{v_i}{v_0}, \tag{129}$$

where we have used the notation

$$v_\alpha = g_{\alpha\beta} v^\beta.$$

Consider now a smooth function f on the tangent bundle TM. Its restriction to the mass shell is denoted by $\bar{f}$. An immediate application of the chain rule brings the following relations:

$$\frac{\partial \bar{f}}{\partial \bar{x}^\alpha} = \frac{\partial f}{\partial x^\alpha} + \frac{\partial v^0}{\partial x^\alpha} \frac{\partial f}{\partial v^0} = \frac{\partial f}{\partial x^\alpha} - \frac{1}{2v_0} \frac{\partial g_{\beta\gamma}}{\partial x^\alpha} v^\beta v^\gamma \frac{\partial f}{\partial v^0}, \tag{130}$$

$$\frac{\partial \bar{f}}{\partial \bar{v}^i} = \frac{\partial f}{\partial v^i} + \frac{\partial v^0}{\partial \bar{v}^i} \frac{\partial f}{\partial v^0} = \frac{\partial f}{\partial v^i} - \frac{v_i}{v_0} \frac{\partial f}{\partial v^0}. \tag{131}$$

The relations (130), (131) can now be used to determine which vectors are tangent to the mass shell. We notice first that the vector fields e_α, when applied to a function f, satisfy

$$\begin{aligned} e_\alpha(f) &= \frac{\partial f}{\partial x^\alpha} - v^\beta \Gamma^\gamma_{\beta\alpha} \frac{\partial f}{\partial v^\gamma} \\ &= \frac{\partial \bar{f}}{\partial \bar{x}^\alpha} - v^\beta \Gamma^i_{\beta\alpha} \frac{\partial \bar{f}}{\partial \bar{v}^i} - \left(-\frac{1}{2v_0} \frac{\partial g_{\beta\gamma}}{\partial x^\alpha} v^\beta v^\gamma + v^\beta \Gamma^0_{\beta\alpha} + \frac{v_i}{v_0} v^\beta \Gamma^i_{\beta\alpha} \right) \frac{\partial f}{\partial v^0} \\ &= \frac{\partial \bar{f}}{\partial \bar{x}^\alpha} - v^\beta \Gamma^i_{\beta\alpha} \frac{\partial \bar{f}}{\partial \bar{v}^i} + \frac{1}{2v_0} \left(\frac{\partial g_{\beta\gamma}}{\partial x^\alpha} v^\beta v^\gamma - 2 v^\beta v_\gamma \Gamma^\gamma_{\beta\alpha} \right) \frac{\partial f}{\partial v^0}. \end{aligned}$$

A quick calculation shows, using the expression of the Christoffel symbols, that

$$v^\beta v_\gamma \Gamma^\gamma_{\beta\alpha} = \frac{1}{2} v^\beta v^\gamma \left(\frac{\partial g_{\beta\gamma}}{\partial x^\alpha} + \frac{\partial g_{\alpha\gamma}}{\partial x^\beta} - \frac{\partial g_{\beta\alpha}}{\partial x^\gamma} \right) = \frac{1}{2} \frac{\partial g_{\beta\gamma}}{\partial x^\alpha} v^\beta v^\gamma.$$

The expression of e_α is consequently

$$e_\alpha(\bar{f}) = \frac{\partial \bar{f}}{\partial \bar{x}^\alpha} - v^\beta \Gamma^i_{\beta\alpha} \frac{\partial \bar{f}}{\partial \bar{v}^i} = e_\alpha(f).$$

This proves in particular that e_α is tangent to the mass shell, as well as the Liouville vector field

$$T(f) = v^\alpha e_\alpha(f) = v^\alpha e_\alpha(\bar{f}) = v^\alpha \frac{\partial \bar{f}}{\partial \bar{x}^\alpha} - v^\alpha v^\beta \Gamma^i_{\beta\alpha} \frac{\partial \bar{f}}{\partial \bar{v}^i}.$$

In dimension n, the mass shell is of dimension $2n+1$. We have as a consequence completely characterized the generators of the tangent plane to the mass shell, which is generated by the vectors

$$\left\{ e_\alpha, \frac{\partial}{\partial \bar{v}^i} \right\}.$$

From this, it is easy to deduce that the vector

$$v^\alpha \frac{\partial}{\partial v^\alpha}$$

is normal, for the Sasaki metric, to the mass shell (and also tangent in the massless case). The unit normal is consequently given by, for $m > 0$,

$$N = \frac{-1}{m^2} v^\alpha \frac{\partial}{\partial v^\alpha} = \frac{-1}{v_0} \frac{\partial}{\partial v^0} + \frac{1}{m^2} v^i \frac{\partial}{\partial \bar{v}^i}. \tag{132}$$

We will now discuss the conditions ensuring that a complete lift is tangent to the mass shell. The same procedure based on equations (130), (131) can be applied to the complete lift of a vector field X:

$$\begin{aligned} \widehat{X} &= X^\alpha e_\alpha + v^\beta \nabla_\beta X^i \frac{\partial}{\partial \bar{v}^i} + \frac{1}{v_0} v^\beta v^\gamma \nabla_\beta X_\gamma \frac{\partial}{\partial v^0} \\ &= X^\alpha e_\alpha + v^\beta \nabla_\beta X^i \frac{\partial}{\partial \bar{v}^i} + \frac{1}{v_0} \pi^{(X)}_{\beta\gamma} v^\beta v^\gamma \frac{\partial}{\partial v^0}. \end{aligned} \tag{133}$$

We immediately get the following lemma:

Lemma C.9. • *If $m > 0$, then X is Killing if, and only if, $\widehat{X}$ is tangent to P_m.*

• *If $m = 0$, then X is conformal Killing if, and only if, $\widehat{X}$ is tangent to P_0.*

Proof. The proof of this fact consists in noticing that

$$g_S(N, \widehat{X}) = v^\mu v^\nu \nabla_{(\mu} X_{\nu)}.$$

Then, if X is Killing in the massive case, or conformal Killing in the massless case,

$$g_S(N, \widehat{X}) = 0,$$

and then $\widehat{X}$ is tangent to $P_m(x)$.

Assume now that

$$g_S(N, \widehat{X}) = v^\mu v^\nu \nabla_{(\mu} X_{\nu)} = 0.$$

Consider now the symmetric 2-form $\nabla_{(\mu} X_{\nu)}$ on the vertical space, which is endowed with the metric g_{ij}. Then, in the massless case, the symmetric 2-form $\nabla_{(\mu} X_{\nu)}$ vanishes on the light cone of g_{ij} and is, as a consequence, proportional to it:

$$\nabla_{(\mu} X_{\nu)} = \phi g_{\mu\nu};$$

i.e., X is conformal Killing. The conclusion in the massive case follows in the same way. □

Acknowledgements

We would like to thank Martin Taylor for several interesting discussions on his work on the massless Einstein–Vlasov system. We would also like to thank Olivier Sarbach for his geometric introduction to Vlasov fields. J. Joudioux and J. Smulevici are partially funded by ANR-12-BS01-012-01 (AARG). J. Smulevici is also partially funded by ANR SIMI-1-003-01. D. Fajman gratefully acknowledges the travel support by ANR-12-BS01-012-01 (AARG). The comments of the referee are gratefully acknowledged.

References

[Bardos and Degond 1985] C. Bardos and P. Degond, "Global existence for the Vlasov–Poisson equation in 3 space variables with small initial data", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **2**:2 (1985), 101–118. MR Zbl

[Calogero 2003] S. Calogero, "Spherically symmetric steady states of galactic dynamics in scalar gravity", *Classical Quantum Gravity* **20**:9 (2003), 1729–1741. MR Zbl

[Calogero 2006] S. Calogero, "Global classical solutions to the 3D Nordström–Vlasov system", *Comm. Math. Phys.* **266**:2 (2006), 343–353. MR Zbl

[Calogero and Rein 2004] S. Calogero and G. Rein, "Global weak solutions to the Nordström–Vlasov system", *J. Differential Equations* **204**:2 (2004), 323–338. MR Zbl

[Christodoulou 2007] D. Christodoulou, *The formation of shocks in 3-dimensional fluids*, European Mathematical Society, Zürich, 2007. MR Zbl

[Christodoulou and Klainerman 1993] D. Christodoulou and S. Klainerman, *The global nonlinear stability of the Minkowski space*, Princeton Mathematical Series **41**, Princeton University Press, 1993. MR Zbl

[Crampin and Pirani 1986] M. Crampin and F. A. E. Pirani, *Applicable differential geometry*, London Mathematical Society Lecture Note Series **59**, Cambridge University Press, 1986. MR Zbl

[Dafermos 2006] M. Dafermos, "A note on the collapse of small data self-gravitating massless collisionless matter", *J. Hyperbolic Differ. Equ.* **3**:4 (2006), 589–598. MR Zbl

[Friedrich 2004] S. Friedrich, "Global Small Solutions of the Vlasov–Nordström System", preprint, 2004. arXiv

[Georgiev 1992] V. Georgiev, "Decay estimates for the Klein–Gordon equation", *Comm. Partial Differential Equations* **17**:7-8 (1992), 1111–1139. MR Zbl

[Hwang et al. 2011] H. J. Hwang, A. Rendall, and J. J. L. Velázquez, "Optimal gradient estimates and asymptotic behaviour for the Vlasov–Poisson system with small initial data", *Arch. Ration. Mech. Anal.* **200**:1 (2011), 313–360. MR Zbl

[Klainerman 1985a] S. Klainerman, "Global existence of small amplitude solutions to nonlinear Klein–Gordon equations in four space-time dimensions", *Comm. Pure Appl. Math.* **38**:5 (1985), 631–641. MR Zbl

[Klainerman 1985b] S. Klainerman, "Uniform decay estimates and the Lorentz invariance of the classical wave equation", *Comm. Pure Appl. Math.* **38**:3 (1985), 321–332. MR Zbl

[Klainerman 1993] S. Klainerman, "Remark on the asymptotic behavior of the Klein–Gordon equation in $\mathbb{R}^{n+1}$", *Comm. Pure Appl. Math.* **46**:2 (1993), 137–144. MR Zbl

[Klainerman and Nicolò 2003] S. Klainerman and F. Nicolò, *The evolution problem in general relativity*, Progress in Mathematical Physics **25**, Birkhäuser, Boston, 2003. MR Zbl

[LeFloch and Ma 2016] P. G. LeFloch and Y. Ma, "The global nonlinear stability of Minkowski space for self-gravitating massive fields", *Comm. Math. Phys.* **346**:2 (2016), 603–665. MR Zbl

[Lieb and Loss 1997] E. H. Lieb and M. Loss, *Analysis*, Graduate Studies in Mathematics **14**, American Mathematical Society, Providence, RI, 1997. MR

[Lindblad and Rodnianski 2010] H. Lindblad and I. Rodnianski, "The global stability of Minkowski space-time in harmonic gauge", *Ann. of Math.* (2) **171**:3 (2010), 1401–1477. MR Zbl

[Morawetz 1962] C. S. Morawetz, "The limiting amplitude principle", *Comm. Pure Appl. Math.* **15** (1962), 349–361. MR Zbl

[Morawetz 1968] C. S. Morawetz, "Time decay for the nonlinear Klein–Gordon equations", *Proc. Roy. Soc. Ser. A* **306** (1968), 291–296. MR Zbl

[Pallard 2006] C. Pallard, "On global smooth solutions to the 3D Vlasov–Nordström system", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **23**:1 (2006), 85–96. MR Zbl

[Rein and Rendall 1992] G. Rein and A. D. Rendall, "Global existence of solutions of the spherically symmetric Vlasov–Einstein system with small initial data", *Comm. Math. Phys.* **150**:3 (1992), 561–583. Correction in **176**:2 (1996), 475–478. MR Zbl

[Ringström 2013] H. Ringström, *On the topology and future stability of the universe*, Oxford University Press, 2013. MR Zbl

[Sarbach and Zannias 2014a] O. Sarbach and T. Zannias, "The geometry of the tangent bundle and the relativistic kinetic theory of gases", *Classical Quantum Gravity* **31**:8 (2014), art. id. 085013. Zbl

[Sarbach and Zannias 2014b] O. Sarbach and T. Zannias, "Tangent bundle formulation of a charged gas", pp. 192–207 in *Recent developments on physics in strong gravitational fields, V: Leopoldo García-Colín Mexican Meeting on Mathematical and Experimental Physics* (México City, 2013), edited by A. Macías and M. Maceda, American Institute of Physics Conference Series **1577**, AIP Publishing, Melville, NY, 2014.

[Schaeffer 2004] J. Schaeffer, "A small data theorem for collisionless plasma that includes high velocity particles", *Indiana Univ. Math. J.* **53**:1 (2004), 1–34. MR Zbl

[Smulevici 2016] J. Smulevici, "Small data solutions of the Vlasov–Poisson system and the vector field method", *Ann. PDE* **2**:2 (2016), art. id. 11. MR

[Taylor 2017] M. Taylor, "The global nonlinear stability of Minkowski space for the massless Einstein–Vlasov system", *Ann. PDE* **3**:1 (2017), art. id. 9. MR

[Villani 2010] C. Villani, "Landau damping", lecture notes for a course given in Cotonou, Benin and in CIRM, 2010, available at http://cedricvillani.org/wp-content/uploads/2012/08/B13.Landau.pdf.

[Wang 2015a] Q. Wang, "Global stability of Minkowski space for massive scalar fields", talk given at the conference "General Relativity: A celebration of the 100th anniversary", Institut Henri Poincaré, November 2015, available at https://philippelefloch.files.wordpress.com/2015/11/2015-ihp-qianwang.pdf. Video available at https://www.youtube.com/watch?v=-7nTHUPzVL0&feature=youtu.be.

[Wang 2015b] Q. Wang, "Lectures on nonlinear wave equations", lecture notes, University of Oxford, 2015, available at http://people.maths.ox.ac.uk/wangq1/Lecture_notes/nonlinear_wave_9.pdf.

Received 28 Apr 2016. Revised 13 Apr 2017. Accepted 9 May 2017.

DAVID FAJMAN: david.fajman@univie.ac.at
Gravitational Physics, Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria

JÉRÉMIE JOUDIOUX: jeremie.joudioux@univie.ac.at
Gravitational Physics, Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria

JACQUES SMULEVICI: jacques.smulevici@math.u-psud.fr
Laboratoire de Mathématiques, Université Paris-Sud 11, Bât. 425, 91405 Orsay Cedex, France

ANALYSIS AND PDE
Vol. 10, No. 7, 2017
dx.doi.org/10.2140/apde.2017.10.1613

ANALYTIC HYPOELLIPTICITY FOR SUMS OF SQUARES AND THE TREVES CONJECTURE, II

ANTONIO BOVE AND MARCO MUGHETTI

We are concerned with the problem of real analytic regularity of the solutions of sums of squares with real analytic coefficients. The Treves conjecture defines a stratification and states that an operator of this type is analytic hypoelliptic if and only if all the strata in the stratification are symplectic manifolds.

Albano, Bove, and Mughetti (2016) produced an example where the operator has a single symplectic stratum, according to the conjecture, but is not analytic hypoelliptic.

If the characteristic manifold has codimension 2 and if it consists of a single symplectic stratum, defined again according to the conjecture, it has been shown that the operator is analytic hypoelliptic.

We show here that the above assertion is true only if the stratum is single, by producing an example with two symplectic strata which is not analytic hypoelliptic.

1. Introduction

The purpose of the paper is to discuss the real analytic regularity of the distribution solutions to sums of squares equations

$$P(x,D)u=\sum_{j=1}^{N}X_j(x,D)^2u=f, \tag{1-1}$$

where $X_j(x,D)$ denote vector fields with real analytic coefficients defined in an open set $\Omega\subset\mathbb{R}^n$, u is a distribution in Ω and $f\in C^\omega(\Omega)$, the space of all real analytic functions in Ω.

We suppose that the vector fields verify Hörmander's condition

(H) The Lie algebra generated by the vector fields and their commutators has dimension n, equal to the dimension of the ambient space.

In 1996 F. Treves [1999], see also [Bove and Treves 2004] for a formulation closer to the following, as well as [Treves 2006] for variants, stated a conjecture for the sums of squares of vector fields to be analytic hypoelliptic. In this paper we give neither the motivations nor the details about its statement; for both the motivations and a short introduction to the conjecture, as well as a brief review of the existing literature, we refer to [Albano et al. 2016].

Let us first give a very sketchy idea of how the conjecture was formulated. The main concept it uses is a stratification of the characteristic variety. This is a partition of the set $\{(x,\xi)\mid X_j(x,\xi)=0,\ j=1,\dots,N\}$ into real analytic manifolds as follows.

MSC2010: primary 35H10, 35H20; secondary 35B65, 35A20, 35A27.
Keywords: sums of squares of vector fields, analytic hypoellipticity, Treves conjecture.

Let P be as in (1-1). Then the characteristic variety of P is

$$\operatorname{Char}(P) = \{(x, \xi) \mid X_j(x, \xi) = 0,\ j = 1, \dots, N\},$$

where $X_j(x, \xi)$ denotes the symbol of the j-th vector field. This is a real analytic variety and, as such, it can be stratified, locally, in real analytic manifolds Σ_i for i in a finite family of indices $\mathscr{I}$. This means that

$$\operatorname{Char}(P) = \bigcup_{i \in \mathscr{I}} \Sigma_i,$$

and the Σ_i have the property that for $i \neq i'$, we have $\Sigma_i \neq \Sigma_{i'}$ and either $\Sigma_i \cap \overline{\Sigma}_{i'} = \varnothing$ or, if $\Sigma_i \cap \overline{\Sigma}_{i'} \neq \varnothing$, then $\Sigma_i \subset \partial \Sigma_{i'}$ (the boundary of $\Sigma_{i'}$). We refer to [Treves $\geq$ 2017] for more details.

Next we examine the rank of the restriction of the symplectic form, that is, of the form $\sigma = \sum_{j=1}^n d\xi_j \wedge dx_j$, to the strata Σ_i, meaning that at any point $\rho \in \Sigma_i$ in a certain fixed neighborhood of $\rho_0 \in \operatorname{Char}(P)$, we restrict σ to the tangent space to Σ_i at ρ, denoted by $T_\rho \Sigma_i$. We want σ to have constant rank on each stratum Σ_i.

If this is not the case, we may consider the analytic variety where there is a change of rank, since the symplectic form restricted to Σ_i has a matrix whose entries are the Poisson brackets of the defining functions of Σ_i. Hence the rank is not maximal on a closed analytic subvariety where the determinant of a maximal minor vanishes. We may start over the procedure described above and further stratify this subvariety. The procedure ends after a finite number of steps yielding a stratification of $\operatorname{Char}(P)$ with real analytic manifolds where the restriction of the symplectic form has constant rank.

In the final step one considers the multiple Poisson brackets of the symbols of the vector fields. Let $I = (i_1, i_2, \dots, i_r)$, where $i_j \in \{1, \dots, N\}$. Write $|I| = r$ and define

$$X_I(x, \xi) = \{X_{i_1}(x, \xi), \{X_{i_2}(x, \xi), \{\cdots \{X_{i_{r-1}}(x, \xi), X_{i_r}(x, \xi)\} \cdots\}\}\}.$$

Here r is called the length of the multiple Poisson bracket $X_I(x, \xi)$. We recall that the Poisson bracket is defined as

$$\{X_i(x, \xi), X_j(x, \xi)\} = \sum_{\ell=1}^n \left(\frac{\partial X_i}{\partial \xi_\ell} \frac{\partial X_j}{\partial x_\ell} - \frac{\partial X_i}{\partial x_\ell} \frac{\partial X_j}{\partial \xi_\ell} \right).$$

We recall that L. Hörmander[1967] solved the problem of the C^∞ hypoellipticity of sums of squares formulating his well known condition using the algebra built with the Poisson brackets of the vector fields.

It is clear that since all the strata defined above are submanifolds of the characteristic variety, the symbols of the vector fields vanish on each stratum.

Next we examine all the Poisson brackets of two (symbols of) vector fields on a stratum in a neighborhood U of a fixed point ρ_0. Denote again by Σ_i the stratum. A few things may happen: there is at least a nonzero Poisson bracket at ρ_0 and hence, possibly shrinking U, on all of it, in which case we stop. Otherwise all brackets may vanish identically on $\Sigma_i \cap U$. Finally there may be Poisson brackets that vanish on a subvariety of $\Sigma_i \cap U$.

At this point we repeat the stratification construction above using the equations defining this subvariety. Then we pick a new stratum, say Σ_j, and we have that either there is a nonzero Poisson bracket on $\Sigma_j \cap U$ or every Poisson bracket identically vanishes in $\Sigma_j \cap U$.

This procedure may be iterated by considering Poisson brackets of length 3 etc.

In the end, after a finite number of steps, we wind up with a stratification such that every stratum Σ has the following properties:

(a) Σ is a real analytic submanifold of $\operatorname{Char}(P)$.

(b) The symplectic form restricted to Σ has constant rank.

(c) Denoting again by U a neighborhood of the point $\rho_0 \in \operatorname{Char}(P)$, there is an index say $m = m(\Sigma)$ such that at least one bracket of length m is nonzero on $U \cap \Sigma$ and all the brackets of length less than m identically vanish on $\Sigma \cap U$.

Conjecture 1.1 [Treves 1999]. *Consider the operator P in (1-1) and define the stratification as sketched above. Then P is analytic hypoelliptic if and only if every stratum in the stratification is symplectic.*

Remark 1.2. The above conjecture is subtle and depends essentially on two ingredients: first the way the stratification is defined and, second, the fact that the strata must be symplectic manifolds. The condition that the strata, whatever that might mean, should be symplectic seems quite reasonable, since there are examples of operators with a nonsymplectic characteristic manifold which are known not to be analytic hypoelliptic. A different problem is the definition of the strata. It seems that, because of [Albano et al. 2016] and the present result, the way of defining the strata has to be changed in the statement of the conjecture. The authors have no solution to this problem right now.

The necessary part of the conjecture, i.e., the nonanalytic hypoellipticity in the presence of nonsymplectic strata, is, as far as we know, still an open problem, although it might be of limited interest if the definition of the stratification is changed.

In [Albano et al. 2016] it was shown that the sufficient part of the above conjecture is false by exhibiting an operator with a single symplectic stratum, defined according to Conjecture 1.1, of dimension 4 (and codimension 4) and proving that the operator is not analytic hypoelliptic. Actually its Gevrey regularity has been completely characterized.

It is not difficult to exhibit, based on [Albano et al. 2016], examples of sums of squares having a single symplectic stratum Σ defined according to Conjecture 1.1, and such that $\operatorname{codim} \Sigma = 2\nu$, $2 \le \nu \le n-2$, for which a proof analogous to that of the same paper implies the nonanalytic hypoellipticity of the operator. Here is an example.

Let us define $x' = (x_1, \dots, x_\nu)$ and $x'' = (x_{\nu+1}, \dots, x_n)$, so that $x = (x', x'')$, $x \in \mathbb{R}^n$, where ν satisfies the above conditions. Define $x'_\nu = (x_1, \dots, x_{\nu-1})$, $x''_\nu = (x_{\nu+2}, \dots, x_n)$ and set

$$|D_{x'_\nu}|^2 + [x'_\nu]_r^2 |D_{x''}|^2 + D_{x_\nu}^2 + x_\nu^{2(p-1)} D_{x_{\nu+1}}^2 + x_\nu^{2(q-1)} |D_{x''_\nu}|^2,$$

where $[x'_\nu]_r^2 = \sum_{j=1}^{\nu-1} x_j^{2(r_j-1)}$, $|D_{x''}|^2 = \sum_{j=\nu+1}^{n} D_{x_j}^2$, and analogously for $|D_{x'_\nu}|^2$, $|D_{x''_\nu}|^2$, with the condition that $1 < \min r_j \le \max r_j < p < q$.

A number of papers have been written both in the case of a single stratum and of a more complex stratification. The meaning of the word "stratum" is henceforth that defined in Conjecture 1.1. If we are in the presence of a single symplectic stratum of codimension 2, then the conjecture has been proved true by Métivier [1981], Ōkaji [1985], Cordaro and Hanges [2009], and Albano and Bove [2013]. The papers by Métivier, Ōkaji and Albano and Bove include also a higher codimension single symplectic stratum, provided additional assumptions are satisfied.

We say that an operator exhibits nested strata, always according to Conjecture 1.1, if the associated stratification has at least two strata Σ_1, Σ_2 such that $\Sigma_1 \neq \Sigma_2$ and $\Sigma_1 \cap \bar{\Sigma}_2 \neq \varnothing$. By definition this implies that $\Sigma_1 \subset \partial\Sigma_2$ and, in particular, that the dimension of Σ_1 is smaller than that of Σ_2.

Theorem 2.1 proved below implies that the conjecture does not hold, in general, when there are several symplectic nested Poisson–Treves strata.

The conjecture fails even if the characteristic manifold has codimension 2, as in the case we are going to examine. This is actually proved in the remainder of the present paper.

Thus we can state:

Theorem 1.3. *Let us consider the class of all sums of squares with analytic coefficients such that the associated stratification near a point $\rho_0 \in \operatorname{Char}(P)$ has not a single stratum. Then the sufficient part of Conjecture 1.1 is false even in the case of a characteristic manifold of codimension* 2.

We remark that if the characteristic variety is a real analytic manifold of dimension 2, and the Treves strata are symplectic, then we may have only a single symplectic stratum of Treves type — obviously coinciding with the characteristic manifold.

If the characteristic variety is a manifold of codimension 2 as well as of dimension 2, by the results quoted above one deduces that the operator is analytic hypoelliptic. If, on the other hand, the codimension of the characteristic manifold is larger than 2, we do not think that analytic hypoellipticity ensues and thus Conjecture 1.1 would be contradicted.

To clarify the above sentence let us consider the operator

$$Q(x,D) = D_1^2 + D_2^2 + x_1^2 D_3^2 + x_2^4 D_3^2 + x_2^2 x_3^2 D_3^2.$$

It is easily seen that $\operatorname{Char}(Q) = \{(x,\xi) \mid x_i = \xi_i = 0,\ i = 1,2,\ \xi_3 \neq 0\}$, which has dimension 2 and codimension 4. $\operatorname{Char}(Q)$ is a single symplectic stratum for Q and it is not too difficult to prove, either by using the subelliptic estimate (see Section 4 below and formula (4-2)) or the method described in [Bove and Mughetti 2016], that Q is Gevrey $\frac{4}{3}$-hypoelliptic (see Definition 2.2 for a definition of the Gevrey classes).

We think that the Gevrey $\frac{4}{3}$-regularity is optimal for the operator Q, but the proof of optimality is however an open problem. The difficulty of the proof is due to the following fact: in the examples of [Albano et al. 2016] and (2-1) of this paper, as well as in Q, there are strata which do not appear in Treves stratification. The Hamilton bicharacteristics associated to these phantom strata either project injectively on the base space, like in the case of [Albano et al. 2016] and (2-1), or project injectively onto the fibers of the cotangent bundle. The latter is the case for Q. In this sense the operator Q shares this difficulty with the Métivier operator, [1981] where optimality is very hard to prove.

We would also like to recall that due to [Bove et al. 2013] the conjecture does not hold for sums of squares of complex vector fields.

Here is the structure of the paper. The result is stated in Section 2. Section 3 is devoted to the proof of the optimality of the s_ℓ Gevrey regularity. We construct a solution to $Pu = g$, where g is real analytic, which is not better than Gevrey s_ℓ. In Section 4 we prove that every solution to $Pu = f$ is Gevrey s_ℓ, if $f \in G^{s_\ell}$. This is done using the subelliptic estimate for the operator.

2. Statement of the result

Let $\ell, r, p, q \in \mathbb{N}$, $1 < r < p < q$, and $x = (x_1, \dots, x_4) \in \mathbb{R}^4$. The objective of this section is to state the optimal Gevrey regularity for the operator

$$P(x, D) = D_1^2 + x_1^{2(\ell+r-1)}(D_3^2 + D_4^2) + x_1^{2\ell}[D_2^2 + x_2^{2(p-1)} D_3^2 + x_2^{2(q-1)} D_4^2]. \tag{2-1}$$

Hörmander's condition is satisfied by P and thus P is C^∞ hypoelliptic.

The characteristic manifold of P is the real analytic manifold

$$\mathrm{Char}(P) = \{(x, \xi) \in T^*\mathbb{R}^4 \setminus \{0\} \mid \xi_1 = 0,\ x_1 = 0,\ \xi_2^2 + \xi_3^2 + \xi_4^2 > 0\}. \tag{2-2}$$

According to Treves' conjecture one has to look at the strata associated with P.

The stratification associated with P is made up of two symplectic strata Σ_1 and Σ_2:

$$\Sigma_1 = \{(0, x_2, x_3, x_4; 0, \xi_2, \xi_3, \xi_4) \mid \xi_2^2 + x_2^2 > 0,\ \xi_2^2 + \xi_3^2 + \xi_4^2 > 0\} \tag{α}$$

at depth $\ell + 1$.[1] Σ_1 is a symplectic stratum and the restriction of the symplectic form to it has rank 6.

$$\Sigma_2 = \{(0, 0, x_3, x_4; 0, 0, \xi_3, \xi_4) \mid \xi_3^2 + \xi_4^2 > 0\} \tag{β}$$

at depth $\ell + r$. This is also a symplectic stratum and the restriction of the symplectic form to it has rank 4.

We point out that the above stratification does not depend on the choice of the indices p and q.

According to the conjecture we would expect local real analyticity near the origin for the distribution solutions u of $Pu = f$, with a real analytic right-hand side.

We are ready to state the theorem that is proved in the next two sections of the paper.

Theorem 2.1. *Let*

$$\frac{1}{s_\ell} = \frac{\ell+1}{\ell+r} + \frac{r-1}{\ell+r}\,\frac{p-1}{q-1}.$$

Then P is locally Gevrey s_ℓ-hypoelliptic and not better near the origin.

We recall here the definition of the Gevrey classes:

Definition 2.2. If Ω is an open subset of $\mathbb{R}^n$ and $s \geq 1$ we denote by $G^s(\Omega)$ the class of all functions $u \in C^\infty(\Omega)$ such that for every compact set $K \Subset \Omega$ there is a positive constant $C = C_K$ such that

$$|\partial_x^\alpha u(x)| \leq C^{|\alpha|+1} \alpha!^s \quad \text{for every } x \in K.$$

[1] This means that all the Poisson brackets of the (symbols of) the vector fields of length less than $\ell + 1$ are identically zero on the characteristic manifold and that there is at least one bracket of length $\ell + 1$ which is nonzero.

We observe that $G^1(\Omega)$ coincides with the class of all real analytic functions in Ω.

As a consequence of Theorem 2.1 we have:

Corollary 2.3. *The operator P is analytic hypoelliptic if and only if $p = q$.*

The proof of the corollary is contained in Section 4.

Moreover from Theorem 2.1 we deduce that Theorem 1.3 holds since the operator P above has a stratification made of two nested symplectic strata.

Remark 2.4. The geometric interpretation of the above result is not known. We believe that a different definition of the associated stratification should be given, allowing the existence of an additional stratum for P.

More precisely the missing stratum seems to be

$$\widetilde{\Sigma} = \{(0, 0, x_3, x_4; 0, 0, 0, \xi_4) \mid \xi_4 \neq 0\},$$

which can be seen as the set where ξ_3 vanishes in Σ_2.

$\widetilde{\Sigma}$ is not symplectic; its Hamilton curves are the x_3-lines and this fact gives us a lead as to why, in the following proof of Theorem 2.1, we may conclude that the operator P is not analytic hypoelliptic. We shall come back on this further on.

We would also like to observe that the point $(0, e_4)$ is the only interesting characteristic point where we have a lack of analytic hypoellipticity. In fact the operator P is microlocally analytic hypoelliptic at all points in $\Sigma_1 \setminus \Sigma_2$, as well as at points in Σ_2 where $\xi_3 \neq 0$. This can be proved via L^2 (microlocalized) estimates of the type of (but easier than those) used in Section 4.

It is also worth noting that if we accept, without any proof or other justification, that the stratification associated to P is made of (the connected components of) $\Sigma_1 \setminus \Sigma_2$, $\Sigma_2 \setminus \widetilde{\Sigma}$ and $\widetilde{\Sigma}$, where $\widetilde{\Sigma}$ is given above, then all points of the first two "strata" are points of analytic hypoellipticity and the strata are symplectic, while the nonanalytic hypoellipticity comes in at points of the nonsymplectic stratum $\widetilde{\Sigma}$.[2]

3. Proof of Theorem 2.1

In this section we prove the optimality of the Gevrey regularity in Theorem 2.1.

We construct a solution to the equation $Pu = f$, for a real analytic function f, which is not Gevrey s for any $s < s_\ell$ and is defined in a neighborhood of the origin.

In fact we look for a function $u(x, y, t)$ defined in $\widetilde{U} \times [1, +\infty[\subset \mathbb{R}_x \times \mathbb{R}_y \times [1, +\infty[$, where $\widetilde{U}$ denotes a neighborhood of the origin in $\mathbb{R}^2_{x,y}$, and such that

$$P(x, D)A(u) = g, \tag{3-1}$$

where

$$A(u)(x) = \int_1^{+\infty} e^{-i\rho x_4 + x_3 z(\rho)\rho^\theta - \rho^\theta} u(\rho^{\frac{1}{\ell+r}} x_1, x_2, \rho)\, d\rho, \tag{3-2}$$

and

$$\theta = \frac{1}{s_\ell}.$$

[2]We are indebted to the referees for pointing out to us the need of such a remark.

The function $z(\rho)$ is to be determined. Here we assume that $x \in U$, a suitable neighborhood of the origin in $\mathbb{R}^4$ whose size will ultimately depend on $z(\rho)$. Furthermore $g \in C^\omega(U)$.

We have

$$P(x,D)A(u)(x)=\int_1^{+\infty} e^{-i\rho x_4+x_3 z(\rho)\rho^\theta-\rho^\theta}\Big[-\rho^{\frac{2}{\ell+r}}\partial_{x_1}^2 u-x_1^{2(\ell+r-1)}(z(\rho))^2\rho^{2\theta}u+x_1^{2(\ell+r-1)}\rho^2 u \\ +x_1^{2\ell}\big(-\partial_{x_2}^2 u-x_2^{2(p-1)}(z(\rho))^2\rho^{2\theta}u+x_2^{2(q-1)}\rho^2 u\big)\Big]\,d\rho.$$

Rewriting the right-hand side of the above relation in terms of the variable $y_1=\rho^{\frac{1}{\ell+r}}x_1$, we obtain

$$P(x,D)A(u)(x)=\int_1^{+\infty} e^{-i\rho x_4+x_3 z(\rho)\rho^\theta-\rho^\theta} \\ \times\Big[-\rho^{\frac{2}{\ell+r}}\big(\partial_1^2-y_1^{2(\ell+r-1)}(1-(z(\rho))^2\rho^{2(\theta-1)})\big)u \\ +\rho^{-\frac{2\ell}{\ell+r}}y_1^{2\ell}\big[\big(-\partial_2^2-x_2^{2(p-1)}(z(\rho))^2\rho^{2\theta}+x_2^{2(q-1)}\rho^2\big)\big]u\Big]_{y_1=\rho^{1/(\ell+r)}x_1}\,d\rho.$$

We point out that

$$\theta-1<0.$$

Choose

$$u(y_1,x_2,\rho)=u_1(y_1,\rho)u_2(x_2,\rho), \tag{3-3}$$

where $u_j,\ j=1,2$, will be chosen later. Plugging this into the above formula yields

$$P(x,D)A(u)(x) \\ =\int_1^{+\infty} e^{-i\rho x_4+x_3 z(\rho)\rho^\theta-\rho^\theta} \\ \times\Big[-\rho^{\frac{2}{\ell+r}}u_2(x_2,\rho)\big(\partial_1^2-y_1^{2(\ell+r-1)}(1-(z(\rho))^2\rho^{2(\theta-1)})\big)u_1(y_1,\rho) \\ +\rho^{-\frac{2\ell}{\ell+r}}y_1^{2\ell}u_1(y_1,\rho)\big[\big(-\partial_2^2-x_2^{2(p-1)}(z(\rho))^2\rho^{2\theta}+x_2^{2(q-1)}\rho^2\big)\big]u_2(x_2,\rho)\Big]_{y_1=\rho^{1/(\ell+r)}x_1}\,d\rho. \tag{3-4}$$

We want to determine u_1, u_2 so that $P(x,D)A(u)(x)=0$. In particular

$$-\rho^{\frac{2}{\ell+r}}u_2(x_2,\rho)\big(\partial_1^2-y_1^{2(\ell+r-1)}(1-(z(\rho))^2\rho^{2(\theta-1)})\big)u_1(y_1,\rho) \\ +\rho^{-\frac{2\ell}{\ell+r}}y_1^{2\ell}u_1(y_1,\rho)\big[\big(-\partial_2^2-x_2^{2(p-1)}(z(\rho))^2\rho^{2\theta}+x_2^{2(q-1)}\rho^2\big)\big]u_2(x_2,\rho)=0 \tag{3-5}$$

for ρ large.

Let us start by considering the operator in the x_2-variable:

$$\rho^{2\frac{\ell+1}{\ell+r}}L(x_2,\partial_{x_2})=-\partial_2^2-x_2^{2(p-1)}(z(\rho))^2\rho^{2\theta}+x_2^{2(q-1)}\rho^2.$$

Performing the dilation

$$x_2=y_2\rho^{-\mu},$$

where

$$\mu=\frac{r-1}{\ell+r}\frac{1}{q-1}, \tag{3-6}$$

we obtain

$$L(y_2, \partial_{y_2}) = -\rho^{2\left(\frac{r-1}{\ell+r}\frac{1}{q-1}-\frac{\ell+1}{\ell+r}\right)}\partial_{y_2}^2 - z(\rho)^2 y_2^{2(p-1)} + y_2^{2(q-1)}. \tag{3-7}$$

Set

$$h = \rho^{\frac{r-1}{\ell+r}\frac{1}{q-1}-\frac{\ell+1}{\ell+r}} = \rho^{-\varkappa}. \tag{3-8}$$

Observe that the exponent above is negative, since $r-1 < (\ell+1)(q-1)$ if and only if $r-q < \ell(q-1)$, which is obviously true for any value of ℓ as $r < q$ by assumption. Thus for large ρ, we know h tends to zero and hence we have to study the semiclassical stationary Schrödinger operator

$$L(y_2, \partial_{y_2}) = -h^2\partial_{y_2}^2 - z(h)^2 y_2^{2(p-1)} + y_2^{2(q-1)} \tag{3-9}$$

exhibiting a double-well potential. We point out that the dilation (3-6) has been chosen in such a way to get rid of the parameter ρ, i.e., h, from the double-well potential.

Let us make the following ansatz: the quantity $z(h)$ above is positive and such that there is an $h_0 > 0$ for which

$$0 < \inf_{0<h<h_0} z(h) < +\infty.$$

We shall return to this ansatz and show that it is actually compatible with our findings.

We may further dilate the operator in (3-9) in such a way that the quantity $z(h)$ appears as a coefficient of the second derivative, modulo a multiplying factor. Set

$$y_2 = (z(h))^{\frac{1}{q-p}}y.$$

Note that the above dilation is well defined because of our ansatz when $p < q$. If, on the other hand $p = q$ the whole construction is not needed, since $s_\ell = 1$ then.

Thus (3-9) becomes

$$L(y, \partial_y) = z^{2\frac{q-1}{q-p}}\left[-(z^{-\frac{q}{q-p}}h)^2\partial_y^2 - y^{2(p-1)} + y^{2(q-1)}\right]. \tag{3-10}$$

Let

$$\hat{\gamma} = -\frac{q-p}{q-1}\left(\frac{p-1}{q-1}\right)^{\frac{p-1}{q-p}} < 0$$

denote the minimum of the potential $-y^{2(p-1)} + y^{2(q-1)}$. Then

$$-(z^{-\frac{q}{q-p}}h)^2\partial_y^2 - y^{2(p-1)} + y^{2(q-1)} - \hat{\gamma} \tag{3-11}$$

has a discrete spectrum made of simple positive eigenvalues accumulating at infinity; see, e.g., [Berezin and Shubin 1991]. Hence the eigenvalues are real analytic functions of the parameter $hz^{-\frac{q}{q-p}} > 0$.

At this point we might choose to select the ground state of (3-11). This would allow us to treat only the case of an even eigenfunction with well-known Agmon estimates. However, we would like to emphasize the fact that the Gevrey regularity we find is a consequence of the nature of the spectrum of the operator in (3-11) and that, in particular, any eigenvalue allows us to conclude the Gevrey regularity of the solution to (3-1).

There is a price to pay for this generality: we cannot a priori use the fact that the associated eigenfunction is symmetric and positive — which is true for the fundamental eigenstate.

Denote by $E = E(h(z(h))^{-\frac{q}{q-p}})$ one of the energy levels of (3-11). Also let $u_2(y,h)$ denote the corresponding eigenfunction, i.e.,

$$\left[-(z^{-\frac{q}{q-p}}h)^2\partial_y^2 - y^{2(p-1)} + y^{2(q-1)} - \hat{\gamma}\right]u_2 = Eu_2, \tag{3-12}$$

or, changing back the variables

$$\left[-\partial_{x_2}^2 - z^2\rho^{2\theta}x_2^{2(p-1)} + x_2^{2(q-1)}\rho^2 - \hat{\gamma}z^{2\frac{q-1}{q-p}}\rho^{2\frac{\ell+1}{\ell+r}}\right]u_2 = Ez^{2\frac{q-1}{q-p}}\rho^{2\frac{\ell+1}{\ell+r}}u_2. \tag{3-13}$$

The operator in (3-11) has a symmetric nonnegative double-well potential with two nondegenerate minima and which is unbounded at infinity. Theorem 1.1 in [Simon 1983] asserts:

Theorem 3.1. *For every eigenvalue $E(\mu)$ in the spectrum of*

$$-\mu^2\partial_y^2 - y^{2(p-1)} + y^{2(q-1)} - \hat{\gamma},$$

we have

$$\lim_{\mu\to 0+}\frac{E(\mu)}{\mu} = e^* > 0. \tag{3-14}$$

As a consequence we may continue the function $E(\mu)$ at zero by setting $E(0) = 0$. Thus $E(\mu)$ is differentiable for $0 \le \mu$.

Furthermore we have the following.

Lemma 3.2. *For every $h_0 > 0$, we have that $\partial_h E(h)$ exists and is bounded for $0 \le h \le h_0$.*

Proof. This is basically proved by deriving in this case the Feynman–Hellmann formula expressing the derivative with respect to h of the eigenvalues in terms of the associated eigenfunctions and the derivative of the Hamiltonian; see, e.g., the proof of Lemma 3.1 in [Albano et al. 2016]. □

Let us now go back to equation (3-5). Neglecting the factor $\rho^{\frac{2}{\ell+r}}$ and writing everything as a function of h, we obtain

$$\left[\partial_1^2 - y_1^{2(\ell+r-1)}(1-(z(h))^2h^{2\frac{1-\theta}{\varkappa}})\right]u_1(y_1,h) - y_1^{2\ell}z(h)^{2\frac{q-1}{q-p}}(E(hz(h)^{-\frac{q}{q-p}}) + \hat{\gamma})u_1(y_1,h) = 0, \tag{3-15}$$

where, as specified above, E is an eigenvalue of the operator (3-9) and h is small.

We want to show that, for small h, we can find a bounded positive function $z = z(h)$ such that (3-15) has a nontrivial kernel, which will be made of rapidly decreasing eigenfunctions corresponding to the eigenvalue zero.

First set

$$\tau(h) = 1 - (z(h))^2h^{2\frac{1-\theta}{\varkappa}}.$$

Note that, due to our ansatz, τ is a positive number if h is suitably small. We are thus entitled to perform the dilation

$$t_1 = y_1\tau^{\frac{1}{2(\ell+r)}},$$

so that the differential operator in (3-15) becomes

$$\tau^{\frac{1}{\ell+r}}\big[-\partial_{t_1}^2+t_1^{2(\ell+r-1)}+t_1^{2\ell}z(h)^{2\frac{q-1}{q-p}}\tau^{-\frac{\ell+1}{\ell+r}}\big(E(hz(h)^{-\frac{q}{q-p}})+\hat{\gamma}\big)\big]. \tag{3-16}$$

Proposition 3.3. *There exists a positive number σ such that the operator*

$$-\partial_t^2+t^{2(\ell+r-1)}-\sigma t^{2\ell} \tag{3-17}$$

has a nontrivial kernel.

Proof. The proof is just an analysis of the behavior of the eigenvalues of the operator as functions of the parameter σ; see, for instance, [Mughetti 2014; 2015]. First of all we remark that the function $t\mapsto t^{2(\ell+r-1)}-\sigma t^{2\ell}$ has a (negative) minimum

$$-\frac{r-1}{\ell+r-1}\Big(\frac{\ell}{\ell+r-1}\Big)^{\frac{\ell}{r-1}}\sigma^{\frac{\ell+r-1}{r-1}}=\sigma^{\frac{\ell+r-1}{r-1}}\hat{\mu}.$$

Performing the dilation

$$t=t_1\sigma^{\frac{1}{2(r-1)}},$$

the operator becomes

$$-\sigma^{\frac{-1}{r-1}}\partial_{t_1}^2+\sigma^{\frac{\ell+r-1}{r-1}}(t_1^{2(\ell+r-1)}-t_1^{2\ell}),$$

which can be written as

$$\sigma^{\frac{\ell+r-1}{r-1}}\big[(-\sigma^{-\frac{\ell+r}{r-1}}\partial_{t_1}^2+t_1^{2(\ell+r-1)}-t_1^{2\ell}-\hat{\mu})+\hat{\mu}\big].$$

Now the operator in parentheses is again a Schrödinger operator with a double-well positive potential and hence it has a positive discrete spectrum accumulating at infinity, by [Berezin and Shubin 1991].

Denote by $\lambda(\sigma)$ one such eigenvalue, so that the eigenvalues of the operator above are

$$\sigma^{\frac{\ell+r-1}{r-1}}(\lambda(\sigma)+\hat{\mu}).$$

If $\sigma\to+\infty$, then, by Theorem 3.1, $\lambda(\sigma)\to 0+$, so that the expression above is negative. On the other hand it is obvious that for $\sigma=0$ the eigenvalues of the operator (3-17) are positive. Furthermore they are simple, whatever the value of σ is, and thus they are also continuously dependent on σ. We conclude hence that there is a value $\bar{\sigma}$ of σ for which $\lambda(\bar{\sigma})=0$. This proves the nontriviality of the kernel. □

Going back to (3-16), we see that, in order to solve (3-15) it is enough to choose one such value σ given by Proposition 3.3 and then solve, with respect to z, the equation

$$-\bar{\sigma}=z(h)^{2\frac{q-1}{q-p}}\tau^{-\frac{\ell+1}{\ell+r}}\big(E(hz(h)^{-\frac{q}{q-p}})+\hat{\gamma}\big). \tag{3-18}$$

Set

$$\tilde{z}=\Big(\frac{-\bar{\sigma}}{\hat{\gamma}}\Big)^{\frac{q-p}{2(q-1)}}>0.$$

Observe that the values $h=0,\ z(0)=\tilde{z}$ verify equation (3-18), since $\tau(0)=1$.

Proposition 3.4. *There is an $h_0 > 0$ such that (3-18) implicitly defines a function $z \in C([0, h_0[) \cap C^\omega(]0, h_0[)$. Moreover $\lim_{h\to 0+} z(h) = \tilde{z}$.*

Proof. The proof is analogous to the proof of Proposition 3.1 in [Albano et al. 2016] and we just sketch it.

Consider the function

$$f(h,z) = z^{2\frac{q-1}{q-p}} \tau^{-\frac{\ell+1}{\ell+r}} (E(hz^{-\frac{q}{q-p}}) + \hat{\gamma}) + \bar{\sigma}.$$

Let us compute the derivative with respect to z of the above function in the interval $[0, h_0[\times]\tilde{z} - \delta, \tilde{z} + \delta[$, where δ is a small positive number:

$$\begin{aligned}\partial_z f(h,z) &= 2\frac{q-1}{q-p} z^{2\frac{q-1}{q-p}-1} \tau^{-\frac{\ell+1}{\ell+r}} (E(hz^{-\frac{q}{q-p}}) + \hat{\gamma}) \\ &\quad + 2\frac{\ell+1}{\ell+r} z^{2\frac{q-1}{q-p}+1} h^{2\frac{1-\theta}{\varkappa}} \tau^{-\frac{\ell+1}{\ell+r}-1} (E(hz^{-\frac{q}{q-p}}) + \hat{\gamma}) - \frac{q}{q-p} z^{\frac{q-2}{q-p}-1} h \tau^{-\frac{\ell+1}{\ell+r}} E'(hz^{-\frac{q}{q-p}}).\end{aligned}$$

In view of Lemma 3.2, the derivative above is strictly negative if $(h,z) \in [0, h_0[\times]\tilde{z} - \delta, \tilde{z} + \delta[$ for a suitable choice of small h_0, δ. Note that $f(0, \tilde{z}) = 0$.

Because of the definition of $\tilde{z}$ and the definition of f, we have $f(h, \tilde{z} - \delta) > 0$, $f(h, \tilde{z} + \delta) < 0$ possibly taking a smaller h_0, δ for $0 \le h \le h_0$. Since f is continuous and strictly decreasing on the h-lines, there is a unique zero of the equation $f(h, z(h)) = 0$ with $z(h) \in [\tilde{z} - \delta, \tilde{z} + \delta]$ for $0 \le h \le h_0$.

For positive h, trivially $z(h)$ is real analytic. Let us show that $z(h) \in C([0, h_0[)$. Arguing by contradiction, assume that $z(h) \not\to \tilde{z}$ for $h \to 0+$. Then there is a sequence $h_k \to 0+$ such that $z(h_k) \to \hat{z} \neq \tilde{z}$. Then $0 = f(h_k, z(h_k)) \to f(0, \hat{z})$, which is false since $\tilde{z}$ is the only zero of $f(0, z) = 0$.

The conclusion follows. □

We state also a couple of lemmas that will be needed in the sequel.

Write $V(x) = x^{2(\ell+r-1)} - x^{2\ell} - \hat{\mu}$ and $Q_h = -h^2\partial_x^2 + V$. We have:

Lemma 3.5. *For every $h_0 > 0$ and every $v \in \mathscr{S}(\mathbb{R})$ the following a priori inequality holds*:

$$h^2\|v''\| + \|Vv\| \le C(\|Q_h v\| + h\|v\|) \tag{3-19}$$

for a positive constant C independent of $h \in]0, h_0[$.

Lemma 3.6. *Let $v(x,h)$ denote the $L^2(\mathbb{R})$ normalized eigenfunction of Q_h corresponding to $E(h)$. Then v is rapidly decreasing with respect to x and satisfies the estimates*

$$|v^{(j)}(x,h)| \le C_j h^{-(j+1)/2} \tag{3-20}$$

for $x \in \mathbb{R}$, $C_j > 0$ independent of $0 < h < h_0$, $j = 0, 1, 2$, with h_0 suitably small.

Lemmas 3.5 and 3.6 are rather standard and, for a proof, we refer for instance to the appendix of [Albano et al. 2016].

We can now go back to (3-4). With the choice above of both u_1, u_2 and $z = z(\rho)$, we can satisfy (3-5), provided $\rho \ge \rho_0$, with ρ_0 large enough. Furthermore we also remark that the formal operation of taking derivatives under the integral sign is completely legitimate, due to Lemma 3.6, since a power singularity at infinity does not affect the convergence of the integral. We have then that

$$
\begin{aligned}
&P(x,D)A(u)(x)\\
&=\int_1^{\rho_0} e^{-i\rho x_4+x_3 z(\rho)\rho^\theta-\rho^\theta}\\
&\qquad\times\Big[-\rho^{\frac{2}{\ell+r}}u_2(x_2,\rho)\big(\partial_1^2-y_1^{2(\ell+r-1)}(1-(z(\rho))^2\rho^{2(\theta-1)})\big)u_1(y_1,\rho)\\
&\qquad\qquad+\rho^{-\frac{2\ell}{\ell+r}}y_1^{2\ell}u_1(y_1,\rho)\big[\big(-\partial_2^2-x_2^{2(p-1)}(z(\rho))^2\rho^{2\theta}+x_2^{2(q-1)}\rho^2\big)\big]u_2(x_2,\rho)\Big]_{y_1=\rho^{1/(\ell+r)}x_1}d\rho.
\end{aligned}
\tag{3-21}
$$

Here we used that for $\rho \geq \rho_0$, (3-5) vanishes and we are left with an integral over a finite interval whose upper endpoint depends on the problem data. This defines a real analytic function, $g(x)$.

We need now to check that the growth rates of u_1 and u_2 do not affect the behavior of the integral (3-2) where u has been replaced by the right-hand side of (3-3).

Both u_1 and u_2 are eigenfunctions of the same kind of Schrödinger-type operator with different expressions of Planck's constant: u_2 is an eigenfunction of the operator in (3-11), while u_1 is an eigenfunction of the operator (3-15) where $z(h)$ has been determined according to Proposition 3.3. It is not difficult to see that the two equations are similar, so that discussing one of them is enough.

Let us focus on (3-11). We have to discuss u_2 in a classically forbidden region, i.e., where h is small, which corresponds to large values of ρ, since x is in a neighborhood of the origin. More precisely we need an estimate of the form (3-24), i.e., a bound from below of $u_2(0,\hbar)$. This type of tunneling estimate could be deduced from the results of Helffer and Sjöstrand [1984]; see also [Helffer 1988, Section 2.3]. Another way of deriving such an estimate as a consequence of [Helffer and Sjöstrand 1984] uses [Martinez 1987].

In the present particular case, we may easily reduce the problem of a pointwise estimate to the problem of an L^2 estimate and we actually use a bound, given by Zworski [2012], for the L^2 norm of u_2 in a "forbidden region":

Theorem 3.7 [Zworski 2012, Theorem 7.7]. *Let U be a neighborhood of the origin in $\mathbb{R}$. There exist positive constants $C, \hbar_0$ such that, for $0 < \hbar \leq \hbar_0$,*

$$
\|u_2\|_{L^2(U)} \geq e^{-\frac{C}{\hbar}}\|u_2\|_{L^2(\mathbb{R})}. \tag{3-22}
$$

Here

$$
\hbar = \frac{h}{z(h)^{\frac{q}{q-p}}}, \tag{3-23}
$$

and we note that $\hbar$ is small if and only if h is small, because $z(h)$ is bounded away from zero when h tends to zero.

The Schrödinger operators we deal with have a symmetric potential, so that their eigenfunctions are either even or odd functions with respect to the variable x. The argument is analogous to that in [Albano et al. 2016] and we just sketch it for completeness.

Case of even eigenfunctions u_1, u_2. We may assume that

$$\|u_2\|_{L^2(\mathbb{R})} = 1, \quad u_2(0,\hbar) > 0,$$

since $u_2'(0,\hbar) = 0$ because of its parity and $u_2(0,\hbar) = 0$ would imply that u_2, being a solution of (3-12), is identically zero. A similar conclusion holds for u_1.

Moreover, by (3-12), $\partial_y^2 u_2(0,\hbar) > 0$.

Denote by $x_0 = x_0(\hbar)$ the first positive zero of $V(x) - E(\hbar) = x^{2(q-1)} - x^{2(p-1)} - \hat{\gamma} - E(\hbar)$. Note that u_2 is strictly positive in the interval $0 \le x \le x_0$.

By (3-12), u_2 is strictly convex for $0 \le x \le x_0$ and has its minimum at the origin and its maximum at x_0.

Define $\varphi = \partial_x u_2 / u_2$. We have $\varphi > 0$ if $0 < y \le x_0$. Then, writing φ' for $\partial_y \varphi$,

$$\varphi' = \frac{V - E}{\hbar^2} - \varphi^2.$$

The function φ has a maximum in the interval $]0, x_0[$ because $\varphi'(0) > 0$ and $\varphi'(x_0) = -\varphi^2(x_0) < 0$. Denote by $\bar{x}$ the point where the maximum is attained: it lies in the interior of the interval $[0, x_0]$. Moreover we get

$$\varphi(\bar{x}) = \frac{(V(\bar{x}) - E(\hbar))^{1/2}}{\hbar}.$$

From the definition of φ we obtain

$$u_2(0,\hbar) = e^{-\int_0^{x_0} \varphi(s)\,ds} u_2(x_0,\hbar) \ge e^{-x_0 \varphi(\bar{x})} \frac{1}{\sqrt{2x_0}} \|u_2\|_{L^2([-x_0,x_0])} \ge \frac{1}{\sqrt{2x_0}} e^{-\frac{(-\hat{\gamma})^{1/2}}{\hbar}} e^{-\frac{C}{\hbar}}. \tag{3-24}$$

Here we used Theorem 3.7, as well as the facts that $x_0 < 1$, $E(\hbar) > 0$ and u_2 is normalized. We remark that $\liminf x_0(\hbar) > 0$ when $\hbar > 0$.

We are now in a position to conclude the proof of Theorem 2.1 for an even function u_2. We recall that

$$\hbar = \mathscr{O}(\rho^{\frac{r-1}{\ell+r}\frac{1}{q-1} - \frac{\ell+1}{\ell+r}}) = \mathscr{O}(\rho^{-\varkappa}).$$

Note that

$$A(u) = A_0(u) + A_1(u),$$

where A_0 is defined as the integral in (3-2) over the interval $[1, \rho_0]$, while $A_1(u)$ is the integral in (3-2) over the half-line $[\rho_0, +\infty[$. It is very easy to show that $A_0(u)$, as well as the right-hand side of (3-21), are real analytic functions of x, so that $PA_1(u) = P[A(u) - A_0(u)] \in C^\omega$.

We now compute, assuming that both u_1 and u_2 are even,

$$(-D_{x_4})^k A_1(u)(0) = \int_{\rho_0}^{+\infty} e^{-\rho^\theta} \rho^k u_1(0) u_2(0,\rho)\,d\rho \ge u_1(0) C \int_{\rho_0}^{+\infty} e^{-\rho^\theta - C_1 \rho^\varkappa} \rho^k\,d\rho \ge C_2^{k+1} k!^{s_\ell}.$$

The last inequality above holds since, observing that $\varkappa < \theta$,

$$\int_{\rho_0}^{+\infty} e^{-\rho^\theta - C_1\rho^\varkappa} \rho^k \, d\rho \geq C_1 \int_{\rho_0}^{+\infty} e^{-c\rho^\theta} \rho^k \, d\rho = -C_1 \int_0^{\rho_0} e^{-c\rho^\theta} \rho^k \, d\rho + C_2^{k+1} k!^{s_\ell}$$
$$\geq C_2^{k+1} k!^{s_\ell} \left(1 - C_1 C_2^{-(k+1)} \rho_0 e^{-c\rho_0^\theta} \frac{\rho_0^k}{k!^{s_\ell}}\right) \geq C_3^{k+1} k!^{s_\ell},$$

if k is suitably large and C_3 is suitable and positive.

Case when u_1 is even and u_2 is odd. We may assume that

$$\|u_2\|_{L^2(\mathbb{R})} = 1, \quad u_2'(0,\hbar) > 0.$$

Moreover, due to the parity, $u_2''(0,\hbar) = 0$. Arguing as above we obtain that u_2' is positive in $[0, x_0]$. Set

$$\varphi = \frac{u_2''}{u_2'}.$$

Arguing as above we deduce

$$u_2'(0,\hbar) \geq e^{-\frac{(-\hat{\gamma})^{1/2}}{\hbar}} u_2'(x_0,\hbar) \geq \frac{1}{\sqrt{2x_0}} e^{-\frac{(-\hat{\gamma})^{1/2}}{\hbar}} \|u_2'\|_{L^2([-x_0,x_0])}.$$

Since

$$\|u_2\|_{L^2([-x_0,x_0])} \leq x_0 \|u_2'\|_{L^2([-x_0,x_0])},$$

we get

$$u_2'(0,\hbar) \geq \frac{1}{x_0\sqrt{2x_0}} e^{-\frac{(-\hat{\gamma})^{1/2}}{\hbar}} \|u_2\|_{L^2([-x_0,x_0])}.$$

Using Theorem 3.7 as before we can conclude exactly as in the case of an even eigenfunction.

To finish the proof of Theorem 2.1 we recall that Lemma 3.6 implies that the integral in the definition of $A(\partial_{x_2} u)$ is absolutely convergent, so that, arguing as before, we have

$$(-D_{x_4})^k A_1(\partial_{x_2} u)(0) = \int_{\rho_0}^{+\infty} e^{-\rho^\theta} \rho^k u_1(0)\, \partial_{x_2} u_2(0,\rho)\, d\rho$$
$$\geq u_1(0)\, C \int_{\rho_0}^{+\infty} e^{-\rho^\theta - C_1\rho^\varkappa} \rho^k \, d\rho \geq C_2^{k+1} k!^{s_\ell},$$

again provided k is suitably large.

The other cases, when u_1 is odd and u_2 is even or odd, are treated analogously and we skip them.

This concludes the proof of the optimality part of Theorem 2.1.

4. Proof of Theorem 2.1 (continued)

In this section we prove that the operator P is Gevrey s_ℓ-hypoelliptic. We also point out that the proof given here works when $p = q$, yielding analytic hypoellipticity.

It is useful to establish notation for the vector fields defining P:

$$\begin{aligned} P(x,D) &= D_1^2 + x_1^{2\ell} D_2^2 + x_1^{2(\ell+r-1)} D_3^2 + x_1^{2(\ell+r-1)} D_4^2 + x_1^{2\ell} x_2^{2(p-1)} D_3^2 + x_1^{2\ell} x_2^{2(q-1)} D_4^2 \\ &= \sum_{j=1}^{6} X_j(x,D)^2. \end{aligned} \tag{4-1}$$

We note that, using commutators of the fields up to the length $\ell+r$, we generate the ambient space.

The basic idea for the proof is to use the subelliptic estimate; see, e.g., [Jerison 1986] or [Bolley et al. 1982] for a proof of the inequality

$$\|u\|^2_{\frac{1}{\ell+r}} + \sum_{j=1}^{6} \|X_j(x,D)u\|^2 \leq C\big(\langle P(x,D)u, u\rangle + \|u\|^2\big), \tag{4-2}$$

where C is a positive constant, $\|\cdot\|_{\frac{1}{\ell+r}}$ is the Sobolev norm of order $\frac{1}{\ell+r}$ and $u \in C_0^\infty(\mathbb{R}^4)$.

A further remark is that we may assume $\xi_4 \geq 1$: in fact denoting by ψ a cutoff function such that $\psi \geq 0$, $\psi(\xi_4) = 1$ if $\xi_4 \geq 2$ and $\psi(\xi_4) = 0$ if $\xi_4 \leq 1$, we may apply $\psi(D_4)$ to the equation $Pu = f$, getting $P\psi u = \psi f$, since ψ commutes with P. On the other hand, $\psi f \in G^s$ if $f \in G^s$, for $s \geq 1$, and we are interested in the microlocal Gevrey regularity of u at the point $(0, e_4)$. We write u instead of ψu.

The proof below uses the estimate (4-2) in the following way. Since P is C^∞ hypoelliptic, we may assume the function u to be smooth. If we want to show that it belongs to a Gevrey class we have to bound its derivatives, or, which is easier using (4-2), the L^2 norm of its derivatives by suitable factorials (see Definition 2.2).

To do that, we start with the quantity $X_j \varphi D_4^N u$, where X_j is one of the vector fields appearing in the operator, φ is a cutoff — in general a microlocal cutoff — which is discussed below, and N is an arbitrarily large natural number. If we succeed in bounding $\|X_j \varphi D_4^N u\|$ with $C^{N+1} N!^s$, then we may deduce that u belongs to the Gevrey class G^s on the domain of the cutoff φ.

Thus, feeding the quantity $\varphi D_4^N u$ on the left-hand side of (4-2), we have to estimate the right-hand side: there we have an error term that usually is easy to absorb on the left, but also a term where P appears. In particular we have to treat the term $\langle P\varphi D_4^N u, \varphi D_4^N u\rangle$. Since we are assuming Pu is real analytic or Gevrey s, it is evident that commuting P past the cutoff will lead us to a term $\langle \varphi D_4^N Pu, \varphi D_4^N u\rangle$. This is good, since Pu has analytic estimates and the right factor of the scalar product can be absorbed on the left-hand side of the inequality, like the error term. Unfortunately there is also the commutator $\langle [P, \varphi D_4^N]u, \varphi D_4^N u\rangle$.

Now the commutator either gives vector fields applied to $\varphi D_4^n u$, which are easily absorbed on the left, or gives derivatives of the cutoff and of the coefficients of the vector fields. The derivatives of the cutoff are fine, provided we use suitably chosen cutoffs, see below for this, but the derivatives of the coefficients of the vector fields are more difficult to handle. Actually either one is able to extract another vector field from them — which seldom occurs — or the best resource available is the subelliptic term: we may lower the exponent N by the subellipticity and use the subelliptic part of the estimate to start over with a lower N and derivatives of φ replacing φ. The process terminates when N is completely used and the derivatives of the cutoff give us the final estimate.

Let us start by denoting by $\varphi_N = \varphi_N(x_3, x_4)$, $\chi_N = \chi_N(\xi_4)$ cutoffs of Ehrenpreis type, i.e., $\varphi_N \in C_0^\infty(\mathbb{R}^2)$, $\chi_N \in C^\infty(\mathbb{R})$, with $\varphi_N = 1$ near the origin, $\chi_N = 0$ for $x < 3$ and $\chi_N = 1$ for $x > 4$, and φ_N, χ_N nonnegative.

Ehrenpreis-type functions have the property that $|\partial^k \varphi_N(x)| \le C_\varphi^{k+1} N^k$ for $k \le RN$, $R > R_0$, and C_φ independent of N. See, e.g., [Hörmander 1971, Lemma 2.2] for the definition as well as a construction of such type of functions.

As sketched above we want to estimate the quantity $\|X_j \varphi_N D_4^N u\|$, $j = 1, \dots, 6$, so that getting an estimate of the form $\|X_j \varphi_N D_4^N u\| \le C^{N+1} N^{s_\ell N}$ will be enough to conclude that $u \in G^{s_\ell}$ microlocally at $(0, e_4)$.

As a preliminary remark we point out that if $Pu = f$, $f \in G^{s_\ell}(\Omega)$, then we may assume that $u \in C^\infty(\Omega)$ and u has compact support with respect to the variables x_1, x_2. In fact, if $\theta = \theta(x_1, x_2) \in G^{s_\ell} \cap C_0^\infty$ and is identically equal to 1 in a neighborhood of the origin, we obtain, multiplying the equation $Pu = f$ by θ, that $P(\theta u) = \theta f - [P, \theta] u$ and the commutator term is identically zero in a neighborhood of the origin in the (x_1, x_2)-plane; i.e., it is in G^{s_ℓ}, since u is in G^{s_ℓ} outside of the characteristic manifold. We write u instead of θu.

Now

$$\|X_j \varphi_N D_4^N u\| \le \|X_j \varphi_N (1 - \chi_N(N^{-1} D_4)) D_4^N u\| + \|X_j \varphi_N \chi_N(N^{-1} D_4) D_4^N u\|. \tag{4-3}$$

Consider the first summand above. Since $(1 - \chi_N)\psi$ has support for $1 \le \xi_4 \le 4N$, we deduce immediately a bound of the first summand:

$$\|X_j \varphi_N (1 - \chi_N(N^{-1} D_4)) D_4^N u\| \le C^{N+1} N^N,$$

where C denotes a positive constant independent of N, but depending on u. This means a real analytic growth rate for u. It is enough then to bound the second summand in (4-3).

To do this we plug the quantity $\varphi_N \chi_N D_4^N u$ into (4-2) and, as a consequence, we obtain

$$\begin{aligned} &\|X_j \varphi_N \chi_N(N^{-1} D_4) D_4^N u\|^2 \\ &\qquad \le C\big(\big\langle P \varphi_N \chi_N(N^{-1} D_4) D_4^N u, \varphi_N \chi_N(N^{-1} D_4) D_4^N u\big\rangle + \|\varphi_N \chi_N(N^{-1} D_4) D_4^N u\|^2\big) \end{aligned}$$

Our main concern is the estimate of the scalar product in the next-to-last line of the above formula. We have

$$\begin{aligned} &\big\langle P \varphi_N \chi_N(N^{-1} D_4) D_4^N u, \varphi_N \chi_N(N^{-1} D_4) D_4^N u\big\rangle \\ &\qquad = \big\langle \varphi_N \chi_N(N^{-1} D_4) D_4^N P u, \varphi_N \chi_N(N^{-1} D_4) D_4^N u\big\rangle \\ &\qquad\qquad + \sum_{j=1}^{6} \big\langle [X_j^2, \varphi_N] \chi_N(N^{-1} D_4) D_4^N u, \varphi_N \chi_N(N^{-1} D_4) D_4^N u\big\rangle. \end{aligned}$$

The first term in the right-hand side of the above relation poses no problem: in fact $Pu = f \in G^{s_\ell}$ and thus the scalar product is easily estimated by $C^{N+1} N!^{s_\ell}$, while the right factor can be absorbed on the left.

As for the summands containing a commutator,

$$\begin{aligned}\langle [X_j^2, \varphi_N]\chi_N(N^{-1}D_4)D_4^N u, \varphi_N\chi_N(N^{-1}D_4)D_4^N u\rangle \\ = 2\langle [X_j, \varphi_N]\chi_N(N^{-1}D_4)D_4^N u, X_j\varphi_N\chi_N(N^{-1}D_4)D_4^N u\rangle \\ -\langle N^{-1}[X_j, [X_j, \varphi_N]]\chi_N(N^{-1}D_4)D_4^N u, N\varphi_N\chi_N(N^{-1}D_4)D_4^N u\rangle. \quad (4\text{-}4)\end{aligned}$$

Here we multiplied and divided by N the factors of the second scalar product to compensate for the second derivative landing on φ_N because of the double commutator. The naïve idea behind this is that one derivative of φ_N is worth N, since φ_N is an Ehrenpreis-type cutoff function.

We are going to examine the terms with a single commutator first. Both X_1, X_2 commute with φ_N at this moment, since φ_N depends on x_3 and x_4, even though we shall see shortly that this is not going to be true any longer, for both X_1 and X_2 will have a nonzero commutator with the coefficients of the vector fields. Moreover

$$[X_j, \varphi_N]\chi_N(N^{-1}D_4)D_4^N u = x_1^{\ell+r-1}\varphi_N'\chi_N(N^{-1}D_4)D_4^N u \quad (4\text{-}5)$$

for $j = 3, 4$. Here we just denote by φ_N' a (self-adjoint) derivative with respect to x_3 or x_4, since a more precise notation would only burden the exposition. Furthermore we have

$$[X_5, \varphi_N]\chi_N(N^{-1}D_4)D_4^N u = x_1^{\ell}x_2^{p-1}\varphi_N'\chi_N(N^{-1}D_4)D_4^N u$$

and

$$[X_6, \varphi_N]\chi_N(N^{-1}D_4)D_4^N u = x_1^{\ell}x_2^{q-1}\varphi_N'\chi_N(N^{-1}D_4)D_4^N u.$$

Let us consider the terms corresponding to $j = 3, 4$ first:

$$\begin{aligned}2\left|\langle x_1^{\ell+r-1}\varphi_N'\chi_N(N^{-1}D_4)D_4^N u, X_j\varphi_N\chi_N(N^{-1}D_4)D_4^N u\rangle\right| \\ \le \delta\|X_j\varphi_N\chi_N(N^{-1}D_4)D_4^N u\|^2 + \frac{1}{\delta}\|x_1^{\ell+r-1}\varphi_N'\chi_N(N^{-1}D_4)D_4^N u\|^2, \quad (4\text{-}6)\end{aligned}$$

where δ is a positive number so small to allow us to absorb the first summand in the right-hand side of (4-6) on the left of the subelliptic estimate.

In order to be able to apply again the subelliptic estimate to the second summand above we need to use the formula

$$\varphi_N' D_4^N = \sum_{j=0}^{N-1}(-1)^j D_4\varphi_N^{(j+1)}D_4^{N-j-1} + (-1)^N\varphi_N^{(N+1)}. \quad (4\text{-}7)$$

Thus, since $\chi_N(N^{-1}D_4)$ commutes with D_4^N,

$$\begin{aligned}\|x_1^{\ell+r-1}\varphi_N'\chi_N(N^{-1}D_4)D_4^N u\| \\ \le \sum_{j=0}^{N-1}\|X_4\varphi_N^{(j+1)}\chi_N(N^{-1}D_4)D_4^{N-j-1}u\| + \|\varphi_N^{(N+1)}\chi_N(N^{-1}D_4)u\|, \quad (4\text{-}8)\end{aligned}$$

where we used the fact that the field X_4 can be reconstructed using the factor $x_1^{\ell+r-1}$ and just "pulling back" one x_4-derivative. A completely analogous treatment leads to an analogous conclusion when $j=6$:

$$\begin{aligned}\|[X_6,\varphi_N]\chi_N(N^{-1}D_4)D_4^N u\| &= \|x_1^\ell x_2^{q-1}\varphi_N'\chi_N(N^{-1}D_4)D_4^N u\| \\ &\le \sum_{j=0}^{N-1}\|X_6\varphi_N^{(j+1)}\chi_N(N^{-1}D_4)D_4^{N-j-1}u\|+\|\varphi_N^{(N+1)}\chi_N(N^{-1}D_4)u\|.\end{aligned} \tag{4-9}$$

Furthermore it is clear that the terms on the right of inequalities (4-8) and (4-9) yield a real analytic growth estimate: in fact a typical term in the sum loses $j+1$ derivatives of u with respect to x_4 while the cutoff function φ_N picks up $j+1$ derivatives. Using the properties of φ_N we see that, arguing inductively, this gives analytic growth with respect to N. Same argument for the last terms in (4-8) and (4-9).

We are thus left with the commutator term for $j=5$ in (4-4):

$$\begin{aligned}&2\left|\langle x_1^\ell x_2^{p-1}\varphi_N'\chi_N(N^{-1}D_4)D_4^N u, X_5\varphi_N\chi_N(N^{-1}D_4)D_4^N u\rangle\right| \\ &\qquad\le \delta\|X_5\varphi_N\chi_N(N^{-1}D_4)D_4^N u\|^2+\frac{1}{\delta}\|x_1^\ell x_2^{p-1}\varphi_N'\chi_N(N^{-1}D_4)D_4^N u\|^2.\end{aligned} \tag{4-10}$$

Here, again, δ is chosen so that the first term in the right-hand side above can be absorbed on the left of the subelliptic estimate, as before. We just need to be concerned with the second term. Contrary to what has been done before, pulling back one derivative is of no help, since $p<q$ and the derivative with respect to x_4 is the only derivative available here. Note that if $p=q$ then we may act at this point as we did for $j=3,4,6$, obtaining analytic growth estimates. So let us go on assuming that $p<q$.

Hence we have to resort to the subelliptic part of the subelliptic estimate, i.e., the $1/(\ell+r)$–Sobolev norm. To do this we pull back $D_4^{1/(\ell+r)}$. This is well defined since $\xi_4>1$, but is a pseudodifferential operator, and its commutator with φ_N needs some care.

We actually have the following lemmas. Let $\omega_N\in C^\infty(\mathbb{R})$ be an Ehrenpreis-type cutoff such that $\omega_N=1$ for $x>2$ and $\omega_N=0$ for $x<1$, ω_N nonnegative and such that $\omega_N\chi_N=\chi_N$. Then we have:

Lemma 4.1. *Let $0<\theta<1$. Then*

$$[\omega_N(N^{-1}D)D^\theta,\varphi_N(x)]\chi_N(N^{-1}D)D^{N-\theta}=\sum_{k=1}^{N}a_{N,k}(x,D)\chi_N(N^{-1}D)D^N, \tag{4-11}$$

where $a_{N,k}$ is a pseudodifferential operator of order $-k$ such that

$$|\partial_\xi^\alpha a_{N,k}(x,\xi)|\le C_a^{k+1}N^{k+\alpha}\xi^{-k-\alpha},\quad 1\le k\le N,\ \alpha\le N. \tag{4-12}$$

Corollary 4.2. *For $1\le k\le N-1$ in (4-11) we have that*

$$a_{N,k}(x,D)\chi_N(N^{-1}D)D^N=\frac{\theta(\theta-1)\cdots(\theta-k+1)}{k!}D_x^k\varphi_N(x)\chi_N(N^{-1}D)D^{N-k}. \tag{4-13}$$

For the proofs we refer to [Albano et al. 2016, Appendix B].

Applying Corollary 4.2, we find that

$$\begin{aligned}\|x_1^\ell x_2^{p-1}\varphi_N'\chi_N(N^{-1}D_4)D_4^N u\| &= \|x_1^\ell x_2^{p-1}\varphi_N'\omega_N(N^{-1}D_4)D_4^{\frac{1}{\ell+r}}\chi_N(N^{-1}D_4)D_4^{N-\frac{1}{\ell+r}}u\| \\ &\le c_0\|x_1^\ell x_2^{p-1}\varphi_N'\chi_N(N^{-1}D_4)D_4^{N-\frac{1}{\ell+r}}u\|_{\frac{1}{\ell+r}} \\ &\quad+\sum_{k=1}^{N-1} c_k\|x_1^\ell x_2^{p-1}\varphi_N^{(k+1)}\chi_N(N^{-1}D_4)D_4^{N-k}u\| \\ &\quad+c_N\|x_1^\ell x_2^{p-1}a_{N,N}(x,D)\chi_N(N^{-1}D_4)D_4^N u\|. \end{aligned} \tag{4-14}$$

Here the constants c_j, $j=0,1,\dots,N$, are bounded independently of N by some absolute constant.

The last term in the right-hand side of (4-14) has analytic growth, because $a_{N,N}$ has order $-N$, so that it balances the N-th derivative on u, and is bounded by $C_a^{N+1}N^N$, according to (4-12). Thus we may forget about it because it gives better estimates than those we are going to get.

The first term on the right-hand side of (4-14) can be resubjected to the subelliptic estimate — Sobolev part — and treated as we just did. This means that again we have to consider the commutator of the vector fields in the operator P with $x_1^\ell x_2^{p-1}\varphi_N'$, since D_4 commutes with P. Hence the quantity we have to estimate is

$$\|[X_j, x_1^\ell x_2^{p-1}\varphi_N']\chi_N(N^{-1}D_4)D_4^{N-\frac{1}{\ell+r}}u\|;$$

see (4-4), (4-5). As we said before, the commutators with X_1, X_2 are no longer zero, because of the monomials $x_1^\ell x_2^{p-1}$. Note that commuting with X_1 has just the effect of lowering the exponent ℓ by one unit, while commuting with X_2 lowers the exponent $p-1$ by one unit and increases the exponent ℓ by ℓ units.

On the other hand commuting with X_3, X_4 and X_6 ignores these types of monomials and allows us to treat the above quantity exactly as we did before, yielding analytic-type growth estimates. Note that the monomials can be estimated by a constant, since we are in a neighborhood of the origin in the (x_1,x_2)-plane.

Finally commuting with X_5 is done again as before, but its outcome is to double the exponents of the monomial above, increase the number of derivatives of φ_N and lower the x_4-derivatives on u by $1/(\ell+r)$-th of a unit. In other words, the first term on the right-hand side of (4-14) at the second application of the subelliptic estimate would be

$$\|x_1^{2\ell}x_2^{2(p-1)}\varphi_N''\chi_N(N^{-1}D_4)D_4^{N-\frac{2}{\ell+r}}u\|_{\frac{1}{\ell+r}}.$$

We use the same argument for every summand in the sum in (4-14), except that here we have to pull back a $D_4^{1/(\ell+r)}$ once more in order to use the subelliptic estimate employing Lemma 4.1 as well as Corollary 4.2.

Hence applying the subelliptic estimate — Sobolev part — alters the exponents in the monomials $x_1^\ell x_2^{p-1}$, lowers the x_4-derivatives on u, and increases the derivatives of φ_N.

When the exponent of x_2 becomes greater than or equal to $q-1$ or the exponent of x_1 becomes greater than or equal to $\ell+r-1$, whichever comes first, we do not need to use the Sobolev part of the subelliptic estimate anymore. At that point we are able to pull back a whole x_4-derivative, reconstructing the field X_6 in the first case or X_4 in the second. The powers of x_1 and x_2 are decreased in the first case, but only the power of x_1 is decreased in the second case.

Before writing the product of such an iteration, we discuss also the terms from (4-4) containing a double commutator. For $j=3,4$ we have that

$$\begin{aligned}\left|\langle N^{-1}x_1^{\ell+r-1}\varphi_N''\chi_N(N^{-1}D_4)D_4^N u, Nx_1^{\ell+r-1}\varphi_N\chi_N(N^{-1}D_4)D_4^N u\rangle\right|&\\ \le \tfrac{1}{2}\|N^{-1}x_1^{\ell+r-1}\varphi_N''\chi_N(N^{-1}D_4)D_4^N u\|^2+\tfrac{1}{2}\|Nx_1^{\ell+r-1}\varphi_N\chi_N(N^{-1}D_4)D_4^N u\|^2.&\end{aligned}$$

Each of the summands in the right-hand side is then treated as we did before with the terms involving a single commutator. Note that $N^{-1}\varphi_N''$ counts as a first derivative and so does $N\varphi_N$. The same argument holds for $j=5,6$.

For $j=1,\dots,6$, we iterate this procedure. The following notation is useful.

Let us denote by a_j the number of times we take a commutator, according to the procedure outlined above, with X_j. If, as we saw before, $j=3,4,6$, the net result is a decrease by one unit of the derivatives of u and a parallel increase by one unit of the derivatives of φ_N. No Sobolev part of the subelliptic estimate is used in this case.

If $j=1,2,5$, the situation is more complicated. As we said, the vector field X_5 contributes a monomial $x_1^{\ell}x_2^{p-1}$ and the Sobolev part will decrease the derivatives on u by $1/(\ell+r)$. However we do not always need the Sobolev part of the estimate. Let us call α, β the number of times we do not apply it, but are able to reconstruct the vector field X_4 — using $x_1^{\ell+r-1}$ — or X_6 — using the monomial $x_1^{\ell}x_2^{p-1}$. Of course this has an effect also on the number of derivatives on u, which lose $\alpha+\beta$ units instead of just $(\alpha+\beta)/(\ell+r)$.

Finally we must take into account all terms deriving from the application of Lemma 4.1 and Corollary 4.2. Let us denote by h the number of times we apply them and by k_j, $j=1,\dots,h$, the summation indices in (4-11). As we saw, these also have an effect on the derivatives landing on u and φ_N.

Hence we wind up with the estimate

$$\begin{aligned}&\|X_j\varphi_N\chi_N(N^{-1}D_4)D_4^N u\|\\ &\lesssim\sum\Big\|x_1^{a_5\ell-\alpha(\ell+r-1)-\beta\ell-a_1}x_2^{a_5(p-1)-\beta(q-1)-a_2}\varphi_N^{(1+a_3+a_4+a_5+a_6+k_1+\dots+k_h)}\\ &\qquad\qquad\cdot\chi_N(N^{-1}D_4)D_4^{N-k_1-\dots-k_h-\frac{a_1+a_2+a_5}{\ell+r}-a_3-a_4-a_6-(\alpha+\beta)\frac{\ell+r-1}{\ell+r}}u\Big\|\\ &\quad+\sum N^{-(a_3+a_4+a_5+a_6)}\Big\|x_1^{a_5\ell-\alpha(\ell+r-1)-\beta\ell-a_1}x_2^{a_5(p-1)-\beta(q-1)-a_2}\varphi_N^{(1+2(a_3+a_4+a_5+a_6)+k_1+\dots+k_h)}\\ &\qquad\qquad\cdot\chi_N(N^{-1}D_4)D_4^{N-k_1-\dots-k_h-\frac{a_2+a_5}{r}-a_3-a_4-a_6-b\frac{r-1}{r}}u\Big\|\\ &\quad+\sum N^{a_3+a_4+a_5+a_6}\Big\|x_1^{a_5\ell-\alpha(\ell+r-1)-\beta\ell-a_1}x_2^{a_5(p-1)-b(q-1)-a_2}\varphi_N^{(1+k_1+\dots+k_h)}\\ &\qquad\qquad\cdot\chi_N(N^{-1}D_4)D_4^{N-k_1-\dots-k_h-\frac{a_2+a_5}{r}-a_3-a_4-a_6-b\frac{r-1}{r}}u\Big\|,\end{aligned}\tag{4-15}$$

where each sum is taken on the indices $a_1, a_2, \dots, a_6, k_1, \dots, k_h$ such that

$$0 \le N - k_1 - \cdots - k_h - \frac{a_1 + a_2 + a_5}{\ell + r} - a_3 - a_4 - a_6 - (\alpha + \beta)\frac{\ell + r - 1}{\ell + r} < 1, \tag{4-16}$$

and the indices satisfy the conditions

$$0 \le a_5(p-1) - \beta(q-1) - a_2 < q - 1 \tag{4-17}$$

and

$$0 \le a_5\ell - a_1 - \alpha(\ell + r - 1) - \beta\ell < \ell. \tag{4-18}$$

A few explanations are in order. Note that (4-16) expresses the fact that the number of derivatives on u is not negative, but is less than 1 at the end. Conditions (4-17) and (4-18) express the fact that the exponent of x_2 is less than $q-1$ and that the exponent of x_1 is less than ℓ, i.e., when we cannot pull back a derivative and need to apply the Sobolev part of the subelliptic estimate when doing the iteration.

Finally we observe that (4-15) has three sums on the right-hand side. The first is the sum coming from the single commutator terms in (4-4), while the others come from the double commutator terms in (4-4). Note that in the last one the power of N is positive, but φ_N has less derivatives than in the second sum. The balance though is the same in both terms.

We have to evaluate the supremum of the right-hand side in (4-15). From the second condition,

$$\frac{a_5(p-1) - a_2}{q-1} - 1 < \beta \le \frac{a_5(p-1) - a_2}{q-1},$$

while from the third we get

$$\frac{a_5\ell - a_1 - \frac{a_5(p-1)-a_2}{q-1}\ell - \ell}{\ell + r - 1} < \alpha < \frac{a_5\ell - a_1 - \frac{a_5(p-1)-a_2}{q-1}\ell + \ell}{\ell + r - 1},$$

so that from the first condition we deduce

$$N - 1 - \frac{\ell}{\ell + r} < \sum_{j=1}^{h} k_j + a_2\gamma_2 + a_3 + a_4 + a_5\frac{1}{s_\ell} + a_6 < N + 1 + \frac{\ell - 1}{\ell + r},$$

where

$$\gamma_2 = \frac{1}{\ell + r} - \frac{r-1}{(\ell + r)(q-1)} > 0.$$

In order to estimate (4-15) we need to compute

$$\max\left\{\sum_{j=1}^{h} k_j + a_3 + a_4 + a_5 + a_6\right\},$$

where the maximum is on all indices verifying conditions (4-16)–(4-18). More precisely, the three sums in (4-15) give the same contribution, because when an index is missing among the derivatives of φ_N it is found at the exponent of the factor N and, vice versa, when it appears twice, it appears with a negative sign at the exponent of N.

It is then clear that the maximum is $\sim N^{Ns_\ell}$ so that we finally get

$$\|X_j \varphi_N \chi_N(N^{-1} D_4) D_4^N u\| \leq C^{N+1} N!^{s_\ell},$$

and this achieves the proof of the theorem, since the above inequality shows that the function u is in the Gevrey class G^{s_ℓ} in the support of φ_N, $N \in \mathbb{N}$.

Note that if $p = q$, a much simpler proof — which means without any use of the Sobolev part of the subelliptic estimate — yields analytic regularity.

Acknowledgements

We would like to express our thanks to Bernard Helffer for some useful criticism and for suggesting a number of important references. We also thank André Martinez for useful discussions on some results of semiclassical analysis.

References

[Albano and Bove 2013] P. Albano and A. Bove, *Wave front set of solutions to sums of squares of vector fields*, Mem. Amer. Math. Soc. **1039**, Amer. Math. Soc., Providence, RI, 2013. MR Zbl

[Albano et al. 2016] P. Albano, A. Bove, and M. Mughetti, "Analytic hypoellipticity for sums of squares and the Treves conjecture", preprint, 2016. arXiv

[Berezin and Shubin 1991] F. A. Berezin and M. A. Shubin, *The Schrödinger equation*, Mathematics and its Applications (Soviet Series) **66**, Kluwer Academic Publishers Group, Dordrecht, 1991. MR Zbl

[Bolley et al. 1982] P. Bolley, J. Camus, and J. Nourrigat, "La condition de Hörmander–Kohn pour les opérateurs pseudo-différentiels", *Comm. Partial Differential Equations* **7**:2 (1982), 197–221. MR Zbl

[Bove and Mughetti 2016] A. Bove and M. Mughetti, "On a new method of proving Gevrey hypoellipticity for certain sums of squares", *Adv. Math.* **293** (2016), 146–220. MR Zbl

[Bove and Treves 2004] A. Bove and F. Treves, "On the Gevrey hypo-ellipticity of sums of squares of vector fields", *Ann. Inst. Fourier* (*Grenoble*) **54**:5 (2004), 1443–1475. MR

[Bove et al. 2013] A. Bove, M. Mughetti, and D. S. Tartakoff, "Hypoellipticity and nonhypoellipticity for sums of squares of complex vector fields", *Anal. PDE* **6**:2 (2013), 371–445. MR Zbl

[Cordaro and Hanges 2009] P. D. Cordaro and N. Hanges, "A new proof of Okaji's theorem for a class of sum of squares operators", *Ann. Inst. Fourier* (*Grenoble*) **59**:2 (2009), 595–619. MR Zbl

[Helffer 1988] B. Helffer, *Semi-classical analysis for the Schrödinger operator and applications*, Lecture Notes in Mathematics **1336**, Springer, 1988. MR Zbl

[Helffer and Sjöstrand 1984] B. Helffer and J. Sjöstrand, "Multiple wells in the semiclassical limit, I", *Comm. Partial Differential Equations* **9**:4 (1984), 337–408. MR Zbl

[Hörmander 1967] L. Hörmander, "Hypoelliptic second order differential equations", *Acta Math.* **119** (1967), 147–171. MR Zbl

[Hörmander 1971] L. Hörmander, "Uniqueness theorems and wave front sets for solutions of linear differential equations with analytic coefficients", *Comm. Pure Appl. Math.* **24** (1971), 671–704. MR Zbl

[Jerison 1986] D. Jerison, "The Poincaré inequality for vector fields satisfying Hörmander's condition", *Duke Math. J.* **53**:2 (1986), 503–523. MR Zbl

[Martinez 1987] A. Martinez, "Estimations de l'effet tunnel pour le double puits, I", *J. Math. Pures Appl.* (9) **66**:2 (1987), 195–215. MR Zbl

[Métivier 1981] G. Métivier, “Non-hypoellipticité analytique pour $D_x^2 + (x^2 + y^2)D_y^2$”, *C. R. Acad. Sci. Paris Sér. I Math.* **292**:7 (1981), 401–404. MR Zbl

[Mughetti 2014] M. Mughetti, “Hypoellipticity and higher order Levi conditions”, *J. Differential Equations* **257**:4 (2014), 1246–1287. MR Zbl

[Mughetti 2015] M. Mughetti, “On the spectrum of an anharmonic oscillator”, *Trans. Amer. Math. Soc.* **367**:2 (2015), 835–865. MR Zbl

[Ōkaji 1985] T. Ōkaji, “Analytic hypoellipticity for operators with symplectic characteristics”, *J. Math. Kyoto Univ.* **25**:3 (1985), 489–514. MR Zbl

[Simon 1983] B. Simon, “Semiclassical analysis of low lying eigenvalues, I: Nondegenerate minima: asymptotic expansions”, *Ann. Inst. H. Poincaré Sect. A* (*N.S.*) **38**:3 (1983), 295–308. Correction in **40**:2 (1984), 224. MR Zbl

[Treves 1999] F. Treves, “Symplectic geometry and analytic hypo-ellipticity”, pp. 201–219 in *Differential equations: La Pietra 1996 (Florence)*, edited by M. Giaquinta et al., Proc. Sympos. Pure Math. **65**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[Treves 2006] F. Treves, “On the analyticity of solutions of sums of squares of vector fields”, pp. 315–329 in *Phase space analysis of partial differential equations*, edited by A. Bove et al., Progr. Nonlinear Differential Equations Appl. **69**, Birkhäuser, Boston, 2006. MR Zbl

[Treves $\geq$ 2017] F. Treves, “Aspects of analytic PDE”, book in preparation.

[Zworski 2012] M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics **138**, Amer. Math. Soc., Providence, RI, 2012. MR Zbl

Received 1 Jun 2016. Revised 24 Feb 2017. Accepted 17 Jun 2017.

ANTONIO BOVE: bove@bo.infn.it
Dipartimento di Matematica, Università di Bologna, Bologna, Italy

MARCO MUGHETTI: marco.mughetti@unibo.it
Dipartimento di Matematica, Università di Bologna, Bologna, Italy

ANALYSIS AND PDE
Vol. 10, No. 7, 2017
dx.doi.org/10.2140/apde.2017.10.1637

msp

PLIABILITY, OR THE WHITNEY EXTENSION THEOREM FOR CURVES IN CARNOT GROUPS

NICOLAS JUILLET AND MARIO SIGALOTTI

The Whitney extension theorem is a classical result in analysis giving a necessary and sufficient condition for a function defined on a closed set to be extendable to the whole space with a given class of regularity. It has been adapted to several settings, including the one of Carnot groups. However, the target space has generally been assumed to be equal to $\mathbb{R}^d$ for some $d \geq 1$.

We focus here on the extendability problem for general ordered pairs (G_1, G_2) (with G_2 nonabelian). We analyse in particular the case $G_1 = \mathbb{R}$ and characterize the groups G_2 for which the Whitney extension property holds, in terms of a newly introduced notion that we call *pliability*. Pliability happens to be related to rigidity as defined by Bryant and Hsu. We exploit this relation in order to provide examples of nonpliable Carnot groups, that is, Carnot groups such that the Whitney extension property does not hold. We use geometric control theory results on the accessibility of control affine systems in order to test the pliability of a Carnot group. In particular, we recover some recent results by Le Donne, Speight and Zimmerman about Lusin approximation in Carnot groups of step 2 and Whitney extension in Heisenberg groups. We extend such results to all pliable Carnot groups, and we show that the latter may be of arbitrarily large step.

1. Introduction

Extending functions is a basic but fundamental tool in analysis. In particular, in 1934 H. Whitney established his celebrated extension theorem, which guarantees the existence of an extension of a function defined on a closed set of a finite-dimensional vector space to a function of class $\mathcal{C}^k$, provided that the minimal obstruction imposed by Taylor series is satisfied. The Whitney extension theorem plays a significant part in the study of ideals of differentiable functions, see [Malgrange 1967], and its variants are still an active research topic of classical analysis; see, for instance, [Fefferman 2005].

Analysis on Carnot groups with a homogeneous distance like the Carnot–Carathéodory distance, as presented in [Folland and Stein 1982], is nowadays a classical topic too. Carnot groups provide a generalization of finite-dimensional vector spaces that is both close to the original model and radically different. This is why Carnot groups provide a wonderful field of investigation in many branches of mathematics. Not only is the setting elegant and rich but it is at the natural crossroads of different fields of mathematics, for instance, of analysis of PDEs and geometric control theory; see [Barilari et al. 2016a; 2016b] for a contemporary account. It is therefore natural to recast the Whitney extension theorem in the context of Carnot groups. As far as we know, the first generalization of a Whitney extension

MSC2010: 22E25, 41A05, 53C17, 54C20, 58C25.
Keywords: Whitney extension theorem, Carnot group, rigid curve, horizontal curve.

theorem to Carnot groups can be found in [Franchi et al. 2001; 2003], where De Giorgi's result on sets of finite perimeter is adapted first to the Heisenberg group and then to any Carnot group of step 2. This generalization is used in [Kirchheim and Serra Cassano 2004], where the authors stress the difference between intrinsic regular hypersurfaces and classical $\mathcal{C}^1$ hypersurfaces in the Heisenberg group. The recent paper [Vodop'yanov and Pupyshev 2006a] gives a final statement for the Whitney extension theorem for scalar-valued functions on Carnot groups: the most natural generalization that one can imagine holds in its full strength (for more details, see Section 2).

The study of the Whitney extension property for Carnot groups is however not closed. Following a suggestion by F. Serra Cassano [2016], one might consider maps between Carnot groups instead of solely scalar-valued functions on Carnot groups. The new question presents richer geometrical features and echoes classical topics of metric geometry. We think in particular of the classification of Lipschitz embeddings for metric spaces and of the related question of the extension of Lipschitz maps between metric spaces. We refer to [Balogh and Fässler 2009; Wenger and Young 2010; Rigot and Wenger 2010; Balogh et al. 2016] for the corresponding results for the most usual Carnot groups: abelian groups $\mathbb{R}^m$ or Heisenberg groups $\mathbb{H}_n$ (of topological dimension $2n+1$). In view of the Pansu–Rademacher theorem on Lipschitz maps (see Proposition 2.1), the most directly related Whitney extension problem is the one for $\mathcal{C}^1_H$-maps, the so-called horizontal maps of class $\mathcal{C}^1$ defined on Carnot groups. This is the framework of our paper.

Simple arguments show that the Whitney extension theorem does not generalize to every ordered pair of Carnot groups. Basic facts in contact geometry suggest that *the extension does not hold for* $(\mathbb{R}^{n+1}, \mathbb{H}_n)$, i.e., for maps from $\mathbb{R}^{n+1}$ to $\mathbb{H}_n$. It is actually known that local algebraic constraints of first order make n the maximal dimension for a Legendrian submanifold in a contact manifold of dimension $2n+1$. In fact if the derivative of a differentiable map has range in the kernel of the contact form, the range of the map has dimension at most n. A map from $\mathbb{R}^{n+1}$ to $\mathbb{H}_n$ is $\mathcal{C}^1_H$ if it is $\mathcal{C}^1$ with horizontal derivatives, i.e., if its derivatives take value in the kernel of the canonical contact form. In particular, a $\mathcal{C}^1_H$-map defined on $\mathbb{R}^{n+1}$ is nowhere of maximal rank. Moreover, it is a consequence of the Pansu–Rademacher theorem that Lipschitz maps from $\mathbb{R}^{n+1}$ to $\mathbb{H}_n$ are differentiable at almost every point with only horizontal derivatives. Again n is their maximal rank. In order to contradict the extendability of Lipschitz maps, it is enough to define a function on a subset whose topological constraints force any possible extension to have maximal rank at some point. *Let us sketch a concrete example that provides a constraint for the Lipschitz extension problem*: It is known that $\mathbb{R}^n$ can be isometrically embedded in $\mathbb{H}_n$ with the exponential map (for the Euclidean and Carnot–Carathéodory distances). One can also consider two "parallel" copies of $\mathbb{R}^n$ in $\mathbb{R}^{n+1}$ mapped to parallel images in $\mathbb{H}_n$; the second is obtained from the first by a vertical translation. Aiming for a contradiction, suppose that there exists an extending Lipschitz map F. It provides on $\mathbb{R}^n \times [0, 1]$ a Lipschitz homotopy between $F(\mathbb{R}^n \times \{0\})$ and $F(\mathbb{R}^n \times \{1\})$. Using the definition of a Lipschitz map and some topology, the topological dimension of the range is at least $n+1$ and its $(n+1)$-Hausdorff measure is positive. This is not possible because of the dimensional constraints explained above. See [Balogh and Fässler 2009] for a more rigorous proof using a different set as a domain for the function to be extended. That proof is formulated in terms of index theory and the purely $(n+1)$-unrectifiability of $\mathbb{H}_n$.

The latter property means that the $(n+1)$-Hausdorff measure of the range of a Lipschitz map is zero. This construction and some other ideas from the works on the Lipschitz extension problem [Balogh and Fässler 2009; Wenger and Young 2010; Rigot and Wenger 2010; Balogh et al. 2016] can probably be adapted to the Whitney extension problem. It is not really our concern in the present article to list the similarities between the two problems, but rather to exhibit a class of ordered pairs of Carnot groups for which the validity of the Whitney extension problem depends on the geometry of the groups. Note that a different type of counterexample to the Whitney extension theorem, involving groups which are neither Euclidean spaces nor Heisenberg groups, has been obtained by A. Kozhevnikov [2015]. It is described in Example 2.6.

Our work is motivated by a suggestion of Serra Cassano [2016]. He proposed (i) to choose general Carnot groups $\mathbb{G}$ as target space, and (ii) to look at $\mathcal{C}^1_H$ curves only, i.e., $\mathcal{C}^1$ maps from $\mathbb{R}$ to $\mathbb{G}$ with horizontal derivatives. As we will see, the problem is very different from the Lipschitz extension problem for $(\mathbb{R}, \mathbb{G})$ and from the Whitney extension problem for $(\mathbb{G}, \mathbb{R})$. Indeed, both such problems can be solved for every $\mathbb{G}$, while the answer to the extendibility question asked by Serra Cassano depends on the choice of $\mathbb{G}$. More precisely, we provide a geometric characterization of those $\mathbb{G}$ for which the $\mathcal{C}^1_H$-Whitney extension problem for $(\mathbb{R}, \mathbb{G})$ can always be solved. We say in this case that the pair $(\mathbb{R}, \mathbb{G})$ has the $\mathcal{C}^1_H$ extension property. Examples of target nonabelian Carnot groups for which $\mathcal{C}^1_H$ extendibility is possible have been identified by S. Zimmerman [2017], who proved that for every $n \in \mathbb{N}$ the pair $(\mathbb{R}, \mathbb{H}_n)$ has the $\mathcal{C}^1_H$ extension property.

The main component of the characterization of Carnot groups $\mathbb{G}$ for which $(\mathbb{R}, \mathbb{G})$ has the $\mathcal{C}^1_H$ extension property is the notion of *pliable horizontal vector*. A horizontal vector X (identified with a left-invariant vector field) is pliable if for every $p \in \mathbb{G}$ and every neighborhood Ω of X in the horizontal layer of $\mathbb{G}$, the support of all $\mathcal{C}^1_H$ curves with derivative in Ω starting from p in the direction X form a neighborhood of the integral curve of X starting from p (for details, see Definition 3.4 and Proposition 3.7). This notion is close but not equivalent to the property of the integral curves of X not being rigid in the sense introduced by Bryant and Hsu [1993], as we illustrate in Example 3.5. We say that *a Carnot group* $\mathbb{G}$ *is pliable if all its horizontal vectors are pliable*. Since any rigid integral curve of a horizontal vector X is not pliable, it is not hard to show that there exist nonpliable Carnot groups of any dimension larger than 3 and of any step larger than 2 (see Example 3.3). On the other hand, we give some criteria ensuring the pliability of a Carnot group, notably the fact that it has step 2 (Theorem 6.5). We also prove the existence of pliable groups of any positive step (Proposition 6.6).

Our main theorem is the following.

Theorem 1.1. *The pair* $(\mathbb{R}, \mathbb{G})$ *has the* $\mathcal{C}^1_H$ *extension property if and only if* $\mathbb{G}$ *is pliable.*

The paper is organized as follows: in Section 2 we recall some basic facts about Carnot groups and we present the $\mathcal{C}^1_H$-Whitney condition in the light of the Pansu–Rademacher theorem. In Section 3 we introduce the notion of pliability, we discuss its relation with rigidity, and we show that pliability of $\mathbb{G}$ is necessary for the $\mathcal{C}^1_H$ extension property to hold for $(\mathbb{R}, \mathbb{G})$ (Theorem 3.8). The proof of this result assumes that a nonpliable horizontal vector exists and uses it to provide an explicit construction of a $\mathcal{C}^1_H$ map defined on a closed subset of $\mathbb{R}$ which cannot be extended on $\mathbb{R}$. Section 4 is devoted to proving that pliability is also a sufficient condition (Theorem 4.4). In Section 5 we use our result to extend a Lusin-like theorem proved recently by G. Speight [2016] for Heisenberg groups; see also [Zimmerman 2017] for an alternative proof. More precisely, it is proved in [Le Donne and Speight 2016] that an absolutely continuous curve in a group of step 2 coincides on a set of arbitrarily small complement with a $\mathcal{C}^1_H$ curve. We show that this is the case for pliable Carnot groups (Proposition 5.2). Finally, in Section 6 we give some criteria for testing the pliability of a Carnot group. We first show that the zero horizontal vector is always pliable (Proposition 6.1). Then, by applying some results of control theory providing criteria under which the endpoint mapping is open, we show that $\mathbb{G}$ is pliable if its step is equal to 2.

2. The Whitney condition in Carnot groups

A nilpotent Lie group $\mathbb{G}$ is said to be a *Carnot group* if it is stratified in the sense that its Lie algebra $\mathfrak{G}$ admits a direct sum decomposition

$$\mathfrak{G}_1 \oplus \cdots \oplus \mathfrak{G}_s,$$

called *stratification*, such that $[\mathfrak{G}_i, \mathfrak{G}_j] = \mathfrak{G}_{i+j}$ for every $i, j \in \mathbb{N}^*$ with $i + j \leq s$ and $[\mathfrak{G}_i, \mathfrak{G}_j] = \{0\}$ if $i + j > s$. We recall that $[\mathfrak{G}_i, \mathfrak{G}_j]$ denotes the linear space spanned by $\{[X, Y] \in \mathfrak{G} \mid X \in \mathfrak{G}_i,\ Y \in \mathfrak{G}_j\}$. The subspace $\mathfrak{G}_1$ is called the *horizontal layer* and it is also denoted by $\mathfrak{G}_H$. We say that s is the *step* of $\mathbb{G}$ if $\mathfrak{G}_s \neq \{0\}$. The group product of two elements $x_1, x_2 \in \mathbb{G}$ is denoted by $x_1 \cdot x_2$. Given $X \in \mathfrak{G}$ we write $\mathrm{ad}_X : \mathfrak{G} \to \mathfrak{G}$ for the operator defined by $\mathrm{ad}_X Y = [X, Y]$.

The Lie algebra $\mathfrak{G}$ can be identified with the family of left-invariant vector fields on $\mathbb{G}$. The exponential is the application that maps a vector X of $\mathfrak{G}$ into the endpoint at time 1 of the integral curve of the vector field X starting from the identity of $\mathbb{G}$, denoted by $0_{\mathbb{G}}$. That is, if

$$\gamma(0) = 0_{\mathbb{G}} \quad \text{and} \quad \dot{\gamma}(t) = X \circ \gamma(t),$$

then $\gamma(1) = \exp(X)$. We also denote by $e^{tX} : \mathbb{G} \to \mathbb{G}$ the flow of the left-invariant vector field X at time t. Notice that $e^{tX}(p) = p \cdot \exp(tX)$. Integral curves of left-invariant vector fields are said to be *straight curves*.

The Lie group $\mathbb{G}$ is diffeomorphic to $\mathbb{R}^N$ with $N = \sum_{k=1}^{s} \dim(\mathfrak{G}_k)$. A usual way to identify $\mathbb{G}$ and $\mathbb{R}^N$ through a global system of coordinates is to pull-back by exp the group structure from $\mathbb{G}$ to $\mathfrak{G}$, where it can be expressed by the Baker–Campbell–Hausdorff formula. In this way exp becomes a mapping of $\mathfrak{G} = \mathbb{G}$ onto itself that is simply the identity.

For any $\lambda \in \mathbb{R}$ we introduce the dilation $\Delta_\lambda : \mathfrak{G} \to \mathfrak{G}$ uniquely characterized by

$$\begin{cases} \Delta_\lambda([X, Y]) = [\Delta_\lambda(X), \Delta_\lambda(Y)] & \text{for any } X, Y \in \mathfrak{G}, \\ \Delta_\lambda(X) = \lambda X & \text{for any } X \in \mathfrak{G}_1. \end{cases}$$

Using the decomposition $X = X_1 + \cdots + X_s$ with $X_k \in \mathfrak{G}_k$, it holds that $\Delta_\lambda(X) = \sum_{k=1}^{s} \lambda^k X_k$. For any $\lambda \in \mathbb{R}$ we also define on $\mathbb{G}$ the dilation $\delta_\lambda = \exp \circ \Delta_\lambda \circ \exp^{-1}$.

Given an absolutely continuous curve $\gamma : [a, b] \to \mathbb{G}$, the velocity $\dot{\gamma}(t)$, which exists from almost every $t \in [a, b]$, is identified with the element of $\mathfrak{G}$ whose associated left-invariant vector field, evaluated at $\gamma(t)$, is equal to $\dot{\gamma}(t)$. An absolutely continuous curve γ is said to be *horizontal* if $\dot{\gamma}(t) \in \mathfrak{G}_H$ for almost every t. For any interval I of $\mathbb{R}$, we denote by $\mathcal{C}^1_H(I,\mathbb{G})$ the space of all curves $\phi\in\mathcal{C}^1(I,\mathbb{G})$ such that $\dot\phi(t)\in\mathfrak{G}_H$ for every $t\in I$.

Assume that the horizontal layer $\mathfrak{G}_H$ of the algebra is endowed with a quadratic norm $\|\cdot\|_{\mathfrak{G}_H}$. The Carnot–Carathéodory distance $d_{\mathbb{G}}(p,q)$ between two points $p,q\in\mathbb{G}$ is then defined as the minimal length of a horizontal curve connecting p and q; i.e.,

$$d_{\mathbb{G}}(p,q)=\inf\left\{\int_a^b\|\dot\gamma(t)\|_{\mathfrak{G}_H}\,dt\;\middle|\;\gamma:[a,b]\to\mathbb{G}\text{ horizontal},\ \gamma(a)=p,\ \gamma(b)=q\right\}.$$

Note that $d_{\mathbb{G}}$ is left-invariant. It is known that $d_{\mathbb{G}}$ provides the same topology as the usual one on $\mathbb{G}$. Moreover, it is homogeneous; i.e., $d_{\mathbb{G}}(\delta_\lambda p,\delta_\lambda q)=|\lambda|d_{\mathbb{G}}(p,q)$ for any $\lambda\in\mathbb{R}$.

Observe that the Carnot–Carathéodory distance depends on the norm $\|\cdot\|_{\mathfrak{G}_H}$ considered on $\mathfrak{G}_H$. However, all Carnot–Carathéodory distances are in fact metrically equivalent. They are even equivalent with any left-invariant homogeneous distance [Folland and Stein 1982], similar to the way all norms on a finite-dimensional vector space are equivalent.

Notice that $d_{\mathbb{G}}(p,\cdot)$ can be seen as the value function of the optimal control problem

$$\begin{cases}\dot\gamma=\displaystyle\sum_{i=1}^m u_iX_i(\gamma), & (u_1,\dots,u_m)\in\mathbb{R}^m,\\ \gamma(a)=p, & \\ \displaystyle\int_a^b\sqrt{u_1(t)^2+\cdots+u_m(t)^2}\,dt\to\min, & \end{cases}$$

where $X_1,\dots,X_m$ is a $\|\cdot\|_{\mathfrak{G}_H}$-orthonormal basis of $\mathfrak{G}_H$.

Finally, the space $\mathcal{C}^1_H([a,b],\mathbb{G})$ of horizontal curves of class $\mathcal{C}^1$ can be endowed with a natural $\mathcal{C}^1$ metric associated with $(d_{\mathbb{G}},\|\cdot\|_{\mathfrak{G}_H})$ as follows: the distance between two curves γ_1 and γ_2 in $\mathcal{C}^1_H([a,b],\mathbb{G})$ is

$$\max\Bigl(\sup_{t\in[a,b]}d_{\mathbb{G}}(\gamma_1(t),\gamma_2(t)),\ \sup_{t\in[a,b]}\|\dot\gamma_2(t)-\dot\gamma_1(t)\|_{\mathfrak{G}_H}\Bigr).$$

In the following, we will write $\|\dot\gamma_2-\dot\gamma_1\|_{\infty,\mathfrak{G}_H}$ to denote the quantity $\sup_{t\in[a,b]}\|\dot\gamma_2(t)-\dot\gamma_1(t)\|_{\mathfrak{G}_H}$.

Whitney condition. A *homogeneous homomorphism* between two Carnot groups $\mathbb{G}_1$ and $\mathbb{G}_2$ is a group morphism $L:\mathbb{G}_1\to\mathbb{G}_2$ with $L\circ\delta^{\mathbb{G}_1}_\lambda=\delta^{\mathbb{G}_2}_\lambda\circ L$ for any $\lambda\in\mathbb{R}$. Moreover, L is a homogeneous homomorphism if and only if $\exp^{-1}_{\mathbb{G}_2}\circ L\circ\exp_{\mathbb{G}_1}$ is a homogeneous Lie algebra morphism. It is in particular a linear map on $\mathbb{G}_1$ identified with $\mathfrak{G}_{\mathbb{G}_1}$. The first layer is mapped on the first layer so that a homogeneous homomorphism from $\mathbb{R}$ to $\mathbb{G}_2$ has the form $L(t)=\exp_{\mathbb{G}_2}(tX)$, where $X\in\mathfrak{G}^{\mathbb{G}_2}_H$.

Proposition 2.1 (Pansu–Rademacher theorem). *Let f be a locally Lipschitz map from an open subset U of $\mathbb{G}_1$ into $\mathbb{G}_2$. Then for almost every $p\in U$, there exists a homogeneous homomorphism L_p such that*

$$\mathbb{G}_1\ni q\mapsto\delta^{\mathbb{G}_2}_{1/r}\bigl(f(p)^{-1}\cdot f(p\cdot\delta^{\mathbb{G}_1}_r(q))\bigr)\tag{1}$$

tends to L_p uniformly on every compact set $K\subset\mathbb{G}_1$ as r goes to zero.

Note that in Proposition 2.1 the map L_p is uniquely determined. It is called the *Pansu derivative* of f at p and denoted by Df_p.

We denote by $\mathcal{C}^1_H(\mathbb{G}_1,\mathbb{G}_2)$ the space of functions f such that (1) holds at every point $p\in\mathbb{G}_1$ and $p\mapsto Df_p$ is continuous for the usual topology. For $\mathbb{G}_1=\mathbb{R}$ this coincides with the definition of $\mathcal{C}^1_H(I,\mathbb{G}_2)$ given earlier; see [Pansu 1989, Proposition 4.1]. We have the following.

Proposition 2.2 (Taylor expansion). *Let $f\in\mathcal{C}^1_H(\mathbb{G}_1,\mathbb{G}_2)$, where $\mathbb{G}_1$ and $\mathbb{G}_2$ are Carnot groups. Let $K\subset\mathbb{G}_1$ be compact. Then there exists a function ω from $\mathbb{R}^+$ to $\mathbb{R}^+$ with $\omega(t)=o(t)$ at 0^+ such that for any $p,q\in K$,*

$$d_{\mathbb{G}_2}\big(f(q),\,f(p)\cdot Df_p(p^{-1}\cdot q)\big)\le\omega(d_{\mathbb{G}_1}(p,q)),$$

where Df_p is the Pansu derivative.

Proof. This is a direct consequence the mean value inequality by Magnani [2013, Theorem 1.2]. □

The above proposition hints at the suitable formulation of the $\mathcal{C}^1$-Whitney condition for Carnot groups. This generalization already appeared in the literature in [Vodop'yanov and Pupyshev 2006a].

Definition 2.3 ($\mathcal{C}^1_H$-Whitney condition).
- Let K be a compact subset of $\mathbb{G}_1$ and consider $f:K\to\mathbb{G}_2$ and a map L which associates with any $p\in K$ a homogeneous group homomorphism $L(p)$. We say that the $\mathcal{C}^1_H$-*Whitney condition holds* for (f,L) on K if L is continuous and there exists a function ω from $\mathbb{R}^+$ to $\mathbb{R}^+$ with $\omega(t)=o(t)$ at 0^+ such that, for any $p,q\in K$,
$$d_{\mathbb{G}_2}\big(f(q),\,f(p)\cdot L(p)(p^{-1}\cdot q)\big)\le\omega(d_{\mathbb{G}_1}(p,q)).\tag{2}$$
- Let K_0 be a closed set of $\mathbb{G}_1$, $f:K_0\to\mathbb{G}_2$, and L be such that $K_0\ni p\mapsto L(p)$ is continuous. We say that the $\mathcal{C}^1_H$-*Whitney condition holds* for (f,L) on K_0 if for any compact set $K\subset K_0$ it holds for the restriction of (f,L) to K.

Of course, according to Proposition 2.2, if $f\in\mathcal{C}^1_H(\mathbb{G}_1,\mathbb{G}_2)$, then the restriction of (f,Df) to any closed K_0 satisfies the $\mathcal{C}^1_H$-Whitney condition on K_0.

In this paper we focus on the case $\mathbb{G}_1=\mathbb{R}$. The condition on a compact set K is $r_{K,\eta}\to0$ as $\eta\to0$, where

$$r_{K,\eta}=\sup_{\substack{\tau,t\in K\\0<|\tau-t|<\eta}}\frac{d_{\mathbb{G}_2}\big(f(t),\,f(\tau)\cdot\exp[(t-\tau)X(\tau)]\big)}{|\tau-t|},\tag{3}$$

because for every $\tau\in\mathbb{R}$ one has $[L(\tau)](h)=\exp(hX(\tau))$ for some $X(\tau)\in\mathfrak{G}^{\mathbb{G}_2}_H$ and every $h\in\mathbb{R}$. With a slight abuse of terminology, we say that the $\mathcal{C}^1_H$-Whitney condition holds for (f,X) on K.

In the classical setting, the Whitney condition is equivalent to the existence of a $\mathcal{C}^1$ map $\bar f:\mathbb{R}^{n_1}\to\mathbb{R}^{n_2}$ such that $\bar f$ and $D\bar f$ have respectively restrictions f and L on K. This property is usually known as the $\mathcal{C}^1$-*Whitney extension theorem* or simply the *Whitney extension theorem*, as for instance in [Evans and Gariepy 2015], even though the original theorem by Whitney [1934a; 1934b] is more general and in particular includes higher-order extensions and considers the extension $f\to\bar f$ as a linear operator. This theorem is of broad use in analysis and is still the subject of dedicated research. See, for instance, [Brudnyi and Shvartsman 1994; Fefferman 2005; Fefferman et al. 2014].

Definition 2.4. We say that the pair $(\mathbb{G}_1, \mathbb{G}_2)$ *has the $\mathcal{C}^1_H$ extension property* if for every (f, L) satisfying the $\mathcal{C}^1_H$-Whitney condition on some closed set K_0 there exists $\bar{f} \in \mathcal{C}^1_H(\mathbb{G}_1, \mathbb{G}_2)$ which extends f on $\mathbb{G}_1$ and such that $D\bar{f}_p = L(p)$ for every $p \in K_0$.

We now state the $\mathcal{C}^1_H$-extension theorem that Franchi, Serapioni, and Serra Cassano proved in [Franchi et al. 2003, Theorem 2.14]. It has been generalized by Vodop'yanov and Pupyshev [2006a; 2006b] to a form closer to Whitney's original result including higher-order extensions and the linearity of the operator $f \mapsto \bar{f}$.

Theorem 2.5 (Franchi, Serapioni, Serra Cassano). *For any Carnot group $\mathbb{G}_1$ and any $d \in \mathbb{N}$, the pair $(\mathbb{G}_1, \mathbb{R}^d)$ has the $\mathcal{C}^1_H$ extension property.*

The proof proposed by Franchi, Serapioni, and Serra Cassano is established for Carnot groups of step 2 only, but is identical for general Carnot groups. It is inspired by the proof in [Evans and Gariepy 2015], which corresponds to the special case $\mathbb{G}_1 = \mathbb{R}^{n_1}$ for $n_1 \geq 1$.

Let us mention an example from the literature of nonextension with $\mathbb{G}_1 \neq \mathbb{R}$. This remarkable fact was explained to us by A. Kozhevnikov.

Example 2.6. If $\mathbb{G}_1$ and $\mathbb{G}_2$ are the ultrarigid Carnot groups of dimensions 17 and 16 respectively, presented in [Le Donne et al. 2014] and analyzed in Lemma A.2.1 of [Kozhevnikov 2015], one can construct an example (f, L) satisfying the $\mathcal{C}^1_H$-Whitney condition on some compact K without any possible extension $(\bar{f}, D\bar{f})$ on $\mathbb{G}_1$. For this, one exploits the rarity of $\mathcal{C}^1_H$ maps of maximal rank in ultrarigid Carnot groups. The definition of ultrarigid from [Le Donne et al. 2014, Definition 3.1] is that all quasimorphisms are Carnot similitudes, i.e., compositions of dilations and left-translations. We do not use here directly the definition of ultrarigid groups but just the result stated in Lemma A.2.1 of [Kozhevnikov 2015] for $\mathbb{G}_1$ and $\mathbb{G}_2$. Concretely, let us set

$$K = \left\{(p_1, \dots, p_{17}) \in \mathbb{G}_1 \;\middle|\; p_2 = \cdots = p_{16} = 0,\;\; p_1 \in [-1, 1],\;\; p_{17} = p_1\right\}.$$

Let the map f be constantly equal to 0 on K and L be the constant projection

$$\Lambda : \mathbb{G}_1 \to \mathbb{G}_2, \qquad (q_1, \dots, q_{17}) \mapsto (q_1, \dots, q_{16}).$$

Lemma A.2.1 in [Kozhevnikov 2015] applied at the point $0_{\mathbb{G}_1}$ implies that the only possible extension of f is the projection $L(0) = \Lambda$. But this map vanishes only on $\{p \in \mathbb{G}_1 \mid p_1 = \cdots = p_{16} = 0\}$, which does not contain K. It remains for us to prove that Whitney's condition holds. In fact for two points $p = (x, 0_{\mathbb{R}^{15}}, x)$ and $q = (y, 0_{\mathbb{R}^{15}}, y)$ in K, we look at the distance from $f(x) = 0_{\mathbb{G}_2}$ to

$$f(p) \cdot L(0)(p^{-1} \cdot q) = L(0)\big((x, 0_{\mathbb{R}^{15}}, x)^{-1} \cdot (y, 0_{\mathbb{R}^{15}}, y)\big) = (y - x, 0_{\mathbb{R}^{15}})$$

on the one side and from p to q on the other side. The first one is $|y - x|$, up to a multiplicative constant, and when $|y - x|$ goes to zero, the second one is $c|y - x|^{1/3}$ for some constant $c > 0$. This proves the $\mathcal{C}^1_H$-Whitney condition for (f, L) on K.

In the present paper we provide examples of ordered pairs $(\mathbb{G}_1, \mathbb{G}_2)$ with $\mathbb{G}_1 = \mathbb{R}$ such that the $\mathcal{C}^1_H$ extension property does or does not hold, depending on the geometry of $\mathbb{G}_2$. We do not address the problem of Whitney extensions for orders larger than 1. A preliminary step for considering higher-order extensions would be to provide a suitable Taylor expansion for $\mathcal{C}^m_H$-functions from $\mathbb{R}$ to $\mathbb{G}_2$, in the spirit of what was recalled for $m = 1$ in Proposition 2.2.

Let us conclude the section by assuming that the $\mathcal{C}^1_H$ extension property holds for some ordered pair $(\mathbb{G}_1, \mathbb{G}_2)$ of Carnot groups and by showing how to deduce it for other pairs. We describe below three such possible implications:

(1) Let $\mathbb{S}_1$ be a homogeneous subgroup of $\mathbb{G}_1$ that admits a complementary group $\mathbb{K}$ in the sense of [Serra Cassano 2016, Section 4.1.2]: both $\mathbb{S}_1$ and $\mathbb{K}$ are homogeneous Lie groups and the intersection is reduced to $\{0\}$. Assume moreover that $\mathbb{S}_1$ is a Carnot group and $\mathbb{K}$ is normal, so that one can define canonically a projection $\pi : \mathbb{G}_1 \to \mathbb{S}_1$ that is a homogeneous homomorphism. Moreover, π is Lipschitz continuous; see [Serra Cassano 2016, Proposition 4.13]. For the rest of the section, we say that $\mathbb{S}_1$ is an *appropriate Carnot subgroup of* $\mathbb{G}_1$. It can be easily proved that $(\mathbb{S}_1, \mathbb{G}_2)$ has the $\mathcal{C}^1_H$ extension property.

(2) Assume now that $\mathbb{S}_2$ is an appropriate Carnot subgroup of $\mathbb{G}_2$. Using the Lipschitz continuity of the projection $\pi : \mathbb{G}_2 \to \mathbb{S}_2$, one easily deduces from the definition of the $\mathcal{C}^1_H$-Whitney condition that $(\mathbb{G}_1, \mathbb{S}_2)$ has the $\mathcal{C}^1_H$ extension property.

(3) Finally assume that $(\mathbb{G}_1, \mathbb{G}'_2)$ has the $\mathcal{C}^1_H$ extension property, where $\mathbb{G}'_2$ is a Carnot group. Then one checks without difficulty that the same is true for $(\mathbb{G}_1, \mathbb{G}_2 \times \mathbb{G}'_2)$.

As a consequence of Theorem 1.1, we can use these three implications to infer pliability statements. Namely, a Carnot group $\mathbb{G}$ is pliable if (i) $(\mathbb{G}_0, \mathbb{G})$ has the $\mathcal{C}^1_H$ extension property for some Carnot group $\mathbb{G}_0$ of positive dimension, (ii) $\mathbb{G}$ is the appropriate Carnot subgroup of a pliable Carnot group, and (iii) $\mathbb{G}$ is the product of two pliable Carnot groups.[1]

3. Rigidity, a necessary condition for the $\mathcal{C}^1_H$ extension property

Let us first adapt to the case of horizontal curves on Carnot groups the notion of rigid curve introduced by Bryant and Hsu [1993]. We will show in the following that the existence of rigid curves in a Carnot group $\mathbb{G}$ can be used to identify obstructions to the validity of the $\mathcal{C}^1_H$ extension property for $(\mathbb{R}, \mathbb{G})$.

Definition 3.1 (Bryant, Hsu). Let $\gamma \in \mathcal{C}^1_H([a, b], \mathbb{G})$. We say that γ *is rigid* if there exists a neighborhood $\mathcal{V}$ of γ in the space $\mathcal{C}^1_H([a, b], \mathbb{G})$ such that if $\beta \in \mathcal{V}$ and $\gamma(a) = \beta(a)$, $\gamma(b) = \beta(b)$ then β is a reparametrization of γ.

A vector $X \in \mathfrak{G}_H$ is said to be *rigid* if the curve $[0, 1] \ni t \mapsto \exp(tX)$ is rigid.

A celebrated existence result of rigid curves for general sub-Riemannian manifolds has been obtained by Bryant and Hsu [1993] and further improved in [Liu and Sussman 1995; Agrachev and Sarychev 1996]. Examples of Carnot groups with rigid curves have been illustrated in [Golé and Karidi 1995] and extended in [Huang and Yang 2012], where it is shown that, for any $N \geq 6$, there exists a Carnot group of

[1] Added in print: both the notions of Carnot subgroup and Carnot quotient group can be defined, based on the definitions of Lie subgroup and Lie quotient group. The additional structure to care about is the grading and the fact that the first layer generates the Lie algebra. Appropriate Carnot subgroups are in fact those groups that are both Carnot subgroups and Carnot quotient groups. As E. Le Donne pointed out to us, the only necessary setup for (2) and (ii) is the quotient structure.

topological dimension N having rigid curves. Nevertheless, such curves need not be straight. Actually, the construction proposed in [Huang and Yang 2012] produces curves which are necessarily not straight.

Following [Agrachev and Sarychev 1996], see also [Montgomery 1995], and focusing on rigid straight curves in Carnot groups, we can formulate Theorem 3.2 below. In order to state it, let $\pi : T^*\mathbb{G} \to \mathbb{G}$ be the canonical projection and recall that a curve $p : I \to T^*\mathbb{G}$ is said to be an *abnormal path* if $\pi \circ p : I \to \mathbb{G}$ is a horizontal curve, $p(t) \neq 0$ and $p(t)X = 0$ for every $t \in I$ and $X \in \mathfrak{G}_H$, and, moreover, for every $Y \in \mathfrak{G}$ and almost every $t \in I$,

$$\frac{d}{dt} p(t)Y = p(t)[Z(t), Y], \tag{4}$$

where

$$Z(t) = \frac{d}{dt}\pi \circ p(t) \in \mathfrak{G}_H.$$

Theorem 3.2. *Let $X \in \mathfrak{G}_H$ and assume that $p : [0, 1] \to T^*\mathbb{G}$ is an abnormal path with $\pi \circ p(t) = \exp(tX)$.*

If $t \mapsto \exp(tX)$ is rigid, then $p(t)[V, W] = 0$ for every $V, W \in \mathfrak{G}_H$ and every $t \in [0, 1]$. Moreover, denoting by $Q_{p(t)}$ the quadratic form $Q_{p(t)}(V) = p(t)[V, [X, V]]$ defined on $\{V \in \mathfrak{G}_H \mid V \perp X\}$, we have $Q_{p(t)} \geq 0$ for every $t \in [0, 1]$.

Conversely, if $p(t)[V, W] = 0$ for every $V, W \in \mathfrak{G}_H$ and every $t \in [0, 1]$ and $Q_{p(t)} > 0$ for every $t \in [0, 1]$ then $t \mapsto \exp(tX)$ is rigid.

Example 3.3. An example of Carnot structure having rigid straight curves is the standard Engel structure. In this case $s = 3$, $\dim \mathfrak{G}_1 = 2$, $\dim \mathfrak{G}_2 = \dim \mathfrak{G}_3 = 1$ and one can pick two generators X, Y of the horizontal distribution whose only nontrivial bracket relations are $[X, Y] = W_1$ and $[Y, W_1] = W_2$, where W_1 and W_2 span $\mathfrak{G}_2$ and $\mathfrak{G}_3$ respectively.

Let us illustrate how the existence of rigid straight curves can be deduced from Theorem 3.2 (one could also prove rigidity by direct computations of the same type as those of Example 3.5 below).

One immediately checks that p with $p(t)X = p(t)Y = p(t)W_1 = 0$ and $p(t)W_2 = 1$ is an abnormal path such that $\pi \circ p(t) = \exp(tX)$. The rigidity of $t \mapsto \exp(tX)$ then follows from Theorem 3.2, thanks to the relation $Q_p(Y) = 1$.

An extension of the previous construction can be used to exhibit, for every $N \geq 4$, a Carnot group of topological dimension N and step $N-1$ having straight rigid curves. It suffices to consider the N-dimensional Carnot group with Goursat distribution, that is, the group such that $\dim \mathfrak{G}_1 = 2$, $\dim \mathfrak{G}_i = 1$ for $i = 2, \ldots, N-1$, and there exist two generators X, Y of $\mathfrak{G}_1$ whose only nontrivial bracket relations are $[X, Y] = W_1$ and $[Y, W_i] = W_{i+1}$ for $i = 1, \ldots, N-3$, where $\mathfrak{G}_{i+1} = \mathrm{Span}(W_i)$ for $i = 1, \ldots, N-2$.

The following definition introduces the notion of *pliable* horizontal curve, in contrast to a rigid one.

Definition 3.4. We say that a curve $\gamma \in \mathcal{C}^1_H([a, b], \mathbb{G})$ is *pliable* if for every neighborhood $\mathcal{V}$ of γ in $\mathcal{C}^1_H([a, b], \mathbb{G})$ the set

$$\{(\beta(b), \dot\beta(b)) \mid \beta \in \mathcal{V},\ (\beta, \dot\beta)(a) = (\gamma, \dot\gamma)(a)\}$$

is a neighborhood of $(\gamma(b), \dot\gamma(b))$ in $\mathbb{G} \times \mathfrak{G}_H$.

A vector $X \in \mathfrak{G}_H$ is said to be *pliable* if the curve $[0, 1] \ni t \mapsto \exp(tX)$ is pliable.

We say that $\mathbb{G}$ *is pliable* if every vector $X \in \mathfrak{G}_H$ is pliable.

By metric equivalence of all Carnot–Carathéodory distances, it follows that the pliability of a horizontal vector does not depend on the norm $\|\cdot\|_{\mathfrak{G}_H}$ considered on $\mathfrak{G}_H$.

Notice that, by the definition of pliability, in every $\mathcal{C}^1_H$ neighborhood of a pliable curve $\gamma : [a, b] \to \mathbb{G}$ there exists a curve β with $\beta(a) = \gamma(a)$, $(\beta, \dot\beta)(b) = (\gamma(b), W)$, and $W \neq \dot\gamma(b)$. This shows that pliable curves are not rigid. It should be noticed, however, that the converse is not true in general, as will be discussed in Example 3.5. In this example we show that there exist horizontal straight curves that are neither rigid nor pliable.

Example 3.5. We consider the 6-dimensional Carnot algebra $\mathfrak{G}$ of step 3 that is spanned by X, Y, Z, $[X, Z]$, $[Y, Z]$, and $[Y, [Y, Z]]$, where X, Y, Z is a basis of $\mathfrak{G}_1$ and, except for permutations, all brackets different from the ones above are zero.

According to [Bonfiglioli et al. 2007, Chapter 4] there is a group structure on $\mathbb{R}^6$ with coordinates (x, y, z, z_1, z_2, z_3) isomorphic to the corresponding Carnot group $\mathbb{G}$ such that the vectors of $\mathfrak{G}_1$ are the left-invariant vector fields

$$X = \partial_x, \quad Y = \partial_y, \quad Z = \partial_z + x\,\partial_{z_1} + y\,\partial_{z_2} + y^2\,\partial_{z_3}.$$

Consider the straight curve $[0, 1] \ni t \mapsto \gamma(t) = \exp(tZ) \in \mathbb{G}$. First notice that γ *is not pliable*, since for all horizontal curves in a small enough $\mathcal{C}^1$ neighborhood of γ the component of the derivative along Z is positive, which implies that the coordinate z_3 is nondecreasing. No endpoint of a horizontal curve starting from $0_{\mathbb{G}}$ and belonging to a small enough $\mathcal{C}^1$ neighborhood of γ can have negative z_3-component.

Let us now show that γ *is not rigid* either. Consider the solution β of

$$\dot\beta(t) = Z(\beta(t)) + u(t)X(\beta(t)), \quad \beta(0) = 0_{\mathbb{G}}.$$

Notice that the y-component of β is identically equal to zero. As a consequence, the same is true for the components z_2 and z_3, while the x-, z- and z_1-components of $\beta(t)$ are, respectively, $\int_0^t u(\tau)\,d\tau$, t, and $\int_0^t \int_0^\tau u(\theta)\,d\theta\,d\tau$. In order to disprove the rigidity, it is then sufficient to take a nontrivial continuous $u : [0, 1] \to \mathbb{R}$ such that $\int_0^1 u(\tau)\,d\tau = 0 = \int_0^1 \int_0^\tau u(\theta)\,d\theta\,d\tau$.

Let us list some useful manipulations which transform horizontal curves into horizontal curves. Let γ be a horizontal curve defined on $[0, 1]$ and such that $\gamma(0) = 0_{\mathbb{G}}$.

(T1) For every $\lambda > 0$, the curve $t \in [0, \lambda] \mapsto \delta_\lambda \circ \gamma(\lambda^{-1}t)$ is horizontal and its velocity at time t is $\dot\gamma(\lambda^{-1}t)$.

(T2) For every $\lambda < 0$, the curve $t \in [0, |\lambda|] \mapsto \delta_\lambda \circ \gamma(|\lambda|^{-1}t)$ is horizontal and its velocity at time t is $-\dot\gamma(|\lambda|^{-1}t)$.

(T3) The curve $\bar\gamma$ defined by $\bar\gamma(t) = \gamma(1)^{-1} \cdot \gamma(1-t)$ is horizontal. It starts in $0_{\mathbb{G}}$ and finishes in $\gamma^{-1}(1)$. Its velocity at time t is $-\dot\gamma(1-t)$.

(T4) If one composes the (commuting) transformations (T2) with $\lambda = -1$ and (T3), one obtains a curve with derivative $\dot\gamma(1-t)$ at time t.

(T5) It is possible to define the concatenation of two curves $\gamma_1 : [0, t_1] \to \mathbb{G}$ and $\gamma_2 : [0, t_2] \to \mathbb{G}$ both starting from $0_{\mathbb{G}}$ as follows: the concatenated curve $\tilde{\gamma} : [0, t_1 + t_2] \to \mathbb{G}$ satisfies $\tilde{\gamma}(0) = 0_{\mathbb{G}}$, has the same velocity as γ_1 on $[0, t_1]$ and the velocity of $\gamma_2(\cdot - t_1)$ on $[t_1, t_1 + t_2]$. We have $\tilde{\gamma}(t_1 + t_2) = \gamma_1(t_1) \cdot \gamma(t_2)$ as a consequence of the invariance of Lie algebra for the left-translation.

A consequence of (T1) and (T2) is that $X \in \mathfrak{G}_H \setminus \{0\}$ is rigid if and only if λX is rigid for every $\lambda \in \mathbb{R} \setminus \{0\}$. Similarly, $X \in \mathfrak{G}_H$ is pliable if and only if λX is pliable for every $\lambda \in \mathbb{R} \setminus \{0\}$.

Proposition 3.7 below gives a characterization of pliable horizontal vectors in terms of a condition which is a priori easier to check than the one appearing in Definition 3.4. Before proving the proposition, let us give a technical lemma. From now on, we write $\mathcal{B}_{\mathbb{G}}(x, r)$ to denote the ball of center x and radius r in $\mathbb{G}$ for the distance $d_{\mathbb{G}}$ and, similarly, $\mathcal{B}_{\mathfrak{G}_H}(x, r)$ to denote the ball of center x and radius r in $\mathfrak{G}_H$ for the norm $\|\cdot\|_{\mathfrak{G}_H}$.

Lemma 3.6. *For any $x \in \mathbb{G}$ and $0 < r < R$, there exists $\varepsilon > 0$ such that if $y, z \in \mathbb{G}$ and $\rho \geq 0$ satisfy $d_{\mathbb{G}}(y, 0_{\mathbb{G}}), d_{\mathbb{G}}(z, 0_{\mathbb{G}}), \rho \leq \varepsilon$, then*

$$\mathcal{B}_{\mathbb{G}}(x, r) \subset y \cdot \delta_{1-\rho}(\mathcal{B}_{\mathbb{G}}(x, R)) \cdot z.$$

Proof. Assume, by contradiction, that for every $n \in \mathbb{N}$ there exist $x_n \in \mathcal{B}_{\mathbb{G}}(x, r)$, $y_n, z_n \in \mathcal{B}_{\mathbb{G}}(0_{\mathbb{G}}, 1/n)$ and $\rho_n \in [0, 1/n]$ such that

$$x_n \notin y_n \cdot \delta_{1-\rho_n}(\mathcal{B}_{\mathbb{G}}(x, R)) \cdot z_n.$$

Equivalently,

$$\delta_{(1-\rho_n)^{-1}}(y_n^{-1} \cdot x_n \cdot z_n^{-1}) \notin \mathcal{B}_{\mathbb{G}}(x, R).$$

However, $\limsup_{n\to\infty} d_{\mathbb{G}}\big(x, \delta_{(1-\rho_n)^{-1}}(y_n^{-1} \cdot x_n \cdot z_n^{-1})\big) \leq r$, leading to a contradiction. □

Proposition 3.7. *A vector $V \in \mathfrak{G}_H$ is pliable if and only if for every neighborhood $\mathcal{V}$ of the curve $[0, 1] \ni t \mapsto \exp(tV)$ in the space $C^1_H([0, 1], \mathbb{G})$, the set*

$$\{\beta(1) \mid \beta \in \mathcal{V},\ (\beta, \dot{\beta})(0) = (0_{\mathbb{G}}, V)\}$$

is a neighborhood of $\exp(V)$.

Proof. Let

$$\mathcal{F} : C^1_H([0, 1], \mathbb{G}) \to \mathbb{G} \times \mathfrak{G}_H, \quad \beta \mapsto (\beta, \dot{\beta})(1),$$

and denote by $\pi : \mathbb{G} \times \mathfrak{G}_H \to \mathbb{G}$ the canonical projection.

One direction of the equivalence being trivial, let us take $\varepsilon > 0$ and assume that $\pi \circ \mathcal{F}(\mathcal{U}_\varepsilon)$ is a neighborhood of $\exp(V)$ in $\mathbb{G}$, where

$$\mathcal{U}_\varepsilon = \big\{\beta \in C^1_H([0, 1], \mathbb{G}) \,\big|\, (\beta, \dot{\beta})(0) = (0_{\mathbb{G}}, V),\ \|\dot{\beta} - V\|_{\infty, \mathfrak{G}_H} < \varepsilon\big\}.$$

We should prove that $\mathcal{F}(\mathcal{U}_\varepsilon)$ is a neighborhood of $(\exp(V), V)$ in $\mathbb{G} \times \mathfrak{G}_H$.

Step 1: As an intermediate step, we first prove that there exists $\eta > 0$ such that $\mathcal{B}_{\mathbb{G}}(\exp(V), \eta) \times \{V\}$ is contained in $\mathcal{F}(\mathcal{U}_\varepsilon)$.

Let ρ be a real parameter in $(0, 1)$. Using the transformations among horizontal curves described earlier in this section, let us define a map $T_\rho : \mathcal{U}_\varepsilon \to \mathcal{C}^1_H([0, 1], \mathbb{G})$ associating with a curve $\gamma \in \mathcal{U}_\varepsilon$ the concatenation, i.e., transformation (T5), of $\gamma_1 : t \mapsto \delta_\rho \circ \gamma(\rho^{-1}t)$ on $[0, \rho]$ obtained by transformation (T1) and a curve γ_2 defined as follows. Consider $\gamma_{2,1} : [0, 1-\rho] \ni t \mapsto \delta_{1-\rho} \circ \gamma((1-\rho)^{-1}t)$, again using (T1). The curve γ_2 is defined from $\gamma_{2,1}$ by

$$\gamma_2(t) = \gamma_1(\rho) \cdot \big(\gamma_{2,1}(1-\rho)^{-1} \cdot \delta_{-1} \circ \gamma_{2,1}((1-\rho) - t)\big);$$

see transformation (T4). The derivative of $T_\rho(\gamma)$ at time $t \in [0, \rho)$ is $\dot\gamma(\rho^{-1}t)$. Its derivative at time $\rho + t$ is $\dot\gamma(1 - (1-\rho)^{-1}t)$ for $t \in (0, 1-\rho]$. Hence $T_\rho(\gamma)$ is continuous and has derivative $\dot\gamma(1)$ at limit times ρ^- and ρ^+; i.e., it is a well-defined map from $\mathcal{U}_\varepsilon$ into $\mathcal{C}^1_H([0, 1], \mathbb{G})$. Moreover, $T_\rho(\gamma)$ has the same derivative $V = \dot\gamma(0)$ at times 0 and 1 and its derivative at any time in $[0, 1]$ is in the set of the derivatives of γ. In particular, $T_\rho(\mathcal{U}_\varepsilon) \subset \mathcal{U}_\varepsilon$.

Notice now that, by construction, the endpoint $T_\rho(\gamma)(1)$ of the curve $T_\rho(\gamma)$ is a function of $\gamma(1)$ and ρ only. It is actually equal to

$$F_\rho(x) = \delta_\rho(x) \cdot \delta_{\rho-1}(x)^{-1},$$

where $x = \gamma(1)$; see (T1) and (T4). Let $x_0 = \exp(V)$ and $\gamma_0 : t \mapsto \exp(tV)$. We have $F_\rho(x_0) = x_0$ because $T_\rho(\gamma_0) = \gamma_0$, both curves having derivative constantly equal to V. We prove now that for ρ close enough to 1, the differential of F_ρ at x_0 is invertible. Let us use the coordinate identification of $\mathbb{G}$ with $\mathbb{R}^N$. For every $y \in \mathbb{G}$, the limits of $\delta_\rho(y)$ and $\delta_{1-\rho}(y)$ as ρ tends to 1 are y and $0_\mathbb{G}$ respectively, while $D\delta_\rho(y)$ and $D\delta_{\rho-1}(y)$ converge to Id and 0 respectively. One can check — see, e.g., [Bonfiglioli et al. 2007, Proposition 2.2.22] — that the inverse function has derivative $-\mathrm{Id}$ at $0_\mathbb{G}$. Finally the left and right translations are global diffeomorphisms. Collecting this information and applying the chain rule, we get that $DF_\rho(x_0)$ tends to an invertible operator as ρ goes to 1. Hence for ρ great enough, $F_\rho(x_0)$ is a local diffeomorphism.

We know by assumption on V that, for any $\varepsilon > 0$, the endpoints of the curves of $\mathcal{U}_\varepsilon$ form a neighborhood of x_0. We have shown that this is also the case if we replace $\mathcal{U}_\varepsilon$ by $T_\rho(\mathcal{U}_\varepsilon)$ for ρ close to 1. The curves of $T_\rho(\mathcal{U}_\varepsilon)$ are in $\mathcal{U}_\varepsilon$ and have, moreover, derivative V at time 1. We have thus proved that for every $\varepsilon > 0$ there exists $\eta > 0$ such that $\mathcal{B}_\mathbb{G}(x_0, \eta) \times \{V\}$ is contained in $\mathcal{F}(\mathcal{U}_\varepsilon)$.

Step 2: Let us now prove that $\mathcal{F}(\mathcal{U}_\varepsilon)$ is a neighborhood of (x_0, V) in $\mathbb{G} \times \mathfrak{G}_H$.

Let β be a curve in $\mathcal{U}_\varepsilon$ with $\dot\beta(1) = V$ and consider for every $W \in \mathcal{B}_{\mathfrak{G}_H}(V, \varepsilon)$ and every $\rho \in (0, 1)$ the curve $\alpha_{\rho,W}$ defined as follows: $\alpha_{\rho,W} = \delta_{1-\rho} \circ \beta((1-\rho)^{-1}t)$ on $[0, 1-\rho]$ (transformation (T1)) and $\dot\alpha_{\rho,W}$ is the linear interpolation between V and W on $[1-\rho, 1]$. Notice that $\alpha_{\rho,W}$ is in $\mathcal{U}_\varepsilon$.

Let $u \in \mathbb{G}$ be the endpoint at time ρ of the curve in $\mathbb{G}$ starting at $0_\mathbb{G}$ whose derivative is the linear interpolation between V and W on $[0, \rho]$. Then $(\alpha_{\rho,W}, \dot\alpha_{\rho,W})(1) = (\delta_{1-\rho}(\beta(1)) \cdot u, W)$ and u depends only on V, ρ and W, and not on the curve β. Moreover, u tends to $0_\mathbb{G}$ as ρ goes to 1, uniformly with respect to $W \in \mathcal{B}_{\mathfrak{G}_H}(V, \varepsilon)$. Lemma 3.6 implies that for ρ sufficiently close to 1, for every $W \in \mathcal{B}_{\mathfrak{G}_H}(V, \varepsilon)$, it holds that $\delta_{1-\rho}(\mathcal{B}_\mathbb{G}(x_0, \eta)) \cdot u \supset \mathcal{B}_\mathbb{G}\big(x_0, \frac{1}{2}\eta\big)$. We proved that $\mathcal{B}_\mathbb{G}\big(x_0, \frac{1}{2}\eta\big) \times \mathcal{B}_{\mathfrak{G}_H}(V, \varepsilon) \subset \mathcal{F}(\mathcal{U}_\varepsilon)$, concluding the proof of the proposition. $\square$

The main result of this section is the following theorem, which constitutes the necessary part of the characterization of $\mathcal{C}^1_H$ extendability stated in Theorem 1.1.

Theorem 3.8. *Let $\mathbb{G}$ be a Carnot group. If $(\mathbb{R}, \mathbb{G})$ has the $\mathcal{C}^1_H$ extension property, then $\mathbb{G}$ is pliable.*

Proof. Suppose, by contradiction, that there exists $V \in \mathfrak{G}_H$ which is not pliable. We are going to prove that $(\mathbb{R}, \mathbb{G})$ does not have the $\mathcal{C}^1_H$ extension property.

Let $\gamma(t) = \exp(tV)$ for $t \in [0, 1]$. Since V is not pliable, it follows from Proposition 3.7 that there exist a neighborhood $\mathcal{V}$ of γ in the space $\mathcal{C}^1_H([0, 1], \mathbb{G})$ and a sequence $(x_n)_{n\geq 1}$ converging to $0_{\mathbb{G}}$ such that for every $n \geq 1$ no curve β in $\mathcal{V}$ satisfies $(\beta(0), \dot\beta(0)) = (0_{\mathbb{G}}, V)$ and $\beta(1) = \gamma(1)\cdot x_n$. In particular, there exists a neighborhood Ω of V in $\mathfrak{G}_H$ such that for every $\beta \in \mathcal{C}^1_H([0, 1], \mathbb{G})$ with $(\beta(0), \dot\beta(0)) = (0_{\mathbb{G}}, V)$ and

$$\dot\beta(t) \in \Omega \quad \forall t \in [0, 1],$$

we have $(\beta(1), \dot\beta(1)) \neq (\gamma(1) \cdot x_n, V)$ for every $n \in \mathbb{N}$. Since $\lim_{n\to\infty} x_n = 0_{\mathbb{G}}$, we can assume without loss of generality that, for every $n \geq 1$,

$$\max\Big\{d\big(\delta_\rho(x_n) \cdot \exp(tV), \exp(tV)\big) \;\Big|\; \rho \in [0, 1],\ t \in [-1, 1]\Big\} \leq 2^{-n}. \tag{5}$$

By homogeneity and left-invariance, we deduce that for every $y \in \mathbb{G}$ and every $\rho > 0$, for every $\beta \in \mathcal{C}^1([0, \rho], \mathbb{G})$ with $(\beta(0), \dot\beta(0)) = (y, V)$ and

$$\dot\beta(t) \in \Omega \quad \forall t \in [0, \rho],$$

we have $(\beta(\rho), \dot\beta(\rho)) \neq (y \cdot \gamma(\rho) \cdot \delta_\rho(x_n), V)$ for every $n \in \mathbb{N}$.

Define

$$\rho_n = \frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}$$

and $\tilde x_n = \delta_{\rho_n}(x_n)$ for every $n \in \mathbb{N}$. It follows from (5) that

$$\max\Big\{d(\tilde x_n \cdot \exp(tV), \exp(tV)) \;\Big|\; t \in [-1, 1]\Big\} \leq 2^{-n} \quad \forall n \geq 1. \tag{6}$$

We introduce the sequence defined recursively by $y_0 = 0_{\mathbb{G}}$ and

$$y_{n+1} = y_n \cdot \gamma(\rho_n) \cdot \tilde x_n. \tag{7}$$

Notice that $(y_n)_{n\geq 1}$ is a Cauchy sequence and denote by y_∞ its limit as $n \to \infty$.

By construction, for every $n \in \mathbb{N}$ and every $\beta \in \mathcal{C}^1_H([0, \rho_n], \mathbb{G})$ with $(\beta(0), \dot\beta(0)) = (y_n, V)$ and $\dot\beta(t) \in \Omega$ for all $t \in [0, \rho_n]$, we have $(\beta(\rho_n), \dot\beta(\rho_n)) \neq (y_{n+1}, V)$. The proof that $(\mathbb{R}, \mathbb{G})$ does not have the $\mathcal{C}^1_H$ extension property is then concluded if we show that the $\mathcal{C}^1_H$-Whitney condition holds for (f, X) on K, where

$$K = \left(\bigcup_{n=1}^{\infty} \left\{1 - \frac{1}{n}\right\}\right) \cup \{1\},$$

and $f : K \to \mathbb{G}$ and $X : K \to \mathfrak{G}_H$ are defined by

$$f(1 - n^{-1}) = y_n, \quad X(1 - n^{-1}) = V, \quad n \in \mathbb{N}^* \cup \{\infty\}.$$

For $i, j \in \mathbb{N}^* \cup \{\infty\}$, let

$$\begin{aligned} D(i, j) &= d_{\mathbb{G}}\big(f(1-i^{-1}), f(1-j^{-1}) \cdot \exp[(j^{-1}-i^{-1})X(1-j^{-1})]\big) \\ &= d_{\mathbb{G}}\big(y_i, y_j \cdot \exp[(j^{-1}-i^{-1})V]\big). \end{aligned}$$

We have to prove that

$$D(i, j) = o(j^{-1}-i^{-1})$$

as $i, j \to \infty$; that is, for every $\varepsilon > 0$, there exists $i_\varepsilon \in \mathbb{N}^*$ such that $D(i, j) < \varepsilon |j^{-1}-i^{-1}|$ for $i, j \in \mathbb{N}^* \cup \{\infty\}$ with $i, j > i_\varepsilon$.

By the triangle inequality we have

$$D(i, j) \le \sum_{k=\min(i,j)}^{\max(i,j)-1} d_{\mathbb{G}}\big(y_{k+1} \cdot \exp[((k+1)^{-1}-i^{-1})V], y_k \cdot \exp[(k^{-1}-i^{-1})V]\big).$$

Notice that

$$\begin{aligned} &d_{\mathbb{G}}\big(y_{k+1} \cdot \exp[((k+1)^{-1}-i^{-1})V], y_k \cdot \exp[(k^{-1}-i^{-1})V]\big) \\ &\qquad = d_{\mathbb{G}}\big(y_{k+1} \cdot \exp[((k+1)^{-1}-i^{-1})V], [y_k \cdot \gamma(\rho_k)] \cdot \exp[((k+1)^{-1}-i^{-1})V]\big) \\ &\qquad = d_{\mathbb{G}}\big(\tilde{x}_k \cdot \exp[((k+1)^{-1}-i^{-1})V], \exp[((k+1)^{-1}-i^{-1})V]\big), \end{aligned}$$

where the last equality follows from (7) and the invariance of $d_{\mathbb{G}}$ by left-multiplication. Thanks to (6), one then concludes that

$$d_{\mathbb{G}}\big(y_{k+1} \cdot \exp[((k+1)^{-1}-i^{-1})V], y_k \cdot \exp[(k^{-1}-i^{-1})V]\big) \le 2^{-k}.$$

Hence,

$$D(i, j) \le \sum_{k=\min(i,j)}^{\max(i,j)-1} 2^{-k} = o(j^{-1}-i^{-1}). \qquad \square$$

4. Sufficient condition for the $\mathcal{C}^1_H$ extension property

We have seen in the previous section that, differently from the classical case, for a general Carnot group $\mathbb{G}$ the suitable Whitney condition for (f, X) on K is not sufficient for the existence of an extension $(f, \dot{f})$ of (f, X) on $\mathbb{R}$. More precisely, it follows from Theorem 3.8 that if $\mathbb{G}$ has horizontal vectors which are not pliable, then there exist triples (K, f, X) such that the $\mathcal{C}^1_H$-Whitney condition holds for (f, X) on K but there is not a $\mathcal{C}^1_H$-extension of (f, X). In this next section we prove the converse to the result above, showing that the $\mathcal{C}^1_H$ extension property holds when all horizontal vectors are pliable, i.e., when $\mathbb{G}$ is pliable.

We start by introducing the notion of a *locally uniformly pliable* horizontal vector.

Definition 4.1. A horizontal vector X is called *locally uniformly pliable* if there exists a neighborhood $\mathcal{U}$ of X in $\mathfrak{G}_H$ such that for every $\varepsilon > 0$, there exists $\eta > 0$ so that for every $W \in \mathcal{U}$

$$\big\{(\gamma, \dot{\gamma})(1) \,\big|\, \gamma \in \mathcal{C}^1_H([0,1], \mathbb{G}),\ (\gamma, \dot{\gamma})(0) = (0_{\mathbb{G}}, W),\ \|\dot{\gamma} - W\|_{\infty, \mathfrak{G}_H} \le \varepsilon\big\} \supset \mathcal{B}_{\mathbb{G}}(\exp(W), \eta) \times \mathcal{B}_{\mathfrak{G}_H}(W, \eta).$$

Remark 4.2. As it happens for pliability, if X is locally uniformly pliable then, for every $\lambda \in \mathbb{R} \setminus \{0\}$, λX is locally uniformly pliable.

We are going to see in Remark 6.2 that pliability and local uniform pliability are not equivalent properties. The following proposition, however, establishes the equivalence between pliability and local uniform pliability of *all* horizontal vectors.

Proposition 4.3. *If $\mathbb{G}$ is pliable, then all horizontal vectors are locally uniformly pliable.*

Proof. Assume that $\mathbb{G}$ is pliable. For every $V \in \mathfrak{G}_H$ and $\varepsilon > 0$ denote by $\eta(V, \varepsilon)$ a positive constant such that

$$\left\{(\gamma, \dot\gamma)(1) \,\middle|\, \gamma \in \mathcal{C}^1_H([0,1], \mathbb{G}),\ (\gamma, \dot\gamma)(0) = (0_{\mathbb{G}}, V),\ \|\dot\gamma - V\|_{\infty, \mathfrak{G}_H} \le \varepsilon\right\}$$
$$\supset \mathcal{B}_{\mathbb{G}}(\exp(V), \eta(V, \varepsilon)) \times \mathcal{B}_{\mathfrak{G}_H}(V, \eta(V, \varepsilon)).$$

We are going to show that there exists $\nu(V, \varepsilon) > 0$ such that for every $W \in \mathcal{B}_{\mathfrak{G}_H}(V, \nu(V, \varepsilon))$

$$\left\{(\gamma, \dot\gamma)(1) \,\middle|\, \gamma \in \mathcal{C}^1_H([0,1], \mathbb{G}),\ (\gamma, \dot\gamma)(0) = (0_{\mathbb{G}}, W),\ \|\dot\gamma - W\|_{\infty, \mathfrak{G}_H} \le \varepsilon\right\}$$
$$\supset \mathcal{B}_{\mathbb{G}}\big(\exp(W), \tfrac14 \eta\big(V, \tfrac12 \varepsilon\big)\big) \times \mathcal{B}_{\mathfrak{G}_H}\big(W, \tfrac14 \eta\big(V, \tfrac12 \varepsilon\big)\big). \quad (8)$$

The proof of the local uniform pliability of any horizontal vector X is then concluded by simple compactness arguments (taking any compact neighborhood $\mathcal{U}$ of X, using the notation of Definition 4.1).

First fix $\bar\nu(V, \varepsilon) > 0$ in such a way that

$$\exp(W) \in \mathcal{B}_{\mathbb{G}}\big(\exp(V), \tfrac14 \eta\big(V, \tfrac12 \varepsilon\big)\big)$$

for every $W \in \mathcal{B}_{\mathfrak{G}_H}(V, \bar\nu(V, \varepsilon))$.

For every $W \in \mathfrak{G}_H$, every $\rho \in (0,1)$, and every curve $\gamma \in \mathcal{C}^1_H([0,1], \mathbb{G})$ such that $(\gamma, \dot\gamma)(0) = (0_{\mathbb{G}}, V)$, define $\gamma_{W,\rho} \in \mathcal{C}^1_H([0,1], \mathbb{G})$ as follows: $\gamma_{W,\rho}(0) = 0_{\mathbb{G}}$, $\dot\gamma_{W,\rho}(t) = (t/\rho)V + ((\rho - t)/\rho)W$ for $t \in [0, \rho]$, and $\dot\gamma_{W,\rho}(\rho + (1-\rho)t) = \dot\gamma(t)$ for $t \in [0,1]$. In particular,

$$\gamma_{W,\rho}(1) = \gamma_{W,\rho}(\rho) \cdot \delta_{1-\rho}(\gamma(1)), \qquad \dot\gamma_{W,\rho}(1) = \dot\gamma(1),$$

and

$$\|\dot\gamma_{W,\rho} - W\|_{\infty, \mathfrak{G}_H} \le \|\dot\gamma - V\|_{\infty, \mathfrak{G}_H} + \|W - V\|_{\mathfrak{G}_H}.$$

If $\|V - W\|_{\mathfrak{G}_H} \le \frac12 \varepsilon$, we then have

$$\|\dot\gamma_{W,\rho} - W\|_{\infty, \mathfrak{G}_H} \le \varepsilon \quad \forall \gamma \text{ such that } \|\dot\gamma - V\|_{\infty, \mathfrak{G}_H} \le \tfrac12 \varepsilon.$$

Since $\gamma_{W,\rho}(\rho)$ depends on V, W, and ρ, but not on γ, we conclude that, for every $W \in \mathcal{B}_{\mathfrak{G}_H}\big(V, \frac12 \varepsilon\big)$,

$$\left\{(\beta, \dot\beta)(1) \,\middle|\, \beta \in \mathcal{C}^1_H([0,1], \mathbb{G}),\ (\beta, \dot\beta)(0) = (0_{\mathbb{G}}, W),\ \|\dot\beta - W\|_{\infty, \mathfrak{G}_H} \le \varepsilon\right\}$$
$$\supset \big(\gamma_{W,\rho}(\rho) \cdot \delta_{1-\rho}\big(\mathcal{B}_{\mathbb{G}}\big(\exp(V), \eta\big(V, \tfrac12 \varepsilon\big)\big)\big)\big) \times \mathcal{B}_{\mathfrak{G}_H}\big(V, \eta\big(V, \tfrac12 \varepsilon\big)\big).$$

Notice that $d_{\mathbb{G}}(0_{\mathbb{G}}, \gamma_{W,\rho}(\rho)) \le \rho \max(\|V\|_{\mathfrak{G}_H}, \|W\|_{\mathfrak{G}_H})$. Thanks to Lemma 3.6, for ρ sufficiently small,

$$\gamma_{W,\rho}(\rho) \cdot \delta_{1-\rho}\big(\mathcal{B}_{\mathbb{G}}\big(\exp(V), \eta\big(V, \tfrac12 \varepsilon\big)\big)\big) \supset \mathcal{B}_{\mathbb{G}}\big(\exp(V), \tfrac12 \eta\big(V, \tfrac12 \varepsilon\big)\big).$$

Now,

$$\mathcal{B}_{\mathbb{G}}\big(\exp(V), \tfrac{1}{2}\eta\big(V, \tfrac{1}{2}\varepsilon\big)\big) \supset \mathcal{B}_{\mathbb{G}}\big(\exp(W), \tfrac{1}{4}\eta\big(V, \tfrac{1}{2}\varepsilon\big)\big)$$

whenever $W \in \mathcal{B}_{\mathfrak{G}_H}(V, \bar{\nu}(V, \varepsilon))$.

Similarly,

$$\mathcal{B}_{\mathfrak{G}_H}\big(V, \eta\big(V, \tfrac{1}{2}\varepsilon\big)\big) \supset \mathcal{B}_{\mathfrak{G}_H}\big(W, \tfrac{1}{4}\eta\big(V, \tfrac{1}{2}\varepsilon\big)\big),$$

provided that

$$\|V - W\|_{\mathfrak{G}_H} \leq \tfrac{3}{4}\eta\big(V, \tfrac{1}{2}\varepsilon\big).$$

The proof of (8) is concluded by taking

$$\nu(V, \varepsilon) = \min\big(\bar{\nu}(V, \varepsilon), \tfrac{1}{2}\varepsilon, \tfrac{3}{4}\eta\big(V, \tfrac{1}{2}\varepsilon\big)\big). \qquad \square$$

We are now ready to prove the converse of Theorem 3.8, concluding the proof of Theorem 1.1.

Theorem 4.4. *Let $\mathbb{G}$ be a pliable Carnot group. Then $(\mathbb{R}, \mathbb{G})$ has the $\mathcal{C}^1_H$ extension property.*

Proof. By Proposition 4.3, we can assume that all vectors in $\mathfrak{G}_H$ are locally uniformly pliable. Note, moreover, that it is enough to prove the extension for maps defined on compact sets K. The generalization to closed sets K_0 is immediate because the source Carnot group is $\mathbb{R}$. Let (f, X) satisfy the $\mathcal{C}^1_H$-Whitney condition on K, where K is compact. We have to define $\bar{f}$ on the complementary (open) set $\mathbb{R} \setminus K$, which is the countable and disjoint union of open intervals. For the unbounded components of $\mathbb{R} \setminus K$, we simply define $\bar{f}$ as the curve with constant speed $X(i)$ or $X(j)$, where $i = \min(K)$ and $j = \max(K)$. For the finite components (a, b) we proceed as follows. We consider $y = \delta_{1/(b-a)}(f(a)^{-1} \cdot f(b))$. We let ε be the smallest number such that

$$\big\{(\gamma, \dot{\gamma})(1) \bigm| \gamma \in \mathcal{C}^1_H([0, 1], \mathbb{G}),\ (\gamma, \dot{\gamma})(0) = (0_{\mathbb{G}}, X(a)),\ \|\dot{\gamma} - X(a)\|_{\infty, \mathfrak{G}_H} \leq \varepsilon'\big\}$$

contains $(y, X(b))$ for every $\varepsilon' > \varepsilon$. We consider an extension $\bar{f} \in \mathcal{C}^1_H$ of f on $[a, b]$ such that $\dot{\bar{f}}(a) = X(a)$, $\dot{\bar{f}}(b) = X(b)$, and $\|\dot{\bar{f}} - X(a)\|_{\infty, \mathfrak{G}_H} \leq 2\varepsilon$.

By construction, $\bar{f}$ extends f (but we do not know whether the extension is continuous) and $\dot{\bar{f}} = X$ on the interior of K. We prove now that $\bar{f}$ is $\mathcal{C}^1_H$ and that $\dot{\bar{f}} = X$ on the boundary ∂K of K. It is clear that $\bar{f}$ is $\mathcal{C}^1_H$ on $\mathbb{R} \setminus \partial K$. In order to conclude the proof, we pick $x \in \partial K$ and we are left to prove that $\dot{\bar{f}}(x)$ exists, is equal to $X(x)$, and $\dot{\bar{f}}|_{\mathbb{R} \setminus \partial K}$ has the correct limit at x. Because of the $\mathcal{C}^1_H$-Whitney condition, we know that for every sequence x_n in K converging to x, the vector $(\bar{f}(x_n) - \bar{f}(x))/(x_n - x)$ (written in any coordinate system) converges to the coordinate representation of $X(x)$. Let us then pick a sequence x_n in $\mathbb{R} \setminus K$ tending to x, and we are left to show that both $(\bar{f}(x_n) - \bar{f}(x))/(x_n - x)$ and $\dot{\bar{f}}(x_n)$ tend to $X(x)$. Assume for now that $x_n > x$ for every n. The connected component (a_n, b_n) of $\mathbb{R} \setminus K$ containing x_n is either constant for n large (in this case $x = a_n$) or its length goes to zero as $n \to \infty$. In the first case we simply notice that $\bar{f}|_{[a_n, b_n]}$ is $\mathcal{C}^1$ by construction and that the right derivative of $\bar{f}$ at $a_n = x$ is equal to $X(x)$. In the second case we can assume that $a_n < x_n < b_n$ and $b_n - a_n$ goes to zero. The local uniform pliability of $X(x)$ implies that $\|\dot{\bar{f}}|_{[a_n, b_n]} - X(a_n)\|_{\infty, \mathfrak{G}_H}$ goes to zero as $n \to \infty$. It follows that $\dot{\bar{f}}(x_n)$

converges to $X(x)$ as $n \to \infty$. Hence, still in some coordinate system,

$$\begin{aligned}\frac{\bar f(x_n)-\bar f(x)}{x_n-x} &= \frac{\bar f(a_n)+X(a_n)(x_n-a_n)-\bar f(x)}{x_n-x}+\frac{o(x_n-a_n)}{x_n-x}\\ &= \frac{\bar f(a_n)-\bar f(x)}{a_n-x}\frac{a_n-x}{x_n-x}+X(a_n)\left(1-\frac{a_n-x}{x_n-x}\right)+\frac{o(x_n-a_n)}{x_n-x}\to X(x)\end{aligned}$$

as $n \to \infty$. The situation where $x_n < x$ for infinitely many n can be handled similarly. □

5. Application to the Lusin approximation of an absolutely continuous curve

In a recent paper, E. Le Donne and G. Speight [2016, Theorem 1.2] prove the following result.

Proposition 5.1 (Le Donne–Speight). *Let $\mathbb{G}$ be a Carnot group of step 2 and consider a horizontal curve $\gamma : [a, b] \to \mathbb{G}$. For any $\varepsilon > 0$, there exist $K \subset [a, b]$ and a $\mathcal{C}^1_H$-curve $\gamma_1 : [a, b] \to \mathbb{G}$ such that $\mathcal{L}([a, b] \setminus K) < \varepsilon$ and $\gamma = \gamma_1$ on K.*

In the case in which $\mathbb{G}$ is equal to the n-th Heisenberg group $\mathbb{H}_n$, such a result had already been proved in [Speight 2016, Theorem 2.2]; see also [Zimmerman 2017, Corollary 3.8]. In Theorem 3.2 of the same paper, Speight also identified a horizontal curve on the Engel group such that the statement of Proposition 5.1 is not satisfied.

The name "Lusin approximation" for the property stated in Proposition 5.1 comes from the use of the classical theorem of Lusin [1912] in the proof. Let us sketch a proof when $\mathbb{G}$ is replaced by a vector space $\mathbb{R}^n$. The derivative $\dot\gamma$ of an absolutely continuous curve γ is an integrable function. Lusin's theorem states that $\dot\gamma$ coincides with a continuous vector-valued function $X : K \to \mathbb{R}^n$ on a set K of measure arbitrarily close to $b - a$. Thanks to the inner continuity of the Lebesgue measure, one can assume that K is compact. Moreover, K can be chosen so that the Whitney condition is satisfied by $(\gamma|_K, X)$ on K. This is a consequence of the mean value inequality

$$\|\gamma(x+h)-\gamma(x)-h\dot\gamma(x)\| \le o(h), \tag{9}$$

where $o(h)$ depends on $x \in K$. By usual arguments of measure theory, inequality (9) can be made uniform with respect to x if one slightly reduces the measure of K. The (classical) Whitney extension theorem provides a $\mathcal{C}^1$-curve γ_1 defined on $[a, b]$ with $\gamma_1 = \gamma$ and $\dot\gamma_1 = X$ on K.

The proof in [Le Donne and Speight 2016], and also in [Speight 2016], follows the same scheme as the one sketched above. We show here below how the same scheme can be adapted to any pliable Carnot group. The fact that all Carnot groups of step 2 are pliable and that not all pliable Carnot groups are of step 1 or 2 is proved in the next section (Theorem 6.5 and Proposition 6.6), so that our paper actually provides a nontrivial generalization of Proposition 5.1. The novelty of our approach with respect to those in [Le Donne and Speight 2016; Speight 2016; Zimmerman 2017] is to replace the classical Rademacher differentiability theorem for Lipschitz or absolutely continuous curves from $\mathbb{R}$ to $\mathbb{R}^n$ by the more adapted Pansu–Rademacher theorem.

Proposition 5.2 (Lusin approximation of a horizontal curve). *Let $\mathbb{G}$ be a pliable Carnot group and $\gamma : [a, b] \to \mathbb{G}$ be a horizontal curve. Then for any $\varepsilon > 0$ there exist $K \subset [a, b]$ with $\mathcal{L}([a, b] \setminus K) < \varepsilon$ and a curve $\gamma_1 : [a, b] \to \mathbb{G}$ of class $\mathcal{C}^1_H$ such that the curves γ and γ_1 coincide on K.*

Proof. We are going to prove that for any $\varepsilon > 0$ there exists a compact set $K \subset [a, b]$ with $\mathcal{L}([a, b] \setminus K) < \varepsilon$ such the three following conditions are satisfied:

(1) $\dot\gamma(t)$ exists and it is a horizontal vector at every $t \in K$.

(2) $\dot\gamma|_K$ is uniformly continuous.

(3) For every $\varepsilon > 0$ there exists $\eta > 0$ such that, for every $t \in K$ and $|h| \le \eta$ with $t + h \in [a, b]$, it holds that $d_{\mathbb{G}}\big(\gamma(t + h), \gamma(t) \cdot \exp(h\dot\gamma(t))\big) \le \eta\varepsilon$.

With these conditions the $\mathcal{C}^1_H$-Whitney condition holds for $(\gamma, \dot\gamma|_K)$ on K. Since $\mathbb{G}$ is pliable, according to Theorem 4.4 the $\mathcal{C}^1_H$ extension property holds for $(\mathbb{R}, \mathbb{G})$, yielding γ_1 as in the statement of Proposition 5.2.

Case 1: γ is Lipschitz continuous. Let γ be a Lipschitz curve from $[a, b]$ to $\mathbb{G}$. The Pansu–Rademacher theorem (Proposition 2.1) states that there exists $A \subset [a, b]$ of full measure such that, for any $t \in A$, the curve γ admits a derivative at t and it holds that

$$d_{\mathbb{G}}\big(\gamma(t + h), \gamma(t) \cdot \exp(h\dot\gamma(t))\big) = o(h)$$

as h goes to zero. Let ε be positive. By Lusin's theorem, one can restrict A to a compact set $K_1 \subset A$ such that $t \mapsto \dot\gamma(t)$ is uniformly continuous on K_1 and $\mathcal{L}(A \setminus K_1) < \frac{1}{2}\varepsilon$. Moreover, by classical arguments of measure theory, the functions $h \mapsto |h|^{-1} d_{\mathbb{G}}\big(\gamma(t + h), \gamma(t) \cdot \exp(h\dot\gamma(t))\big)$ can be bounded by a function that is $o(1)$ as h goes to zero, uniformly in t on some compact set K_2 with $\mathcal{L}(A \setminus K_2) < \frac{1}{2}\varepsilon$. In other words, for every $\varepsilon > 0$ there exists η such that for $t \in K_2$ and $h \in [t - \eta, t + \eta]$ it holds that

$$d_{\mathbb{G}}\big(\gamma(t + h), \gamma(t) \cdot \exp(h\dot\gamma(t))\big) \le \varepsilon|h|.$$

With $K = K_1 \cap K_2$, the three conditions (1), (2), (3) listed above hold true.

Case 2: γ is a general horizontal curve. Let γ be absolutely continuous on $[a, b]$. It admits a path-length parametrization i.e., there exists a Lipschitz continuous curve $\phi : [0, T] \to \mathbb{G}$ and a function $F : [a, b] \to [0, T]$, absolutely continuous and nondecreasing, such that $\gamma = \phi \circ F$. Moreover, $\dot\phi$ has norm 1 at almost every time. As F is absolutely continuous, for every $\varepsilon > 0$ there exists η such that, for any measurable K, the inequality $\mathcal{L}([0, T] \setminus K) < \eta$ implies $\mathcal{L}([a, b] \setminus F^{-1}(K)) < \varepsilon$.

Let ε be positive and let η be a number corresponding to $\frac{1}{2}\varepsilon$ in the previous sentence. Applying to F the scheme of proof sketched after Proposition 5.1 for $n = 1$, there exists a compact set $K_F \subset [a, b]$ with $\mathcal{L}([a, b] \setminus K_F) < \frac{1}{2}\varepsilon$ such that F is differentiable with a continuous derivative on K_F and the bound in the mean value inequality is uniform on K_F. For the Lipschitz curve ϕ and for every $\eta > 0$, Case 1 provides a compact set $K_\phi \subset [0, T]$ with the listed properties with $\frac{1}{2}\varepsilon$ in place of ε.

Let K be the compact $K_F \cap F^{-1}(K_\phi)$ and note that $\mathcal{L}([a, b] \setminus K) < \varepsilon$. For $t \in K$ it holds that

$$|F(t + h) - F(t) - hF'(t)| = o(h)$$

and

$$d_{\mathbb{G}}\big(\phi(F(t)+H), \phi(F(t)) \cdot \exp(H\dot{\phi}(F(t)))\big) = o(H)$$

as h and H go to zero, uniformly with respect to $t \in K$. We also know that $t \mapsto F'(t)$ and $t \mapsto \dot{\phi}(F(t)) \in \mathfrak{G}_H$ exist and are continuous on K. It is a simple exercise to compose the two Taylor expansions and obtain the wanted conditions for $\gamma = \phi \circ F$. Note that the derivative of γ on K is $F'(t)\dot{\phi}(F(t))$, which is continuous on K. □

Remark 5.3. A set $E \subset \mathbb{R}^n$ is said to be 1-countably rectifiable if there exists a countable family of Lipschitz curves $f_k : \mathbb{R} \to \mathbb{R}^n$ such that

$$\mathcal{H}^1\Big(E \setminus \bigcup_k f_k(\mathbb{R})\Big) = 0.$$

The usual Lusin approximation of curves in $\mathbb{R}^n$ permits one to replace Lipschitz by $\mathcal{C}^1$ in this classical definition of rectifiability. When $\mathbb{R}^n$ is replaced by a pliable Carnot group, the two definitions still make sense and, according to Proposition 5.2, are still equivalent. Rectifiability in metric spaces and Carnot groups is a very active research topic in geometric measure theory; see [Le Donne and Speight 2016] for references.

6. Conditions ensuring pliability

The goal of this section is to identify conditions ensuring that $\mathbb{G}$ is pliable. Let us first focus on the pliability of the zero vector.

Proposition 6.1. *For every Carnot group $\mathbb{G}$, the vector $0 \in \mathfrak{G}$ is pliable.*

Proof. According to Proposition 3.7, we should prove that for every $\varepsilon > 0$ the set

$$\big\{\beta(1) \in \mathbb{G} \,\big|\, \beta \in \mathcal{C}^1_H([0,1], \mathbb{G}),\ \|\dot{\beta}\|_{\infty,\mathfrak{G}_H} < \varepsilon,\ (\beta, \dot{\beta})(0) = (0_{\mathbb{G}}, 0)\big\}$$

is a neighborhood of $0_{\mathbb{G}}$ in $\mathbb{G}$.

Recall that there exist $k \in \mathbb{N}$, $V_1, \ldots, V_k \in \mathfrak{G}_H$ and $t_1, \ldots, t_k > 0$ such that the map

$$\phi : (\tau_1, \ldots, \tau_k) \mapsto e^{\tau_k V_k} \circ \cdots \circ e^{\tau_1 V_1}(0_{\mathbb{G}})$$

has rank equal to $\dim(\mathbb{G})$ at $(\tau_1, \ldots, \tau_k) = (t_1, \ldots, t_k)$ and satisfies $\phi(t_1, \ldots, t_k) = 0_{\mathbb{G}}$; see [Sussmann 1976]. Notice that for every $\nu > 0$, the function

$$\phi_\nu : (\tau_1, \ldots, \tau_k) \mapsto e^{\nu\tau_k V_k} \circ \cdots \circ e^{\nu\tau_1 V_1}(0_{\mathbb{G}}) = e^{(\nu^2\tau_k/\nu)V_k} \circ \cdots \circ e^{\nu^2(\tau_1/\nu)V_1}(0_{\mathbb{G}}) = \delta_{\nu^2}\Big(\phi\Big(\frac{\tau_1}{\nu}, \ldots, \frac{\tau_k}{\nu}\Big)\Big)$$

has also rank equal to $\dim(\mathbb{G})$ at $(\tau_1, \ldots, \tau_k) = (\nu t_1, \ldots, \nu t_k)$ and satisfies $\phi_\nu(\nu t_1, \ldots, \nu t_k) = 0_{\mathbb{G}}$. Hence, up to replacing t_j by νt_j and V_j by $\nu^2 V_j$ for $j = 1, \ldots, k$ and ν small enough, we can assume that $t_1 + \cdots + t_k < 1$ and $\|V_j\|_{\mathfrak{G}_H} < \varepsilon$ for $j = 1, \ldots, k$.

Let O be a neighborhood of $(t_1, \ldots, t_k)$ such that for every $(\tau_1, \ldots, \tau_k) \in O$ we have $\tau_1, \ldots, \tau_k > 0$ and $\tau_1 + \cdots + \tau_k < 1$. Notice that $\{\phi(\tau_1, \ldots, \tau_k) \mid (\tau_1, \ldots, \tau_k) \in O\}$ is a neighborhood of $0_{\mathbb{G}}$ in $\mathbb{G}$.

We complete the proof of the proposition by constructing, for every $\tau = (\tau_1, \dots, \tau_k) \in O$, a curve $\beta_\tau \in \mathcal{C}^1_H([0,1], \mathbb{G})$ such that

$$\|\dot\beta_\tau\|_{\infty,\mathfrak{G}_H} < \varepsilon, \quad (\beta_\tau, \dot\beta_\tau)(0) = (0_{\mathbb{G}}, 0), \quad \beta_\tau(1) = \phi(\tau). \tag{10}$$

For every $X \in \mathfrak{G}_H$, $p \in \mathbb{G}$ and $r > 0$ let us exhibit a curve $\gamma \in \mathcal{C}^1_H([0,r], \mathbb{G})$ such that $\gamma(0) = \gamma(r) = p$, $\dot\gamma(0) = 0$, $\dot\gamma(r) = X$, and $\|\dot\gamma\|_{\infty,\mathfrak{G}_H} = \|X\|_{\mathfrak{G}_H}$. The curve γ can be constructed by imposing $\dot\gamma\left(\frac{1}{2}r\right) = -\frac{1}{2}X$ and by extending $\dot\gamma$ on $\left[0, \frac{1}{2}r\right]$ and $\left[\frac{1}{2}r, r\right]$ by convex interpolation. It is also possible to reverse such a curve by transformation (T4) and connect on any segment $[0,r]$ the point-with-velocity (p, X) with the point-with-velocity $(p, 0)$ by a $\mathcal{C}^1_H$ curve γ respecting, moreover, $\|\dot\gamma\|_{\infty,\mathfrak{G}_H} = \|X\|_{\mathfrak{G}_H}$. Finally just concatenating (transformation (T5)) curves of this type it is possible, for every $r > 0$, to connect (p, X) and (p, Y) on $[0,r]$ with a curve $\gamma_{r,X,Y} \in \mathcal{C}^1_H([0,r], \mathbb{G})$ with $\|\dot\gamma_{r,X,Y}\|_{\infty,\mathfrak{G}_H} = \max(\|X\|_{\mathfrak{G}_H}, \|Y\|_{\mathfrak{G}_H})$.

We then construct β_τ as follows: we fix $r = \left(1 - \sum_{j=1}^k \tau_j\right)/k$, we impose $\beta_\tau(0) = 0_{\mathbb{G}}$ and we define $\dot\beta_\tau$ to be the concatenation of the following $2k$ continuous curves in $\mathfrak{G}_H$: first take $\dot\gamma_{r,0,V_1}$, then the constant equal to V_1 for a time τ_1, then $\dot\gamma_{r,V_1,V_2}$, then the constant equal to V_2 for a time τ_2, and so on up to $\dot\gamma_{r,V_{k-1},V_k}$ and finally the constant equal to V_k for a time τ_k. By construction, $\beta_\tau \in \mathcal{C}^1_H([0,1], \mathbb{G})$ and satisfies (10). □

Remark 6.2. Let us show that, as a consequence of the previous proposition, pliability and local uniform pliability are not equivalent properties (albeit we know from Proposition 4.3 that pliability of all horizontal vectors is equivalent to local uniform pliability of all horizontal vectors).

Recall that local uniform pliability of a horizontal vector X implies pliability of all horizontal vectors in a neighborhood of X (see Definition 4.1). Therefore, if 0 is locally uniformly pliable for a Carnot group $\mathbb{G}$ then every horizontal vector of $\mathfrak{G}$ is pliable (Remark 4.2). Hence 0 cannot be locally uniformly pliable if G is not pliable. The remark is concluded by recalling that nonpliable Carnot groups exist (see Examples 3.3 and 3.5).

Let $\mathbb{G}$ be a Carnot group and let $X_1, \dots, X_m$ be an orthonormal basis of $\mathfrak{G}_H$. Let us consider the control system in $\mathbb{G} \times \mathbb{R}^m$ given by

$$\begin{cases} \dot\gamma = \displaystyle\sum_{i=1}^m u_i X_i(\gamma), \\ \dot u = v, \end{cases} \tag{11}$$

where both $u = (u_1, \dots, u_m)$ and the control $v = (v_1, \dots, v_m)$ vary in $\mathbb{R}^m$.

Let us rewrite $x = (\gamma, u)$,

$$F_0(x) = \begin{pmatrix} \sum_{i=1}^m u_i X_i(\gamma) \\ 0 \end{pmatrix}, \qquad F_i(x) = \begin{pmatrix} 0 \\ e_i \end{pmatrix} \quad \text{for } i = 1, \dots, m,$$

where $e_1, \dots, e_m$ denotes the canonical basis of $\mathbb{R}^m$. System (11) can then be rewritten as

$$\dot x = F_0(x) + \sum_{i=1}^m v_i F_i(x). \tag{12}$$

For every $\bar{u}\in\mathbb{R}^m$, let $\mathcal{F}_{\bar{u}}:L^1([0,1],\mathbb{R}^m)\to\mathbb{G}\times\mathbb{R}^m$ be the endpoint map at time 1 for system (12) with initial condition $(0_{\mathbb{G}},\bar{u})$. Notice that if $x(\cdot)=(\gamma(\cdot),u(\cdot))$ is a solution of (12) with initial condition $(0_{\mathbb{G}},\bar{u})$ corresponding to a control $v\in L^1([0,1],\mathbb{R}^m)$, then $\gamma\in\mathcal{C}^1_H([0,1],\mathbb{G})$ and $\|\dot{\gamma}-\sum_{i=1}^m\bar{u}_iX_i\|_{\infty,\mathfrak{G}_H}\le\|v\|_1$.

We can then state the following criterion for pliability.

Proposition 6.3. *If the map $\mathcal{F}_{\bar{u}}:L^1([0,1],\mathbb{R}^m)\to\mathbb{G}\times\mathbb{R}^m$ is open at 0, then the horizontal vector $\sum_{i=1}^m\bar{u}_iX_i$ is pliable.*

As a consequence, if the restriction of $\mathcal{F}_{\bar{u}}$ to $L^\infty([0,1],\mathbb{R}^m)$ is open at 0, when the L^∞ topology is considered on $L^\infty([0,1],\mathbb{R}^m)$, then $\sum_{i=1}^m\bar{u}_iX_i$ is pliable. We deduce the following property: *if a straight curve is not pliable, then it admits an abnormal lift in $T^*\mathbb{G}$. Indeed, if a horizontal vector $\sum_{i=1}^m\bar{u}_iX_i$ is not pliable, then the differential of $\mathcal{F}_{\bar{u}}|_{L^\infty([0,1],\mathbb{R}^m)}$* at 0 must be singular. Hence — see, for instance, [Agrachev and Sachkov 2004, Section 20.3] or [Trélat 2005, Proposition 5.3.3] — there exist $p_\gamma:[0,1]\to T^*\mathbb{G}$ and $p_u:[0,1]\to(\mathbb{R}^m)^*$ with $(p_\gamma,p_u)\neq 0$ such that

$$\dot{p}_\gamma(t)=-\frac{\partial}{\partial\gamma}H\big(\gamma(t),\bar{u},p_\gamma(t),p_u(t),0\big),\tag{13}$$

$$\dot{p}_u(t)=-\frac{\partial}{\partial u}H\big(\gamma(t),\bar{u},p_\gamma(t),p_u(t),0\big),\tag{14}$$

$$0=\frac{\partial}{\partial v}H\big(\gamma(t),\bar{u},p_\gamma(t),p_u(t),0\big)\tag{15}$$

for $t\in[0,1]$, where $\gamma(t)=\exp\big(t\sum_{i=1}^m\bar{u}_iX_i\big)$ and

$$H(\gamma,u,p_\gamma,p_u,v)=p_\gamma\sum_{i=1}^m u_iX_i(\gamma)+p_uv.$$

From (15) it follows that $p_u(t)=0$ for all $t\in[0,1]$. Equation (14) then implies that $p_\gamma(t)X_i(\gamma(t))=0$ for every $i=1,\dots,m$ and every $t\in[0,1]$. Moreover, p_γ must be different from zero. Comparing (4) and (13), it follows that p_γ is an abnormal path.

The control literature proposes several criteria for testing the openness at 0 of an endpoint map of the type $\mathcal{F}_{\bar{u}}|_{L^\infty([0,1],\mathbb{R}^m)}$. The test presented here below, taken from [Bianchini and Stefani 1990], generalizes previous criteria obtained in [Hermes 1982; Sussmann 1987].

Theorem 6.4 [Bianchini and Stefani 1990, Corollary 1.2]. *Let M be a $\mathcal{C}^\infty$ manifold and $V_0,V_1,\dots,V_m$ be $\mathcal{C}^\infty$ vector fields on M. Assume that the family of vector fields $\mathcal{J}=\{\mathrm{ad}^k_{V_0}V_j\mid k\ge 0,\ j=1,\dots,m\}$ is Lie-bracket generating. Denote by $\mathcal{H}$ the iterated brackets of elements in $\mathcal{J}$ and recall that the length of an element of $\mathcal{H}$ is the sum of the number of times that each of the elements $V_0,\dots,V_m$ appears in its expression. Assume that every element of $\mathcal{H}$ in whose expression each of the vector fields $V_1,\dots,V_m$ appears an even number of times is equal, at every $q\in M$, to the linear combination of elements of $\mathcal{H}$ of smaller length, evaluated at q. Fix $q_0\in M$ and a neighborhood Ω of 0 in $\mathbb{R}^m$. Let $\mathcal{U}\subset L^\infty([0,1],\Omega)$ be the set of those controls v such that the solution of $\dot{q}=V_0(q)+\sum_{i=1}^m v_iV_i(q)$, $q(0)=q_0$, is defined up to time 1 and denote by $\Phi(v)$ the endpoint $q(1)$ of such a solution. Then $\Phi(\mathcal{U})$ is a neighborhood of $e^{V_0}(q_0)$.*

The following two results show how to apply Theorem 6.4 to guarantee that a Carnot group $\mathbb{G}$ is pliable and, hence, that $(\mathbb{R}, \mathbb{G})$ has the $\mathcal{C}^1_H$ extension property.

Theorem 6.5. *Let $\mathbb{G}$ be a Carnot group of step 2. Then $\mathbb{G}$ is pliable and $(\mathbb{R}, \mathbb{G})$ has the $\mathcal{C}^1_H$ extension property.*

Proof. We are going to apply Theorem 6.4 in order to prove that for every horizontal vector $\sum_{i=1}^m u_i X_i$ the endpoint map $\mathcal{F}_u : L^\infty([0,1], \mathbb{R}^m) \to \mathbb{G} \times \mathbb{R}^m$ is open at zero.

Notice that

$$[F_0, F_i](\gamma, w) = -\begin{pmatrix} X_i(\gamma) \\ 0 \end{pmatrix}, \qquad i = 1, \dots, m,$$

and

$$[F_0, [F_0, F_i]](\gamma, w) = \begin{pmatrix} \sum_{j=1}^m w_j [X_i, X_j](\gamma) \\ 0 \end{pmatrix}, \qquad i = 1, \dots, m.$$

Moreover, for every $i, j = 1, \dots, m$,

$$[[F_0, F_i], F_j] = 0, \;\; [[F_0, F_i], [F_0, F_j]](\gamma, w) = \begin{pmatrix} [X_i, X_j](\gamma) \\ 0 \end{pmatrix},$$

and all other Lie brackets in and between elements of $\mathcal{J} = \{\mathrm{ad}^k_{F_0} F_i \mid k \geq 0,\ i = 1, \dots, m\}$ are zero since $\mathbb{G}$ is of step 2.

In particular all Lie brackets between elements of $\mathcal{J}$ in which each of the vector fields $F_1, \dots, F_m$ appears an even number of times is zero.

According to Theorem 6.4, we are left to prove that $\mathcal{J}$ is Lie-bracket generating. This is clearly true, since

$$\mathrm{Span}\big\{F_i(\gamma, w),\ [F_0, F_i](\gamma, w),\ [[F_0, F_i], [F_0, F_j]](\gamma, w) \,\big|\, i, j = 1, \dots, m\big\}$$

is equal to $T_{(\gamma,w)}(\mathbb{G} \times \mathbb{R}^m)$ for every $(\gamma, w) \in \mathbb{G} \times \mathbb{R}^m$. □

We conclude the paper by showing how to construct pliable Carnot groups of arbitrarily large step.

Proposition 6.6. *For every $s \geq 1$ there exists a pliable Carnot group of step s.*

Proof. Fix $s \geq 1$ and consider the free nilpotent stratified Lie algebra $\mathcal{A}$ of step s generated by s elements $Z_1, \dots, Z_s$.

For every $i = 1, \dots, s$, denote by I_i the ideal of $\mathcal{A}$ generated by Z_i and by J_i the ideal $[I_i, I_i]$. Then $J = \bigoplus_{i=1}^s J_i$ is also an ideal of $\mathcal{A}$.

Then the factor algebra $\mathfrak{G} = \mathcal{A}/J$ is nilpotent and inherits the stratification of $\mathcal{A}$. Denote by $\mathbb{G}$ the Carnot group generated by $\mathfrak{G}$. Let $X_1, \dots, X_s$ be the elements of $\mathfrak{G}_H$ obtained by projecting $Z_1, \dots, Z_s$. By construction, every bracket of $X_1, \dots, X_s$ in $\mathfrak{G}$ in which at least one of the X_i appears more than once is zero. Moreover, $\mathfrak{G}$ has step s, since $[X_1, [X_2, [\cdots, X_s], \cdots]]$ is different from zero.

Let us now apply Theorem 6.4 to prove that for every $X \in \mathfrak{G}_H$ the endpoint map $\mathcal{F}_u : L^\infty([0,1], \mathbb{R}^s) \to \mathbb{G} \times \mathbb{R}^s$ is open at zero, where $u \in \mathbb{R}^s$ is such that $X = \sum_{i=1}^s u_i X_i$.

Following the same computations as in the proof of Theorem 6.5,

$$\mathrm{ad}_{F_0}^{k+1} F_i(\gamma, u) = \begin{pmatrix} \mathrm{ad}_X^k X_i(\gamma) \\ 0 \end{pmatrix}, \qquad k \geq 0,\ i = 1, \dots, s.$$

In particular the family $\mathcal{J} = \{\mathrm{ad}_{F_0}^k F_i \mid k \geq 0,\ i = 1, \dots, s\}$ is Lie-bracket generating.

Moreover, every Lie bracket of elements of $\widehat{\mathcal{J}} = \{\mathrm{ad}_{F_0}^{k+1} F_i \mid k \geq 0,\ i = 1, \dots, s\}$ in which at least one of the elements $F_1, \dots, F_s$ appears more than once is zero.

Consider now a Lie bracket W between $h \geq 2$ elements of $\mathcal{J}$. Let $k_1, \dots, k_s$ be the number of times in which each of the elements $F_1, \dots, F_s$ appears in W. Let us prove by induction on h that W is the linear combination of brackets between elements of $\widehat{\mathcal{J}}$ in which each F_i appears k_i times, $i = 1, \dots, s$. Consider the case $h = 2$. Any bracket of the type $[\mathrm{ad}_{F_0}^k F_i, F_j]$, $k \geq 0$, $i, j = 1, \dots, s$, is the linear combination of brackets between elements of $\widehat{\mathcal{J}}$ in which F_i and F_j appear once, as it can easily be proved by induction on k, thanks to the Jacobi identity. The induction step on h also follows directly from the Jacobi identity.

We can therefore conclude that every Lie bracket of elements of $\mathcal{J}$ in which at least one of the elements $F_1, \dots, F_s$ appears more than once is zero. This implies in particular that the hypotheses of Theorem 6.4 are satisfied, concluding the proof that $\mathbb{G}$ is pliable. □

Acknowledgments

We warmly thank Frédéric Chapoton and Gwénaël Massuyeau for suggestions leading us to Proposition 6.6. We are also grateful to Artem Kozhevnikov, Dario Prandi, Luca Rizzi and Andrei Agrachev for many stimulating discussions. This work was initiated during the IHP trimester "Geometry, analysis and dynamics on sub-Riemannian manifolds" and we wish to thank the Institut Henri Poincaré and the Fondation Sciences Mathématiques de Paris for the welcoming working conditions.

Sigalotti was supported by the European Research Council, ERC StG 2009 "GeCoMethods", contract number 239748, by the grant ANR-15-CE40-0018 of the ANR and by the FMJH Program Gaspard Monge in optimization and operation research.

References

[Agrachev and Sachkov 2004] A. A. Agrachev and Y. L. Sachkov, *Control theory from the geometric viewpoint*, Encyclopaedia of Mathematical Sciences **87**, Springer, 2004. MR Zbl

[Agrachev and Sarychev 1996] A. A. Agrachev and A. V. Sarychev, "Abnormal sub-Riemannian geodesics: morse index and rigidity", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **13**:6 (1996), 635–690. MR Zbl

[Balogh and Fässler 2009] Z. M. Balogh and K. S. Fässler, "Rectifiability and Lipschitz extensions into the Heisenberg group", *Math. Z.* **263**:3 (2009), 673–683. MR Zbl

[Balogh et al. 2016] Z. M. Balogh, U. Lang, and P. Pansu, "Lipschitz extensions of maps between Heisenberg groups", *Ann. Inst. Fourier (Grenoble)* **66**:4 (2016), 1653–1665. MR Zbl

[Barilari et al. 2016a] D. Barilari, U. Boscain, and M. Sigalotti (editors), *Geometry, analysis and dynamics on sub-Riemannian manifolds, I* (Luminy, 2014), Eur. Math. Soc., Zürich, 2016. MR Zbl

[Barilari et al. 2016b] D. Barilari, U. Boscain, and M. Sigalotti (editors), *Geometry, analysis and dynamics on sub-Riemannian manifolds, II* (Luminy, 2014), European Mathematical Society, Zürich, 2016. MR Zbl

[Bianchini and Stefani 1990] R. M. Bianchini and G. Stefani, "Graded approximations and controllability along a trajectory", *SIAM J. Control Optim.* **28**:4 (1990), 903–924. MR Zbl

[Bonfiglioli et al. 2007] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians*, Springer, 2007. MR Zbl

[Brudnyi and Shvartsman 1994] Y. Brudnyi and P. Shvartsman, "Generalizations of Whitney's extension theorem", *Internat. Math. Res. Notices* **1994**:3 (1994), 129–139. MR Zbl

[Bryant and Hsu 1993] R. L. Bryant and L. Hsu, "Rigidity of integral curves of rank 2 distributions", *Invent. Math.* **114**:2 (1993), 435–461. MR Zbl

[Evans and Gariepy 2015] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 2015. MR Zbl

[Fefferman 2005] C. L. Fefferman, "A sharp form of Whitney's extension theorem", *Ann. of Math.* (2) **161**:1 (2005), 509–577. MR Zbl

[Fefferman et al. 2014] C. L. Fefferman, A. Israel, and G. K. Luli, "Sobolev extension by linear operators", *J. Amer. Math. Soc.* **27**:1 (2014), 69–145. MR Zbl

[Folland and Stein 1982] G. B. Folland and E. M. Stein, *Hardy spaces on homogeneous groups*, Mathematical Notes **28**, Princeton University Press, 1982. MR Zbl

[Franchi et al. 2001] B. Franchi, R. Serapioni, and F. Serra Cassano, "Rectifiability and perimeter in the Heisenberg group", *Math. Ann.* **321**:3 (2001), 479–531. MR Zbl

[Franchi et al. 2003] B. Franchi, R. Serapioni, and F. Serra Cassano, "On the structure of finite perimeter sets in step 2 Carnot groups", *J. Geom. Anal.* **13**:3 (2003), 421–466. MR Zbl

[Golé and Karidi 1995] C. Golé and R. Karidi, "A note on Carnot geodesics in nilpotent Lie groups", *J. Dynam. Control Systems* **1**:4 (1995), 535–549. MR Zbl

[Hermes 1982] H. Hermes, "Control systems which generate decomposable Lie algebras", *J. Differential Equations* **44**:2 (1982), 166–187. MR Zbl

[Huang and Yang 2012] T. Huang and X. Yang, "Extremals in some classes of Carnot groups", *Sci. China Math.* **55**:3 (2012), 633–646. MR Zbl

[Kirchheim and Serra Cassano 2004] B. Kirchheim and F. Serra Cassano, "Rectifiability and parameterization of intrinsic regular surfaces in the Heisenberg group", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **3**:4 (2004), 871–896. MR Zbl

[Kozhevnikov 2015] A. Kozhevnikov, *Propriétés métriques des ensembles de niveau des applications différentiables sur les groupes de Carnot*, Ph.D. thesis, Université Paris Sud, 2015, available at https://tel.archives-ouvertes.fr/tel-01178864.

[Le Donne and Speight 2016] E. Le Donne and G. Speight, "Lusin approximation for horizontal curves in step 2 Carnot groups", *Calc. Var. Partial Differential Equations* **55**:5 (2016), art. id. 111. MR Zbl

[Le Donne et al. 2014] E. Le Donne, A. Ottazzi, and B. Warhurst, "Ultrarigid tangents of sub-Riemannian nilpotent groups", *Ann. Inst. Fourier* (*Grenoble*) **64**:6 (2014), 2265–2282. MR Zbl

[Liu and Sussman 1995] W. Liu and H. J. Sussman, *Shortest paths for sub-Riemannian metrics on rank-two distributions*, Mem. Amer. Math. Soc. **564**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Lusin 1912] N. Lusin, "Sur les propriétés des fonctions mesurables", *C. R. Acad. Sci. Paris* **154**:25 (1912), 1688–1690. Zbl JFM

[Magnani 2013] V. Magnani, "Towards differential calculus in stratified groups", *J. Aust. Math. Soc.* **95**:1 (2013), 76–128. MR Zbl

[Malgrange 1967] B. Malgrange, *Ideals of differentiable functions*, Tata Institute of Fundamental Research Studies in Mathematics **3**, Tata Institute of Fundamental Research, Bombay, 1967. MR Zbl

[Montgomery 1995] R. Montgomery, "A survey of singular curves in sub-Riemannian geometry", *J. Dynam. Control Systems* **1**:1 (1995), 49–90. MR Zbl

[Pansu 1989] P. Pansu, "Métriques de Carnot–Carathéodory et quasiisométries des espaces symétriques de rang un", *Ann. of Math.* (2) **129**:1 (1989), 1–60. MR Zbl

[Rigot and Wenger 2010] S. Rigot and S. Wenger, "Lipschitz non-extension theorems into jet space Carnot groups", *Int. Math. Res. Not.* **2010**:18 (2010), 3633–3648. MR Zbl

[Serra Cassano 2016] F. Serra Cassano, "Some topics of geometric measure theory in Carnot groups", pp. 1–121 in *Geometry, analysis and dynamics on sub-Riemannian manifolds, I*, edited by D. Barilari et al., Eur. Math. Soc., Zürich, 2016. MR Zbl

[Speight 2016] G. Speight, "Lusin approximation and horizontal curves in Carnot groups", *Rev. Mat. Iberoam.* **32**:4 (2016), 1423–1444. MR Zbl

[Sussmann 1976] H. J. Sussmann, "Some properties of vector field systems that are not altered by small perturbations", *J. Differential Equations* **20**:2 (1976), 292–315. MR Zbl

[Sussmann 1987] H. J. Sussmann, "A general theorem on local controllability", *SIAM J. Control Optim.* **25**:1 (1987), 158–194. MR Zbl

[Trélat 2005] E. Trélat, *Contrôle optimal: théorie et applications*, Vuibert, Paris, 2005. MR Zbl

[Vodop'yanov and Pupyshev 2006a] S. K. Vodop'yanov and I. M. Pupyshev, "Whitney-type theorems on the extension of functions on Carnot groups", *Sibirsk. Mat. Zh.* **47**:4 (2006), 731–752. In Russian; translated in *Sib. Math. J.* **47**:4 (2006), 601–620. MR Zbl

[Vodop'yanov and Pupyshev 2006b] S. K. Vodop'yanov and I. M. Pupyshev, "Whitney-type theorems on the extension of functions on the Carnot group", *Dokl. Akad. Nauk* **406**:5 (2006), 586–590. In Russian. MR

[Wenger and Young 2010] S. Wenger and R. Young, "Lipschitz extensions into jet space Carnot groups", *Math. Res. Lett.* **17**:6 (2010), 1137–1149. MR Zbl

[Whitney 1934a] H. Whitney, "Analytic extensions of differentiable functions defined in closed sets", *Trans. Amer. Math. Soc.* **36**:1 (1934), 63–89. MR Zbl

[Whitney 1934b] H. Whitney, "Differentiable functions defined in closed sets, I", *Trans. Amer. Math. Soc.* **36**:2 (1934), 369–387. MR Zbl

[Zimmerman 2017] S. Zimmerman, "The Whitney extension theorem for C^1, horizontal curves in the Heisenberg Group", *J. Geom. Anal.* (online publication February 2017).

Received 19 Dec 2016. Accepted 11 Jun 2017.

NICOLAS JUILLET: nicolas.juillet@math.unistra.fr
Institut de Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 67000 Strasbourg, France

MARIO SIGALOTTI: mario.sigalotti@inria.fr
Inria, Team GECO & CMAP, École Polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau, France

ANALYSIS AND PDE
Vol. 10, No. 7, 2017
dx.doi.org/10.2140/apde.2017.10.1663

TREND TO EQUILIBRIUM FOR THE BECKER–DÖRING EQUATIONS: AN ANALOGUE OF CERCIGNANI'S CONJECTURE

JOSÉ A. CAÑIZO, AMIT EINAV AND BERTRAND LODS

We investigate the rate of convergence to equilibrium for subcritical solutions to the Becker–Döring equations with physically relevant coagulation and fragmentation coefficients and mild assumptions on the given initial data. Using a discrete version of the log-Sobolev inequality with weights, we show that in the case where the coagulation coefficient grows linearly and the detailed balance coefficients are of typical form, one can obtain a linear functional inequality for the dissipation of the relative free energy. This results in showing Cercignani's conjecture for the Becker–Döring equations and consequently in an exponential rate of convergence to equilibrium. We also show that for all other typical cases, one can obtain an "almost" Cercignani's conjecture, which results in an algebraic rate of convergence to equilibrium.

1. Introduction

The Becker–Döring equations. The Becker–Döring equations are a fundamental set of equations which describe the kinetics of a first-order phase transition. Amongst the phenomena to which they are relevant one can find crystallisation [Kelton et al. 1983], nucleation of polymers [Capasso 2003], vapour condensation, aggregation of lipids [Neu et al. 2002] and phase separation in alloys [Xiao and Haasen 1991]. For more general reviews of nucleation theory, see, for instance, [Schmelzer 2005; Oxtoby 1992].

Cañizo was supported by the Marie-Curie CIG grant KineticCF and the Spanish Ministerio de Economía y Competitividad/European Regional Development Fund (ERDF/FEDER), project MTM2014-52056-P. Einav was supported by EPSRC grant EP/L002302/1. Lods acknowledges support of the "de Castro" Statistics Initiative, Collegio C. Alberto, Moncalieri, Italy.
MSC2010: 34D05, 35Q82, 82C05, 82C40.
Keywords: Becker–Döring, nucleation, exponential convergence, entropy method.

The Becker–Döring equations give the time evolution of the size distribution of clusters of a certain substance. Denoting by $\{c_i(t)\}_{i\in\mathbb{N}}$ the density of clusters of size i at time $t \geqslant 0$ (i.e., the density of clusters that are composed of i particles), the equations read

$$\frac{\mathrm{d}}{\mathrm{d}t}c_i(t) = W_{i-1}(t) - W_i(t), \quad i \in \mathbb{N}\setminus\{1\}, \tag{1-1a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}c_1(t) = -W_1(t) - \sum_{k=1}^{\infty} W_k(t), \tag{1-1b}$$

where

$$W_i(t) := a_i\, c_1(t)c_i(t) - b_{i+1}\, c_{i+1}(t), \quad i \in \mathbb{N}, \tag{1-2}$$

and a_i, b_i, assumed to be strictly positive, are the *coagulation and fragmentation coefficients*. They determine, respectively, the rate at which clusters of size i combine with clusters of size 1 to create clusters of size $i+1$, and the rate at which clusters of size $i+1$ split into clusters of sizes i and 1. This corresponds to the basic assumption of the underlying model: if we represent symbolically by $\{i\}$ the chemical species of clusters of size i, then the only (relevant) chemical reactions that take place are

$$\{i\} + \{1\} \rightleftharpoons \{i+1\}.$$

The quantity $W_i(t)$ defined in (1-2) represents the *net rate* of the reaction $\{i\}+\{1\} \rightleftharpoons \{i+1\}$, and under the above set of equations it is easy to see that the *density*, or *mass*, of the solution, defined by

$$\varrho := \sum_{i=1}^{\infty} i\, c_i(0) = \sum_{i=1}^{\infty} i\, c_i(t), \tag{1-3}$$

is formally conserved under time evolution. The original equations proposed by Becker and Döring [1935] were similar to (1-1), with the slight change that the density of one particle c_1, usually called the monomer density, was assumed to be constant. The current version, motivated by the conservation of total density, was first discussed in [Burton 1977] and [Penrose and Lebowitz 1979] and is widely used in classical nucleation theory.

Much like in other kinetic equations, the study of a state of equilibrium and the convergence to it is a fundamental question in the study of the Becker–Döring equations. Defining the *detailed balance coefficients* Q_i recursively by

$$Q_1 = 1, \qquad Q_{i+1} = \frac{a_i}{b_{i+1}} Q_i, \quad i \in \mathbb{N}, \tag{1-4}$$

one can see that for a given $z \geqslant 0$ the sequence

$$c_i = Q_i z^i \tag{1-5}$$

is formally an equilibrium of (1-1). However, depending on the coagulation and fragmentation coefficients a_i and b_i, many of these formal equilibria do not have a finite mass. The largest $z_s \geqslant 0$, possibly $z_s = +\infty$, for which

$$\sum_{i=1}^{\infty} i Q_i z^i < +\infty \quad \text{for all } 0 \leqslant z < z_s$$

is called the *critical monomer density*, or sometimes the monomer saturation density. The *critical mass* (or, again, saturation mass) is then defined by

$$\varrho_{\mathrm{s}} := \sum_{i=1}^{\infty} i Q_i z_{\mathrm{s}}^i \in [0, +\infty]. \tag{1-6}$$

It is important to note that both z_{s} and ϱ_{s} are uniquely determined by a_i and b_i and that $\{Q_i z^i\}_{i\in\mathbb{N}}$ is a finite-mass equilibrium only for $0 \leqslant z < z_{\mathrm{s}}$, with the possibility for the equality $z = z_{\mathrm{s}}$ only when $\varrho_{\mathrm{s}} < +\infty$. Additionally, it is easy to see that for a given finite mass $\varrho \leqslant \varrho_{\mathrm{s}}$ there exists a unique $\bar{z} \geqslant 0$ such that

$$\varrho = \sum_{i=1}^{\infty} i Q_i \bar{z}^i,$$

giving us a candidate for the asymptotic equilibrium state of (1-1) under a given initial data. These are in fact the only finite-mass equilibria (see [Ball et al. 1986]), and $\bar{z}$ defined above is called the *equilibrium monomer density* for a given mass ρ.

A finite mass solution is called *subcritical* when its mass ϱ is strictly less than ϱ_{s}. It is called *critical* if $\varrho = \varrho_{\mathrm{s}}$ and *supercritical* if $\varrho > \varrho_{\mathrm{s}}$, assuming $\varrho_{\mathrm{s}} < +\infty$. In this paper we will only concern ourselves with subcritical solutions. Thus, to avoid triviality we always assume that $z_{\mathrm{s}} > 0$.

The critical density ϱ_{s}, if finite, marks a change in the behaviour of equilibrium states: if $\varrho < \varrho_{\mathrm{s}}$ then a unique equilibrium state with mass ϱ exists, while if $\varrho > \varrho_{\mathrm{s}}$, no such equilibrium can occur and a phase transition phenomenon takes place — reflected in the fact that the excess density $\varrho - \varrho_{\mathrm{s}}$ is concentrated in larger and larger clusters as time progresses.

Previous results. Let us briefly review existing results on the mathematical theory of the Becker–Döring equations, which has advanced much since the first rigorous works on the topic [Ball and Carr 1988; Ball et al. 1986]. In [Ball et al. 1986] the authors showed (among other things) existence and uniqueness of a global solution to (1-1) when

$$a_i \leqslant C_1 i, \quad b_i \leqslant C_2 i, \quad \sum_{i=1}^{\infty} i^{1+\varepsilon} c_i(0) < +\infty \tag{1-7}$$

for some constants $C_1, C_2, \varepsilon > 0$. As expected, under the above assumptions the unique solution conserves mass (that is, (1-3) holds rigorously). This basic existence theory is applicable to all solutions we consider in this work.

The asymptotic behaviour of solutions to (1-1) is one of the most interesting aspects of the equation. Supercritical behaviour, while not dealt with in this work, has a particularly interesting link to late-stage coarsening and has been studied extensively in [Penrose 1997; Velázquez 1998; Collet et al. 2002; Niethammer 2003], with several questions still open. Asymptotic approximations of such solutions have been developed in [Farjoun and Neu 2008; 2011; Neu et al. 2005].

Regarding the subcritical regime, it was proved in [Ball and Carr 1988; Ball et al. 1986] that solutions with subcritical mass ϱ approach the unique equilibrium with this mass, determined by (1-3). A fundamental quantity in understanding this approach is the *free energy*, $H(\boldsymbol{c})$, defined for any nonnegative sequence $\boldsymbol{c} = \{c_i\}_{i\in\mathbb{N}}$ by

$$H(\boldsymbol{c}) := \sum_{i=1}^{\infty} c_i \left(\log \frac{c_i}{Q_i} - 1\right) \tag{1-8}$$

whenever the sum converges. It can be shown that $H(\boldsymbol{c}(t))$ decreases along solutions $\boldsymbol{c} = \boldsymbol{c}(t)$ to the Becker–Döring equations; in fact, for a (strictly positive, suitably decaying for large i) solution $\boldsymbol{c}(t) = \{c_i(t)\}_{i\in\mathbb{N}}$ of (1-1) we have

$$\frac{\mathrm{d}}{\mathrm{d}t} H(\boldsymbol{c}(t)) = -D(\boldsymbol{c}(t)) := -\sum_{i=1}^{\infty} a_i Q_i \left(\frac{c_1 c_i}{Q_i} - \frac{c_{i+1}}{Q_{i+1}}\right)\left(\log \frac{c_1 c_i}{Q_i} - \log \frac{c_{i+1}}{Q_{i+1}}\right) \leqslant 0. \tag{1-9}$$

This free energy is motivated by physical considerations and constitutes a Lyapunov functional for our equation. Since it does not have a definite sign, we define a more natural candidate to measure the distance of $\boldsymbol{c}(t) = \{c_i(t)\}_{i\in\mathbb{N}}$ to the equilibrium. Using the notation

$$(\mathcal{Q}_z)_i = Q_i z^i$$

and denoting $\mathcal{Q}_{\bar{z}}$ by $\mathcal{Q}$, we can define the *relative* free energy as

$$H(\boldsymbol{c}|\mathcal{Q}) := \sum_{i=1}^{\infty} c_i \left(\log \frac{c_i}{\bar{z}^i Q_i} - 1\right) + \sum_{i=1}^{\infty} \bar{z}^i Q_i = H(\boldsymbol{c}) - \log \bar{z} \sum_{i=1}^{\infty} i c_i + \sum_{i=1}^{\infty} \bar{z}^i Q_i. \tag{1-10}$$

The relative free energy has the same time derivative as the free energy, and thus the same monotonicity property

$$\frac{\mathrm{d}}{\mathrm{d}t} H(\boldsymbol{c}(t)|\mathcal{Q}) = -D(\boldsymbol{c}(t)) \quad \forall t \geqslant 0,$$

where the *free energy dissipation* D is defined in (1-9). The relative free energy also satisfies

- $H(\boldsymbol{c}|\mathcal{Q}) \geqslant 0$, as can be seen by writing

$$H(\boldsymbol{c}|\mathcal{Q}) = \sum_{i=1}^{\infty} \mathcal{Q}_i \varphi\left(\frac{c_i}{\mathcal{Q}_i}\right), \quad \text{with } \varphi(r) := r \log r - r + 1 \geqslant 0, \tag{1-11}$$

- $H(\boldsymbol{c}|\mathcal{Q}) = 0$ if and only if $c_i = \mathcal{Q}_i = Q_i \bar{z}^i$ for any $i \in \mathbb{N}$, which is readily seen from (1-11).

This hints that $H(\boldsymbol{c}|\mathcal{Q})$ is the right "distance" to investigate. Indeed, while $H(\boldsymbol{c}|\mathcal{Q})$ is not a distance strictly speaking, it does control the ℓ^1 distance between $\boldsymbol{c}$ and $\mathcal{Q}$ by the celebrated Csiszár–Kullback inequality,[1] which in our case translates to

$$\|\boldsymbol{c} - \mathcal{Q}\|_{\ell^1(\mathbb{N})} = \sum_{i=1}^{\infty} |c_i - \mathcal{Q}_i| \leqslant \sqrt{2\varrho H(\boldsymbol{c}|\mathcal{Q})}. \tag{1-12}$$

(See also [Jabin and Niethammer 2003, Corollary 2.2] for a version involving the ℓ^1 distance with weight i.) The issue of estimating the rate of convergence to equilibrium of subcritical solutions is the main concern

[1] Sometimes called the Pinsker or Kullback–Pinsker inequality.

of this paper. The first result in this direction was [Jabin and Niethammer 2003], where they investigated the possibility of applying the so-called *entropy method* to the Becker–Döring equation. This consists roughly in looking for functional inequalities between a suitable Lyapunov functional of the equation (generally called the entropy; it corresponds to the relative free energy in our case) and its dissipation, so that one obtains a differential inequality that estimates the rate of convergence to equilibrium. In the case of the Becker–Döring equation, it was proved in [Jabin and Niethammer 2003] that there exists a constant $C > 0$, depending only on the fixed parameters of the problem and the initial data, such that

$$D(\boldsymbol{c}) \geqslant C \frac{H(\boldsymbol{c}|\mathcal{Q})}{(\log H(\boldsymbol{c}|\mathcal{Q}))^2} \tag{1-13}$$

for all nonnegative sequences $\boldsymbol{c}$ with subcritical mass ϱ, satisfying $\varepsilon \leqslant c_1 \leqslant z_s - \varepsilon$ for some $\varepsilon > 0$ and

$$\sum_{i=1}^{\infty} e^{\mu i} c_i =: M^{\exp} < +\infty. \tag{1-14}$$

The constant C depends on ε and $M^{\exp}$. This result applies under reasonable conditions on the coefficients a_i and b_i; in particular, it applies to the coefficients (1-23) and (1-25), which we give as examples below. If we consider now a solution $\boldsymbol{c} = \boldsymbol{c}(t)$ to (1-1), we may apply the inequality (1-13) to $\boldsymbol{c}(t)$ as long as $\boldsymbol{c}(t)$ satisfies the appropriate conditions, obtaining

$$\frac{\mathrm{d}}{\mathrm{d}t} H(\boldsymbol{c}(t)|\mathcal{Q}) = -D(\boldsymbol{c}(t)) \leqslant -C \frac{H(\boldsymbol{c}(t)|\mathcal{Q})}{(\log H(\boldsymbol{c}(t)|\mathcal{Q}))^2}.$$

Adding to this some additional considerations for the times t for which the inequality (1-13) is not applicable to $\boldsymbol{c}(t)$, one can deduce that $H(\boldsymbol{c}(t)|\mathcal{Q})$ is (roughly) bounded above by the solution of the above differential inequality, namely that

$$H(\boldsymbol{c}(t)|\mathcal{Q}) \leqslant H(\boldsymbol{c}(0)|\mathcal{Q}) e^{-Kt^{1/3}}$$

for some $K > 0$. Using inequality (1-12), this gives an almost-exponential rate of convergence to equilibrium for subcritical solutions in the $\ell^1(\mathbb{N})$ norm.

The question remained open of whether the convergence is in fact exponential or not. Recently this has been answered positively in [Cañizo and Lods 2013] by two of the authors of the present paper, through a different approach involving a detailed study of the spectrum of the linearisation of equation (1-1) around a subcritical equilibrium. This is an approach with a strong analogy to results in the theory of the Boltzmann equation; we refer to [Cañizo and Lods 2013; Villani 2003; Desvillettes et al. 2011] for more details on this parallel. The idea of the argument is to use the inequality (1-13) when one is far from equilibrium. Then, once we have reached a region which is close enough to equilibrium, the linearised regime is dominant and one can use the spectral study of the linearised operator in order to show that the convergence is in fact exponential. The outcome of this strategy is the following: for many interesting coefficients (including (1-23) and (1-25)), subcritical solutions $\boldsymbol{c} = \boldsymbol{c}(t)$ to (1-1) with

$$\sum_{i=1}^{\infty} e^{\mu i} c_i(0) =: M^{\exp} < +\infty \quad \text{for some } \mu > 0$$

satisfy

$$\sum_{i=1}^{\infty} e^{\mu' i} |c_i(t) - \mathcal{Q}_i| \leqslant C e^{-\lambda t} \quad \text{for } t \geqslant 0$$

for some $0 < \mu' < \mu$, $C > 0$ and $\lambda > 0$ which depend on the parameters of the problem and on M^{exp}. In fact, μ' and C only depend on the initial data $\boldsymbol{c}(0)$ through its mass and the value of M^{exp}; λ depends only on the coefficients and the initial mass and can be estimated explicitly. The value of λ is bounded above by (and can be taken very close to) the size of the spectral gap of the linearised operator. Recently Murray and Pego [2015] have used this spectral gap and developed the local estimates of the linearised operator in order to obtain convergence to equilibrium at a polynomial rate with milder conditions on the decay of the initial data. These results, like those in [Cañizo and Lods 2013], are local in nature and require the use of some global estimate such as (1-13) in order to provide global rates of convergence to equilibrium.

Main results. Our main goal in this work is to complete the picture of convergence to equilibrium by investigating modified and improved versions of the inequality (1-13). We show optimal inequalities and settle the question of whether full exponential convergence can be obtained through a *linear* inequality of the form

$$D(\boldsymbol{c}) \geqslant K H(\boldsymbol{c}|\mathcal{Q})$$

for some constant $K > 0$. In analogy to the Boltzmann equation, we refer to the question of whether such K exists along solutions to (1-1) as *Cercignani's conjecture for the Becker–Döring equations.* In fact, we show that under relatively mild conditions on the initial data, typical coagulation and fragmentation coefficients (covering the physically relevant situations; see the next subsection) admit an "almost" Cercignani conjecture for the energy dissipation, i.e., an inequality bounding $D(\boldsymbol{c})$ below by a power of $H(\boldsymbol{c}|\mathcal{Q})$, yielding an explicit rate of convergence to equilibrium. Surprisingly, we also find a relevant case ($a_i \sim i$ for all i) for which the conjecture is actually valid.

We will often require the following assumptions on the coagulation and fragmentation coefficients. Some of these are similar to those in [Jabin and Niethammer 2003], and always include physically relevant coefficients, such as those described in the next subsection. We recall that we always assume $a_i, b_i > 0$ for all $i \in \mathbb{N}$, and that the detailed balance coefficients Q_i were defined in (1-4) — given a_i one can determine b_i through Q_i, and vice versa.

Hypothesis 1. $$0 < z_{\text{s}} < +\infty.$$

Hypothesis 2. *For all $i \in \mathbb{N}$, we have $Q_i = z_{\text{s}}^{1-i} \alpha_i$, where $\{\alpha_i\}_{i \in \mathbb{N}}$ is a nonincreasing positive sequence with $\alpha_1 = 1$ and $\lim_{i\to\infty} \alpha_{i+1}/\alpha_i = 1$.*

Hypothesis 3. *There exist $C_1, C_2 > 0$ such that*

$$C_1 i^{\gamma} \leqslant a_i \leqslant C_2 i^{\gamma} \quad \textit{for all } i \in \mathbb{N}.$$

Hypothesis 2 on the form of Q_i is given as a compromise that allows us to give simple quantitative estimates of the constants in our theorems while allowing for the most commonly used types of coefficients. As one can see from the proofs, this assumption may be relaxed at the price of obtaining more involved estimates for our constants, particularly the logarithmic Sobolev constant in Proposition 3.4.

In most of our estimates a crucial role will be played by the *lower free energy dissipation*, $\overline{D}(\boldsymbol{c})$, defined for a given nonnegative sequence $\boldsymbol{c}$ by

$$\overline{D}(\boldsymbol{c}) = \sum_{i=1}^{\infty} a_i Q_i \left(\sqrt{\frac{c_1 c_i}{Q_i}} - \sqrt{\frac{c_{i+1}}{Q_{i+1}}} \right)^2. \tag{1-15}$$

At this point one notices that the elementary inequality $(x-y)(\log x - \log y) \geqslant 4(\sqrt{x} - \sqrt{y})^2$ when $x, y > 0$ implies

$$D(\boldsymbol{c}) \geqslant 4\overline{D}(\boldsymbol{c})$$

for any nonnegative sequence $\boldsymbol{c}$. Thus, any lower bound that is obtained for $\overline{D}(\boldsymbol{c})$ will transfer immediately to $D(\boldsymbol{c})$.

We now state our main result on general functional inequalities for the free energy dissipation, from which later we conclude a quantitative rate of convergence to equilibrium. It can be divided into two parts: functional inequalities when c_1 is not too small and not too far from $\bar{z}$, and inequalities in the case where c_1 escapes the above region.

Theorem 1.1. *Let $\{a_i\}_{i\in\mathbb{N}}$ and $\{Q_i\}_{i\in\mathbb{N}}$ satisfy Hypotheses 1–3 and let $\mathbf{c} = \{c_i\}_{i\in\mathbb{N}}$ be an arbitrary positive sequence with finite total density $0 < \varrho < \varrho_{\mathrm{s}}$.*

(i) ***Estimate for $a_i \sim i$.*** *Assume that $\gamma = 1$ and that there exist $\delta > 0$ such that*

$$\delta < c_1 < z_{\mathrm{s}} - \delta. \tag{1-16}$$

Then there exists $K > 0$ depending only on δ, ϱ and the coefficients $\{a_i\}_{i\geqslant 1}$, $\{b_i\}_{i\geqslant 2}$, such that

$$\overline{D}(\boldsymbol{c}) \geqslant K H(\boldsymbol{c}|\mathcal{Q}). \tag{1-17}$$

(ii) ***Estimate for $a_i \sim i^\gamma$ with $\gamma < 1$.*** *Assume that $0 \leqslant \gamma < 1$ and that c_1 satisfies (1-16) for some $\delta > 0$. If, in addition, there exists $\beta > 1$ with*

$$M_\beta(\boldsymbol{c}) = \sum_{i=1}^{\infty} i^\beta c_i < +\infty \tag{1-18}$$

then there exists $K > 0$ depending only on δ, ϱ, $M_\beta(\mathbf{c})$ and the coefficients $\{a_i\}_{i\geqslant 1}$, $\{b_i\}_{i\geqslant 2}$ such that

$$\overline{D}(\boldsymbol{c}) \geqslant K H(c|\mathcal{Q})^{\frac{\beta-\gamma}{\beta-1}}. \tag{1-19}$$

(iii) ***Estimate for c_1 far from equilibrium.*** *Assume that $\gamma = 1$, or that $0 \leqslant \gamma < 1$ and (1-18) holds for some $\beta > 1$. Assume also that for some $\delta > 0$*

$$c_1 \leqslant \delta$$

or that

$$c_1 \geqslant z_{\mathrm{s}} - \delta;$$

i.e., c_1 is outside of the range given in (1-16). Then if $\delta > 0$ is small enough (depending only on ϱ and $\{Q_i\}_{i\geqslant 1}$), there exists $\varepsilon > 0$ depending only on δ, ϱ and the coefficients $\{a_i\}_{i\geqslant 1}$, $\{b_i\}_{i\geqslant 2}$ if

$\gamma = 1$ (and additionally on $M_\beta(\boldsymbol{c})$ if $\gamma < 1$) such that

$$\overline{D}(\boldsymbol{c}) \geqslant \varepsilon. \tag{1-20}$$

The constants K and ε can be estimated explicitly in all cases.

We emphasise that all constants in the above theorem depend only on ϱ, the coefficients $\{a_i\}_{i\in\mathbb{N}}$, $\{b_i\}_{i\in\mathbb{N}}$, and the additional bounds δ or M_β (notice that ϱ_s is determined by the coefficients alone). Case (ii) of Theorem 1.1 is optimal in the following sense:

Theorem 1.2. *Call X_ϱ the set of nonnegative sequences $\boldsymbol{c} = \{c_i\}_{i\in\mathbb{N}}$ with mass ϱ, i.e., such that $\sum_{i=1}^{\infty} i c_i = \varrho$. Then, there exist $\{a_i\}_{i\in\mathbb{N}}$ and $\{Q_i\}_{i\in\mathbb{N}}$ that satisfy Hypotheses 1–3 with $\gamma < 1$ such that*

$$\inf_{X_\varrho} \frac{D(\boldsymbol{c})}{H(\boldsymbol{c}|\mathcal{Q})} = 0$$

for any $\varrho < \varrho_\mathrm{s}$.

In other words, this shows that a linear inequality as that of Theorem 1.1(i) cannot hold if $a_i \sim i^\gamma$ with $\gamma < 1$.

The idea behind the proof of Theorem 1.1 is to use a discrete logarithmic Sobolev inequality with weights, motivated by works of Bobkov and Götze [1999] and Barthe and Roberto [2003], to show part (i). As the conditions for the validity of the log-Sobolev inequality are *not* satisfied under the conditions of part (ii), a simple interpolation is used to show the desired result in that case. Part (iii) is proved by two estimates: The case where c_1 is too large follows an idea essentially already stated in [Jabin and Niethammer 2003], while the case where c_1 is too small seems to be a new result which we provide.

From Theorem 1.1 one can conclude in a straightforward way the following theorem, our main result on the rate of convergence to equilibrium:

Theorem 1.3. *Let $\{a_i\}_{i\in\mathbb{N}}$ and $\{Q_i\}_{i\in\mathbb{N}}$ satisfy Hypotheses 1–3 with $0 \leqslant \gamma \leqslant 1$, and let $\boldsymbol{c} = \boldsymbol{c}(t) = \{c_i(t)\}_{i\in\mathbb{N}}$ be a solution to the Becker–Döring equations with mass $\varrho \in (0, \varrho_s)$.*

(i) ***Rate for $a_i \sim i$.*** *If $\gamma = 1$ then there exists a constant $K > 0$ depending only on δ, ϱ and the coefficients $\{a_i\}_{i\geqslant 1}$, $\{b_i\}_{i\geqslant 2}$, and a constant $C > 0$ depending only on $H(\boldsymbol{c}(0)|\mathcal{Q})$, ϱ and the coefficients $\{a_i\}_{i\geqslant 1}$, $\{b_i\}_{i\geqslant 2}$ such that*

$$H(\boldsymbol{c}(t)|\mathcal{Q}) \leqslant C e^{-Kt} \quad \text{for } t \geqslant 0.$$

(ii) ***Rate for $a_i \sim i^\gamma$, $\gamma < 1$.*** *If $\gamma < 1$ and $M_\beta(\boldsymbol{c}(0)) < +\infty$ for some $\beta \geqslant \max\{2-\gamma, 1+\gamma\}$ then there exists a constant $K > 0$ depending only on M_β, δ, ϱ and the coefficients $\{a_i\}_{i\geqslant 1}$, $\{b_i\}_{i\geqslant 2}$, and a constant $C > 0$ depending only on $H(\boldsymbol{c}(0)|\mathcal{Q})$, M_β, δ, ϱ and the coefficients $\{a_i\}_{i\geqslant 1}$, $\{b_i\}_{i\geqslant 2}$ such that*

$$H(\boldsymbol{c}(t)|\mathcal{Q}) \leqslant \frac{1}{\left(C + \frac{1-\gamma}{\beta-1} K t\right)^{\frac{\beta-1}{1-\gamma}}} \quad \text{for } t \geqslant 0.$$

The constants K and C can be estimated explicitly.

In order to deduce Theorem 1.3 we use the inequalities in Theorem 1.1 when they are applicable. Of course, the assumption that $c_1(t)$ is in the "good" region given by (1-16) becomes eventually true, since $c_1(t)$ is known to converge to $\bar{z}$. More explicitly, one can apply the Csiszár–Kullback inequality (1-12) to obtain that for any $t > t_0$ we have

$$\bar{z} - H(\boldsymbol{c}(t_0)|\mathcal{Q}) \leqslant c_1(t) \leqslant \bar{z} + H(\boldsymbol{c}(t_0)|\mathcal{Q}), \quad t \geqslant t_0.$$

If $H(\boldsymbol{c}(t_0)|\mathcal{Q})$ is small enough, this implies (1-16). For times t such that $c_1(t)$ is outside this "good" region, we use the inequality in Theorem 1.1(iii); details are given in Section 4.

There are several improvements in these theorems with respect to the existing theory. One of them is that they apply to more general initial conditions, removing the need for a finite exponential moment present in [Cañizo and Lods 2013; Jabin and Niethammer 2003]. Another one is that they answer the question of whether one can obtain a linear inequality such as (1-17) (i.e., whether the equivalent of Cercignani's conjecture holds), making clear the link to discrete logarithmic Sobolev inequalities. It does hold in the case $a_i \sim i$, which is physically relevant, for example, in modelling polymer chains [Farjoun and Neu 2011; Neu et al. 2002]. As a result, the statement for $a_i \sim i$ is quite strong: it gives full exponential convergence, with explicit constants in terms of the parameters, with no restriction on the initial data except that of subcritical mass. Point (ii) in Theorem 1.3 also relaxes the requirements on the initial data, at the price of obtaining a slower convergence than that of [Cañizo and Lods 2013]; we do not know whether this rate is optimal for initial conditions with polynomially decaying tails (so that $M_\beta < \infty$ for some $\beta > 1$, but $M_{\beta'} = +\infty$ for some $\beta' > \beta$). Recently, Murray and Pego [2015] investigated this rate of convergence, concluding an algebraic rate of decay as well. It would be interesting to verify the optimality of this result by determining whether the corresponding linearised operator admits a spectral gap in ℓ^1 spaces with polynomial weights (in ℓ^1 spaces with exponential weights, the answer is positive and an estimate of the spectral gap can be found in [Cañizo and Lods 2013]). We believe that no such spectral gap exists for $0 \leqslant \gamma < 1$, i.e., that the algebraic rate of convergence is optimal even for close to equilibrium initial data.

One may wonder if the method presented here can be used to reach an inequality like Jabin and Niethammer's (1-13) under the additional condition of an exponential moment. The answer is indeed positive:

Theorem 1.4. *Let $\{a_i\}_{i\in\mathbb{N}}$ and $\{Q_i\}_{i\in\mathbb{N}}$ satisfy Hypotheses 1–3 with $0 \leqslant \gamma < 1$.*

(i) ***Functional inequality.*** *Let $\boldsymbol{c} = \{c_i\}_{i\in\mathbb{N}}$ be an arbitrary positive sequence with mass $\varrho \in (0, \varrho_s)$ for which there exists $\mu > 0$ such that*

$$M_\mu^{\exp}(\boldsymbol{c}) := \sum_{i=1}^{\infty} e^{\mu i} c_i < +\infty. \tag{1-21}$$

Then there exist $K_1, K_2, \varepsilon > 0$ depending only on $M_\mu^{\exp}(\boldsymbol{c})$, δ, ϱ and the coefficients $\{a_i\}_{i\geqslant 1}$, $\{b_i\}_{i\geqslant 2}$ such that

$$\bar{D}(\boldsymbol{c}) \geqslant \min\left(\frac{K_1 H(\boldsymbol{c}|\mathcal{Q})}{\left|\log(K_2 H(\boldsymbol{c}|\mathcal{Q}))\right|^{1-\gamma}}, \varepsilon \right). \tag{1-22}$$

Moreover, K_1, K_2 and ε can be given explicitly.

(ii) ***Rate of convergence.*** *If $\boldsymbol{c}(t) = \{c_i(t)\}_{i\in\mathbb{N}}$ is a solution to the Becker–Döring equations with mass $0 < \varrho < \varrho_s$ such that there exists $\mu > 0$ with*

$$M_\mu^{\exp}(\boldsymbol{c}(0)) := \sum_{i=1}^{\infty} e^{\mu i} c_i(0) < +\infty,$$

then there exists a constant $K > 0$ depending only on $M_\mu^{\exp}(\boldsymbol{c}(0))$, δ, ϱ and the coefficients $\{a_i\}_{i\geqslant 1}$, $\{b_i\}_{i\geqslant 2}$, and a constant $C > 0$ depending only on $H(\boldsymbol{c}(0)|\mathcal{Q})$, $M_\mu^{\exp}(\boldsymbol{c}(0))$, δ, ϱ and the coefficients $\{a_i\}_{i\geqslant 1}$, $\{b_i\}_{i\geqslant 2}$ such that

$$H(\boldsymbol{c}(t)|\mathcal{Q}) \leqslant C e^{-Kt^{1/(2-\gamma)}}.$$

Moreover, K and C can be given explicitly.

Typical coefficients. The above results are valid for coagulation and fragmentation coefficients satisfying Hypotheses 1 – 3. To motivate our choice of assumptions, we briefly recall here some physically motivated coagulation and fragmentation coefficients found in the literature.

Common model coefficients appearing in the theory of density-conserving phase transitions (see [Niethammer 2003; Penrose 1989]) are given by

$$a_i = i^\gamma, \quad b_i = a_i\left(z_s + \frac{q}{i^{1-\mu}}\right) \quad \text{for all } i \geqslant 1 \tag{1-23}$$

for some $0 < \gamma \leqslant 1$, $z_s > 0$, $q > 0$ and $0 < \mu < 1$. These coefficients may be derived from simple assumptions on the mechanism of the reactions taking place; we take particular values from [Niethammer 2003]:

$$\begin{aligned}
\gamma &= \tfrac{1}{3}, &\quad \mu &= \tfrac{2}{3} &\quad &\text{(diffusion-limited kinetics in 3-D)},\\
\gamma &= 0, & \mu &= \tfrac{1}{2} & &\text{(diffusion-limited kinetics in 2-D)},\\
\gamma &= \tfrac{2}{3}, & \mu &= \tfrac{2}{3} & &\text{(interface-reaction-limited kinetics in 3-D)},\\
\gamma &= \tfrac{1}{2}, & \mu &= \tfrac{1}{2} & &\text{(interface-reaction-limited kinetics in 2-D)}.
\end{aligned} \tag{1-24}$$

The case $\gamma = 1$ appears, for example, in modelling polymer chains, where the binding energy increases by a constant each time a monomer is added.

A different kind of reasoning, based on a statistical mechanics argument involving the binding energy of clusters, results in the coefficients

$$a_i = i^\gamma, \quad b_i = z_s(i-1)^\gamma \exp(\sigma i^\mu - \sigma(i-1)^\mu), \quad i \in \mathbb{N}, \tag{1-25}$$

for appropriate constants γ, μ and where $\sigma > 0$ is related to the surface tension of the aggregates. The values of μ and γ for various situations are still those in (1-24).

As already mentioned, the choice $\gamma = 1$ corresponds to the physically relevant example in modelling polymer chains, for instance, for proteins aggregating in a cubic phase of lipid bilayers [Farjoun and Neu 2011; Neu et al. 2002].

The behaviour of (1-23) and (1-25) is similar: observe that for large i we have $i^{\mu} - (i-1)^{\mu} \sim \mu i^{\mu-1}$, so the fragmentation coefficients become roughly

$$b_i \sim z_{\mathrm{s}}\, a_i \exp(\sigma\mu i^{\mu-1}) \sim a_i \left(z_{\mathrm{s}} + \frac{z_{\mathrm{s}}\sigma\mu}{i^{1-\mu}} \right),$$

which is like (1-23) with $q = z_{\mathrm{s}}\sigma\mu$. Moreover, for both classes of coefficients, we can write (by the definition of Q_i)

$$Q_i = \frac{a_1 a_2 \cdots a_{i-1}}{b_2 b_3 \cdots b_i} = z_{\mathrm{s}}^{1-i} \alpha_i, \tag{1-26}$$

where $\{\alpha_i\}_{i\in\mathbb{N}}$ is nonincreasing and satisfies

$$\lim_{i\to\infty} \frac{\alpha_{i+1}}{\alpha_i} = 1.$$

In other words, Hypotheses 1–3 hold true for both models.

Application to general coagulation and fragmentation models. The Becker–Döring equations are the simplest form of a coagulation and fragmentation process, assuming that the only relevant reactions are governed by monomers. Other models take into account the fact that clusters of size i and size j, for $i, j \in \mathbb{N}$, may interact. A discrete model — similar to the Becker–Döring equations (1-1) — can be formulated, now with coagulation and fragmentation coefficients of the form $a_{i,j}$, $b_{i,j}$ (see the subsection on page 1698). Together with an assumption of detailed balance, one can once again find equilibria to the process and inquire about the rate of convergence to them. Our study of the Becker–Döring equations allows us to give a quantitative answer (though not optimal) for this question. We leave the detailed description of the model we have in mind for the subsection on page 1698. For such a model, using the same notion of free relative energy we will show that:

Theorem 1.5 (asymptotic behaviour of the coagulation-fragmentation system). *Let $\{a_{i,j}\}_{i,j\in\mathbb{N}}$ and $\{b_{i,j}\}_{i,j\in\mathbb{N}}$ be the coagulation and fragmentation coefficients for equation (5-1), and assume that the detailed balance condition (5-6) holds. Assume that*

$$a_{i,j} = i^{\gamma} + j^{\gamma} \tag{1-27}$$

for some $0 \leqslant \gamma < 1$ and that $\{Q_i\}_{i\in\mathbb{N}}$ satisfies Hypothesis 2. Assume in addition that $M_k(\boldsymbol{c}(0)) < +\infty$ for some $k \in \mathbb{N}$, $k > 1$. Then

$$H(\boldsymbol{c}(t)|\mathcal{Q}) \leqslant \frac{1}{(C_1 + C_2 \log t)^{\frac{k-1}{1-\gamma}}}, \quad t \geqslant 0, \tag{1-28}$$

where $C_1, C_2 > 0$ are constants depending only on $H(\boldsymbol{c}(0)|\mathcal{Q})$, z_{s}, ϱ, $\{\alpha_i\}_{i\in\mathbb{N}}$, k, γ and $M_k(\boldsymbol{c}(0))$.

Organisation of the paper. The structure of the paper is as follows: In Section 2 we will present our main technical tool, a discrete version of the log-Sobolev inequality with weights. Section 3 contains the proof of Theorem 1.1 and uses Section 2 to show the first part of the theorem. We also show in this section that this method is optimal and that Cercignani's conjecture cannot hold when $\gamma < 1$, proving Theorem 1.2, and explore the additional inequality that appears under the assumption of a finite exponential moment. Section 4 deals with the consequences of our functional inequalities for the solutions to the Becker–Döring equation and contains the proof of Theorem 1.3 and part (ii) of Theorem 1.4. In Section 5 we provide the proof of Theorem 1.5 and remark on the difficulties of obtaining stronger results in this general setting. Lastly, we give two appendices where proofs to some technical lemmas can be found.

2. A discrete weighted logarithmic Sobolev inequality

One of the key ingredients in proving Cercignani's conjecture for the Becker–Döring equations in terms of Theorem 1.1 is a discrete log-Sobolev inequality with weights. The theory presented here follows closely the work of Bobkov and Götze [1999], and that of Barthe and Roberto [2003], and can be seen as a discrete version of the aforementioned papers. It is worth noting that a discrete version is explicitly mentioned in [Barthe and Roberto 2003, Section 4], with a remark that the arguments in that paper can be adapted to prove it. Indeed, our proof is essentially an adaptation of the one in [Bobkov and Götze 1999], and we give it in this section for the sake of completeness (and since we have not been able to find an explicit proof in the discrete case). Some further technical details are postponed to Appendix A.

The main log-Sobolev inequality. We start with some basic definitions:

Definition 2.1. We say that $\boldsymbol{\mu} \in P(\mathbb{N})$ if $\boldsymbol{\mu} = \{\mu_i\}_{i\in\mathbb{N}}$ is a nonnegative sequence such that

$$\sum_{i=1}^{\infty} \mu_i = 1.$$

For any nonnegative sequence $\boldsymbol{g} = \{g_i\}_{i\in\mathbb{N}}$ with $\sum_{i=1}^{\infty} \mu_i g_i < +\infty$, we define its *entropy* with respect to $\boldsymbol{\mu}$ as

$$\mathrm{Ent}_{\mu}(\boldsymbol{g}) = \sum_{i=1}^{\infty} \mu_i g_i \log \frac{g_i}{\sum_{i=1}^{\infty} \mu_i g_i}. \tag{2-1}$$

Definition 2.2. Given $\boldsymbol{\mu} \in P(\mathbb{N})$ and positive sequence $\boldsymbol{\nu} = \{\nu_i\}_{i\in\mathbb{N}}$ (not necessarily normalised) we say that $\boldsymbol{\nu}$ admits a log-Sobolev inequality with respect to $\boldsymbol{\mu}$ with constant $0 < C_{\mathrm{LS}} < +\infty$ if, for any sequence $\boldsymbol{f} = \{f_i\}_{i\in\mathbb{N}}$,

$$\mathrm{Ent}_{\mu}(\boldsymbol{f}^2) \leqslant C_{\mathrm{LS}} \sum_{i=1}^{\infty} \nu_i (f_{i+1} - f_i)^2, \tag{2-2}$$

where $\boldsymbol{f}^2 = \{f_i^2\}_{i\in\mathbb{N}}$.

In what follows we will always assume that $\boldsymbol{\mu} \in P(\mathbb{N})$. Setting

$$\Psi(x) = |x| \log(1 + |x|),$$

the main theorem, and its simplified corollary, that we will prove in this section are:

Theorem 2.3. *The following two conditions are equivalent*:

(i) $\boldsymbol{\nu}$ *admits a log-Sobolev inequality with respect to* $\boldsymbol{\mu}$ *with constant* C_{LS}.

(ii) *For any $m \in \mathbb{N}$ such that*

$$\max\Big(\sum_{i=1}^{m-1} \mu_i, \sum_{i=m+1}^{\infty} \mu_i\Big) < \frac{2}{3},$$

we have

$$B_1 = \sup_{k \geqslant m} \frac{\sum_{i=1}^{k} 1/\nu_i}{\Psi^{-1}\big(1/\sum_{i=k+1}^{\infty} \mu_i\big)} < +\infty. \tag{2-3}$$

Moreover, *if* (ii) *is valid then one can choose*

$$C_{\mathrm{LS}} = 40(B_2 + 4B_1), \quad \text{where } B_2 = \frac{\sum_{i=1}^{m-1} 1/\nu_i}{\Psi^{-1}\big(1/\sum_{i=1}^{m-1} \mu_i\big)}. \tag{2-4}$$

A somehow more tractable consequence is the following.

Corollary 2.4. *The following two conditions are equivalent*:

(i) $\boldsymbol{\nu}$ *admits a log-Sobolev inequality with respect to* $\boldsymbol{\mu}$ *with constant* C_{LS}.

(ii) *For any $m \in \mathbb{N}$ such that*

$$\max\Big(\sum_{i=1}^{m-1} \mu_i, \sum_{i=m+1}^{\infty} \mu_i\Big) < \frac{2}{3},$$

we have

$$D_1 = \sup_{k \geqslant m} \Big(-\sum_{i=k+1}^{\infty} \mu_i \log\Big(\sum_{i=k+1}^{\infty} \mu_i\Big)\Big)\Big(\sum_{i=1}^{k} \frac{1}{\nu_i}\Big) < \infty. \tag{2-5}$$

Moreover, *if* (ii) *is valid then one can choose*

$$C_{\mathrm{LS}} = 120(D_2 + 4D_1), \tag{2-6}$$

where $D_2 = \big(-\sum_{i=1}^{m-1} \mu_i \log\big(\sum_{i=1}^{m-1} \mu_i\big)\big)\big(\sum_{i=1}^{m-1} 1/\nu_i\big)$.

Remark 2.5. One can clearly see that if

$$\sup_{k \geqslant 1} \Big(-\sum_{i=k+1}^{\infty} \mu_i \log\Big(\sum_{i=k+1}^{\infty} \mu_i\Big)\Big)\Big(\sum_{i=1}^{k} \frac{1}{\nu_i}\Big) < \infty$$

then one has a log-Sobolev inequality of $\boldsymbol{\nu}$ with respect to $\boldsymbol{\mu}$. However, the introduction of the "approximate median" m allows us to have an explicit estimation on the log-Sobolev constant C_{LS}.

The rest of the Section is dedicated to the proof of the above results and will be divided in various steps — each one corresponding to a subsection.

A reformulation as a Poincaré inequality in Orlicz spaces. As in the work of Bobkov and Götze [1999], a key argument in the proofs of Theorem 2.3 and Corollary 2.4 is to recast the log-Sobolev inequality as a Poincaré inequality in the Orlicz space associated to Ψ. We start with the definition:

Definition 2.6. Given $\boldsymbol{\mu} \in P(\mathbb{N})$ and a Young function, $\Sigma : [0, +\infty) \to [0, +\infty)$, i.e., a convex function such that

$$\frac{\Sigma(x)}{x} \xrightarrow{x \to +\infty} +\infty, \qquad \frac{\Sigma(x)}{x} \xrightarrow{x \to 0} 0,$$

we define the Orlicz space $L_\Sigma^{(\mu)}$ as the space of all sequences $\boldsymbol{f}$ such that there exists $k > 0$ with

$$\sum_{i=1}^{\infty} \mu_i \, \Sigma\left(\frac{|f_i|}{k}\right) < \infty.$$

In that case we define

$$\|\boldsymbol{f}\|_{L_\Sigma^{(\mu)}} = \inf_{k>0} \left\{ \sum_{i=1}^{\infty} \mu_i \, \Sigma\left(\frac{|f_i|}{k}\right) \leqslant 1 \right\}.$$

In what follows we will drop the superscript μ from the Orlicz space of Ψ and its norm. Additionally we set $\Phi(x) = \Psi(x^2)$ and notice that

$$\|\boldsymbol{f}^2\|_{L_\Psi} = \inf_{k>0} \left\{ \sum_{i=1}^{\infty} \mu_i \, \Psi\left(\frac{f^2}{k}\right) \leqslant 1 \right\} = \left(\inf_{\sqrt{k}>0} \left\{ \sum_{i=1}^{\infty} \mu_i \, \Phi\left(\frac{|f|}{\sqrt{k}}\right) \leqslant 1 \right\} \right)^2 = \|\boldsymbol{f}\|_{L_\Phi}^2. \tag{2-7}$$

We have then the following version of Rothaus's lemma:

Lemma 2.7. *Given $\boldsymbol{\mu} \in P(\mathbb{N})$ and a sequence $\boldsymbol{f} = \{f_i\}_{i \in \mathbb{N}}$, we set*

$$\mathcal{L}(\boldsymbol{f}) = \sup_{\alpha \in \mathbb{R}} \operatorname{Ent}_\mu((\boldsymbol{f} + \alpha)^2), \tag{2-8}$$

where $\boldsymbol{f} + \alpha = \{f_i + \alpha\}_{i \in \mathbb{N}}$. Then,

$$\operatorname{Ent}_\mu(\boldsymbol{f}^2) \leqslant \mathcal{L}(\boldsymbol{f}) \leqslant \operatorname{Ent}_\mu(\boldsymbol{f}^2) + 2 \sum_{i=1}^{\infty} \mu_i f_i^2. \tag{2-9}$$

Remark 2.8. This lemma is an adaptation of the appropriate lemma in [Rothaus 1985, Lemma 9]. We leave the proof of it to Appendix A.

We have then the following equivalent formulation of the log-Sobolev inequality:

Proposition 2.9. *The following conditions are equivalent*:

(i) *$\boldsymbol{\nu}$ admits a log-Sobolev inequality with respect to $\boldsymbol{\mu}$ with constant C_{LS}.*

(ii) *For any sequence $\boldsymbol{f}$,*

$$\mathcal{L}(\boldsymbol{f}) \leqslant C_{\mathrm{LS}} \sum_{i=1}^{\infty} \nu_i (f_{i+1} - f_i)^2. \tag{2-10}$$

(iii) *For any sequence* $\boldsymbol{f}$,

$$\|\boldsymbol{f}-\langle\boldsymbol{f}\rangle\|_{L_\Phi}^2 \leqslant \lambda\sum_{i=1}^{\infty}\nu_i(f_{i+1}-f_i)^2, \tag{2-11}$$

where $\langle\boldsymbol{f}\rangle=\sum_{i=1}^{\infty}\mu_i f_i$.

Moreover, if (i) *or* (ii) *are valid, one can choose* $\lambda=\frac{3}{2}C_{\mathrm{LS}}$. *If* (iii) *is valid one can choose* $C_{\mathrm{LS}}=5\lambda$.

The proof of the proposition relies on the following lemma:

Lemma 2.10. *For any sequence* $\boldsymbol{f}$, *one has*

$$\tfrac{2}{3}\|\boldsymbol{f}-\langle\boldsymbol{f}\rangle\|_{L_\Phi}^2 \leqslant \mathcal{L}(\boldsymbol{f}) \leqslant 5\|\boldsymbol{f}-\langle\boldsymbol{f}\rangle\|_{L_\Phi}^2. \tag{2-12}$$

Proof. We start by noticing that we may assume $\langle\boldsymbol{f}\rangle=0$, as well as $\|\boldsymbol{f}-\langle\boldsymbol{f}\rangle\|_{L_\Phi}=1$. This is true as $\mathcal{L}$ is invariant under translations and

$$\mathrm{Ent}_\mu(\alpha\boldsymbol{f})=\alpha\,\mathrm{Ent}_\mu(\boldsymbol{f}).$$

Using Lemma 2.7, we find that

$$\begin{aligned}\mathcal{L}(\boldsymbol{f}) &\leqslant \mathrm{Ent}_\mu(\boldsymbol{f}^2)+2\sum_{i=1}^{\infty}\mu_i f_i^2\\ &=\sum_{i=1}^{\infty}\mu_i f_i^2\log(f_i^2)+2\sum_{i=1}^{\infty}\mu_i f_i^2-\Bigl(\sum_{i=1}^{\infty}\mu_i f_i^2\Bigr)\log\Bigl(\sum_{i=1}^{\infty}\mu_i f_i^2\Bigr)\\ &\leqslant \sum_{i=1}^{\infty}\mu_i\Phi(f_i)+h\Bigl(\sum_{i=1}^{\infty}\mu_i f_i^2\Bigr),\end{aligned}$$

where $h(x)=2x-x\log x$ for $x>0$. As h is an increasing function on $[0,e]$ and

$$\|\boldsymbol{f}\|_{L^1_\mu}\leqslant\|\boldsymbol{f}\|_{L^2_\mu}\leqslant\sqrt{\tfrac{3}{2}}\|\boldsymbol{f}\|_{L_\Phi},$$

(see Lemma A.2 in Appendix A) we have

$$\|\boldsymbol{f}\|_{L^2_\mu}^2\leqslant 2.$$

Thus, as

$$\sum_{i=1}^{\infty}\mu_i\Phi(f_i)=\sum_{i=1}^{\infty}\mu_i\Phi\Bigl(\frac{f_i}{\|f\|_{L_\Phi}}\Bigr)\leqslant 1,$$

we find that

$$\mathcal{L}(\boldsymbol{f})\leqslant 1+h(2)\leqslant 5,$$

proving the right-hand side inequality of (2-12). To show the left-hand side of the inequality we assume that $\mathcal{L}(\boldsymbol{f})=2$. By the definition of $\mathcal{L}$ and the fact that

$$\|\boldsymbol{f}-\langle\boldsymbol{f}\rangle\|_{L^2_\mu}^2=\tfrac{1}{2}\lim_{|a|\to\infty}\mathrm{Ent}_\mu((\boldsymbol{f}+a)^2)$$

(see Lemma A.3 in Appendix A), we know that

$$\|\boldsymbol{f}\|_{L^2_\mu}^2 \leqslant \tfrac{1}{2}\mathcal{L}(\boldsymbol{f}) = 1.$$

This implies

$$\sum_{i=1}^\infty \mu_i \Phi(f_i) \leqslant 1 + \sum_{i=1}^\infty \mu_i f_i^2 \log f_i^2 = 1 + \mathrm{Ent}_\mu(\boldsymbol{f}^2) + \|\boldsymbol{f}\|_{L^2_\mu}^2 \log(\|\boldsymbol{f}\|_{L^2_\mu}^2) \leqslant 1 + \mathcal{L}(\boldsymbol{f}) = 3,$$

where we have used the fact that $x\log(1+x) \leqslant 1 + x\log x$ when $x > 0$.

Since, for any $a \geqslant 1$,

$$\Phi\left(\frac{x}{\sqrt{a}}\right) = \frac{x^2}{a^2}\log\left(1+\frac{x^2}{a^2}\right) \leqslant \frac{1}{a^2}\Phi(x),$$

the above implies

$$\sum_{i=1}^\infty \mu_i \Phi\left(\frac{f_i}{\sqrt{3}}\right) \leqslant 1$$

and as such, by the definition of $\|\cdot\|_{L_\Phi}$, we conclude that

$$\|\boldsymbol{f}\|_{L_\Phi}^2 \leqslant 3 = \tfrac{3}{2}\mathcal{L}(\boldsymbol{f}),$$

and the proof is complete. □

Proof of Proposition 2.9. The equivalence of (ii) and (iii) is immediate following Lemma 2.10, which also proves the desired connection between C_{LS} and λ. To show that (i) implies (ii) we notice that as the right-hand side of (2-2) is invariant under translation, taking the supremum over all possible translations results in (ii). The fact that (ii) implies (i) is immediate as $\mathrm{Ent}_\mu(\boldsymbol{f}^2) \leqslant \mathcal{L}(\boldsymbol{f})$. □

Discrete Hardy inequalities. The above observation that the log-Sobolev inequality with weights is actually a form of a Poincaré inequality brings to mind another inequality with weights that is closely connected to the Poincaré inequality — the Hardy inequality. In its discrete form, we have:

Lemma 2.11. *Let $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ be two sequences of positive numbers and let $m \in \mathbb{N}$. Then, the following two conditions are equivalent*:

(i) *There exists a finite constant $A_{1,m} \geqslant 0$ such that*

$$\sum_{i=m}^\infty \mu_i \left(\sum_{j=m}^i f_j\right)^2 \leqslant A_{1,m} \sum_{i=m}^\infty \nu_i f_i^2$$

for any sequence $\boldsymbol{f}$.

(ii) *We have*

$$B_{1,m} = \sup_{k\geqslant m}\left(\sum_{i=k}^\infty \mu_i\right)\left(\sum_{i=m}^k \frac{1}{\nu_i}\right) < \infty.$$

Moreover, if any of the conditions holds then $B_{1,m} \leqslant A_{1,m} \leqslant 4B_{1,m}$.

The proof for the case $m = 1$ can be found in [Cañizo and Lods 2013], and the general case follows by the same method of proof.

Corollary 2.12. *Let*

$$B_m^{(1)} = \sup_{k \geqslant m} \left(\sum_{i=k+1}^{\infty} \mu_i \right) \left(\sum_{i=m}^{k} \frac{1}{\nu_i} \right).$$

Then for any sequence $\boldsymbol{f}$ such that $f_m = 0$, we have

$$\sum_{i=m}^{\infty} \mu_i f_i^2 \leqslant A_m^{(1)} \sum_{i=m}^{\infty} \nu_i (f_{i+1} - f_i)^2 \tag{2-13}$$

if and only if $B_m^{(1)} < \infty$. In that case $B_m^{(1)} \leqslant A_m^{(1)} \leqslant 4 B_m^{(1)}$. Additionally,

$$B_{1,m} \leqslant B_m^{(1)} \leqslant B_{1,m+1}.$$

Proof. This follows immediately from Lemma 2.11 applied to the sequence $g_i = f_{i+1} - f_i$ and a simple translation argument. ☐

Besides the above, we will also need to have a Hardy-type inequality for sums up to a fixed integer m.

Lemma 2.13. *Let $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ be two sequences of positive numbers and let $m \in \mathbb{N}$. Then, for any sequence $\boldsymbol{f}$ such that $f_m = 0$, we have that if there exists $A > 0$ such that*

$$\sum_{i=1}^{m-1} \mu_i f_i^2 \leqslant A \sum_{i=1}^{m-1} \nu_i (f_{i+1} - f_i)^2, \tag{2-14}$$

then $b_{2,m} \leqslant A$, where

$$b_{2,m} = \sup_{k \leqslant m-1} \sum_{i=1}^{k} \mu_i \left(\sum_{j=k}^{m-1} \frac{1}{\nu_j} \right).$$

Moreover, one can always choose

$$A = B_{2,m} = \sum_{i=1}^{m-1} \mu_i \left(\sum_{j=i}^{m-1} \frac{1}{\nu_j} \right).$$

Proof. We start by noticing that for any $1 \leqslant i \leqslant m-1$, we have

$$f_i^2 = \left(\sum_{j=i}^{m-1} (f_{j+1} - f_j) \right)^2 \leqslant \left(\sum_{j=i}^{m-1} \frac{1}{\nu_j} \right) \left(\sum_{j=i}^{m-1} \nu_j (f_{j+1} - f_j)^2 \right) \leqslant \left(\sum_{j=i}^{m-1} \frac{1}{\nu_j} \right) \left(\sum_{j=1}^{m-1} \nu_j (f_{j+1} - f_j)^2 \right).$$

Thus

$$\sum_{i=1}^{m-1} \mu_i f_i^2 \leqslant \left(\sum_{i=1}^{m-1} \mu_i \left(\sum_{j=i}^{m-1} \frac{1}{\nu_j} \right) \right) \left(\sum_{j=1}^{m-1} \nu_j (f_{j+1} - f_j)^2 \right) = B_{2,m} \sum_{j=1}^{m-1} \nu_j (f_{j+1} - f_j)^2,$$

completing the second statement. Next, for any $j \leqslant m-1$ we set

$$\sigma_j = \sum_{i=j}^{m-1} \frac{1}{\nu_i}.$$

Fix $k \leqslant m-1$ and define $\boldsymbol{f}^{(k)}$ to be such that $f_i^{(k)} = \sigma_k$ when $i \leqslant k$ and $f_i^{(k)} = \sigma_i$ when $i > k$. We have

$$\sum_{i=1}^{m-1} \nu_i (f_{i+1}^{(k)} - f_i^{(k)})^2 = \sum_{i=k}^{m-1} \nu_i (f_{i+1}^{(k)} - f_i^{(k)})^2 = \sum_{i=k}^{m-1} \frac{1}{\nu_i} = \sigma_k.$$

On the other hand,

$$\sum_{i=1}^{m-1} \mu_i (f_i^{(k)})^2 \geqslant \sum_{i=1}^{k} \mu_i (f_i^{(k)})^2 = \sigma_k^2 \Bigl(\sum_{i=1}^{k} \mu_i \Bigr).$$

As (2-14) is valid, we see that

$$A \geqslant \Bigl(\sum_{i=k}^{m-1} \frac{1}{\nu_i} \Bigr) \Bigl(\sum_{i=1}^{k} \mu_i \Bigr)$$

for all k. □

Proof of the main inequality. The last ingredient we need in order to prove Theorem 2.3 is the following lemma:

Lemma 2.14. *The following conditions are equivalent*:

(i) *$\boldsymbol{\nu}$ admits a log-Sobolev inequality with respect to $\boldsymbol{\mu}$ with constant C_{LS}.*

(ii) *There exists $\eta > 0$ such that, for any sequence $\boldsymbol{f} = \{f_i\}$ such that $f_m = 0$ with $m \in \mathbb{N}$ satisfying*

$$\max\Bigl(\sum_{i=1}^{m-1} \mu_i, \sum_{i=m+1}^{\infty} \mu_i \Bigr) < \frac{2}{3},$$

we have

$$\|(\boldsymbol{f}^{(0)})^2\|_{L_\Psi} + \|(\boldsymbol{f}^{(1)})^2\|_{L_\Psi} \leqslant \eta \sum_{i=1}^{\infty} \nu_i (f_{i+1} - f_i)^2,$$

where $\boldsymbol{f}^{(0)} = \boldsymbol{f}\, \mathbb{1}_{i<m}$ and $\boldsymbol{f}^{(1)} = \boldsymbol{f}\, \mathbb{1}_{i>m}$.

Moreover, if condition (ii) *is valid, one can choose $C_{\mathrm{LS}} = 40\eta$.*

Proof. We notice that it is enough for us to show the equivalence of condition (ii) of our lemma and Proposition 2.9(ii).

Assume, to begin with, that Proposition 2.9(ii) is valid. As was shown in that proposition, this implies

$$\|\boldsymbol{f} - \langle \boldsymbol{f} \rangle\|_{L_\Phi}^2 \leqslant \frac{3C_{\mathrm{LS}}}{2} \sum_{i=1}^{\infty} \nu_i (f_{i+1} - f_i)^2. \tag{2-15}$$

Due to the conditions on $\boldsymbol{f}$ and the definitions of $\boldsymbol{f}^{(0)}$ and $\boldsymbol{f}^{(1)}$, one has that

$$\|\langle \boldsymbol{f}^{(0)}\rangle\|_{L_\Phi} \leqslant |\langle \boldsymbol{f}^{(0)}\rangle| \leqslant \|\boldsymbol{f}^{(0)}\|_{L^2_\mu}\Bigl(\sum_{i=1}^{m-1}\mu_i\Bigr)^{\frac{1}{2}},$$

$$\|\langle \boldsymbol{f}^{(1)}\rangle\|_{L_\Phi} \leqslant |\langle \boldsymbol{f}^{(1)}\rangle| \leqslant \|\boldsymbol{f}^{(1)}\|_{L^2_\mu}\Bigl(\sum_{i=m+1}^{\infty}\mu_i\Bigr)^{\frac{1}{2}}$$

(see Lemma A.4 in Appendix A). Thus

$$\|\boldsymbol{f}^{(0)}\|_{L_\Phi} \leqslant \|\boldsymbol{f}^{(0)}-\langle \boldsymbol{f}^{(0)}\rangle\|_{L_\Phi}+\|\langle \boldsymbol{f}^{(0)}\rangle\|_{L_\Phi} \leqslant \|\boldsymbol{f}^{(0)}-\langle \boldsymbol{f}^{(0)}\rangle\|_{L_\Phi}+\Bigl(\frac{3}{2}\sum_{i=1}^{m-1}\mu_i\Bigr)\|\boldsymbol{f}^{(0)}\|_{L_\Phi},$$

implying

$$\|\boldsymbol{f}^{(0)}\|_{L_\Phi} \leqslant \frac{1}{1-\sqrt{\frac{3}{2}\sum_{i=1}^{m-1}\mu_i}}\|\boldsymbol{f}^{(0)}-\langle \boldsymbol{f}^{(0)}\rangle\|_{L_\Phi},$$

and similarly

$$\|\boldsymbol{f}^{(1)}\|_{L_\Phi} \leqslant \frac{1}{1-\sqrt{\frac{3}{2}\sum_{i=m+1}^{\infty}\mu_i}}\|\boldsymbol{f}^{(1)}-\langle \boldsymbol{f}^{(1)}\rangle\|_{L_\Phi}.$$

We can conclude, by applying (2-15) to $\boldsymbol{f}^{(0)}$ and $\boldsymbol{f}^{(1)}$, that

$$\|\boldsymbol{f}^{(0)}\|^2_{L_\Phi} \leqslant \frac{3C_{\mathrm{LS}}}{2\bigl(1-\sqrt{\frac{3}{2}\sum_{i=1}^{m-1}\mu_i}\bigr)^2}\sum_{i=1}^{m-1}\nu_i(f_{i+1}-f_i)^2,$$

$$\|\boldsymbol{f}^{(1)}\|^2_{L_\Phi} \leqslant \frac{3C_{\mathrm{LS}}}{2\bigl(1-\sqrt{\frac{3}{2}\sum_{i=m+1}^{\infty}\mu_i}\bigr)^2}\sum_{i=m}^{\infty}\nu_i(f_{i+1}-f_i)^2.$$

The result now follows from (2-7).

To show the converse, we use the translation invariance of Proposition 2.9(ii) to assume that $f_m=0$. As such we have $\boldsymbol{f}=\boldsymbol{f}^{(0)}+\boldsymbol{f}^{(1)}$. Moreover,

$$\begin{aligned}
\|\boldsymbol{f}-\langle \boldsymbol{f}\rangle\|^2_{L_\Phi} &\leqslant \bigl(\|\boldsymbol{f}^{(0)}-\langle \boldsymbol{f}^{(0)}\rangle\|_{L_\Phi}+\|\boldsymbol{f}^{(1)}-\langle \boldsymbol{f}^{(1)}\rangle\|_{L_\Phi}\bigr)^2\\
&\leqslant \biggl(\Bigl(1+\Bigl(\frac{3}{2}\sum_{i=1}^{m-1}\mu_i\Bigr)^{\frac{1}{2}}\Bigr)\|\boldsymbol{f}^{(0)}\|_{L_\Phi}+\Bigl(1+\Bigl(\frac{3}{2}\sum_{i=m+1}^{\infty}\mu_i\Bigr)^{\frac{1}{2}}\Bigr)\|\boldsymbol{f}^{(1)}\|_{L_\Phi}\biggr)^2\\
&\leqslant 2\Bigl(1+\Bigl(\frac{3}{2}\sum_{i=1}^{m-1}\mu_i\Bigr)^{\frac{1}{2}}\Bigr)^2\|\boldsymbol{f}^{(0)}\|^2_{L_\Phi}+2\Bigl(1+\Bigl(\frac{3}{2}\sum_{i=m+1}^{\infty}\mu_i\Bigr)^{\frac{1}{2}}\Bigr)^2\|\boldsymbol{f}^{(1)}\|^2_{L_\Phi}\\
&\leqslant 2\eta\max\biggl(\Bigl(1+\Bigl(\frac{3}{2}\sum_{i=1}^{m-1}\mu_i\Bigr)^{\frac{1}{2}}\Bigr)^2,\Bigl(1+\Bigl(\frac{3}{2}\sum_{i=m+1}^{\infty}\mu_i\Bigr)^{\frac{1}{2}}\Bigr)^2\biggr)\sum_{i=1}^{\infty}\nu_i(f_{i+1}-f_i)^2,
\end{aligned}$$

where we again used (2-7). This shows the desired result due to Proposition 2.9. □

Proof of Theorem 2.3. Our main tool will be Lemma 2.14. It is known that

$$\|\boldsymbol{f}^2\|_{L_\Psi} = \sup\left\{\sum_{i=1}^{\infty}\mu_i f_i^2 g_i : \sum_{i=1}^{\infty}\mu_i\,\Xi(g_i)\leqslant 1\right\},$$

where Ξ is the Young complement of Ψ. Using Corollary 2.12, we know that if $f_m = 0$ then

$$\sum_{i=m}^{\infty}\mu_i f_i^2 g_i \leqslant C_{\mathrm{LS}}\sum_{i=m}^{\infty}\nu_i(f_{i+1}-f_i)^2$$

if and only if

$$B = \sup_{k\geqslant m}\left(\sum_{i=k+1}^{\infty} g_i\mu_i\right)\left(\sum_{i=1}^{k}\frac{1}{\nu_i}\right) < \infty.$$

Taking supremum over all appropriate $\boldsymbol{g} = \{g_i\}$, we find that

$$\|\boldsymbol{f}^2\mathbb{1}_{i>m}\|_{L_\Psi} \leqslant C_{\mathrm{LS}}\sum_{i=m}^{\infty}\nu_i(f_{i+1}-f_i)^2 \tag{2-16}$$

if and only if

$$B = \sup_{k\geqslant m}\|\mathbb{1}_{[k+1,\infty)}\|_{L_\Psi}\sum_{i=1}^{k}\frac{1}{\nu_i} < \infty.$$

As

$$\begin{aligned}\|\mathbb{1}_{[k+1,\infty)}\|_{L_\Psi} &= \inf_{\alpha>0}\left\{\sum_{i=k+1}^{\infty}\mu_i\Psi\left(\frac{1}{\alpha}\right)\leqslant 1\right\} = \inf_{\alpha>0}\left\{\Psi\left(\frac{1}{\alpha}\right)\leqslant\frac{1}{\sum_{i=k+1}^{\infty}\mu_i}\right\}\\ &= \frac{1}{\Psi^{-1}\left(1/\sum_{i=k+1}^{\infty}\mu_i\right)},\end{aligned}$$

we find that (2-16) is equivalent to $B_1 < \infty$, showing that (i) implies (ii).

Conversely, using Lemma 2.13 we find that if $f_m = 0$ then

$$\begin{aligned}\sum_{i=1}^{m-1}\mu_i f_i^2 g_i &\leqslant \left[\sum_{i=1}^{m-1}\mu_i g_i\left(\sum_{j=i}^{m-1}\frac{1}{\nu_j}\right)\right]\sum_{i=1}^{m-1}\nu_i(f_{i+1}-f_i)^2\\ &\leqslant \left[\left(\sum_{i=1}^{m-1}\mu_i g_i\right)\left(\sum_{j=1}^{m-1}\frac{1}{\nu_j}\right)\right]\sum_{i=1}^{m-1}\nu_i(f_{i+1}-f_i)^2\end{aligned}$$

and again, by taking the supremum over the appropriate $\boldsymbol{g}$, we find that

$$\|\boldsymbol{f}^2\mathbb{1}_{i<m}\|_{L_\Psi} \leqslant B_2\sum_{i=1}^{m-1}\nu_i(f_{i+1}-f_i)^2. \tag{2-17}$$

Thus, if $\boldsymbol{f} = \{f_i\}$ is a sequence such that $f_m = 0$, and if in addition $B_1 < \infty$, we have

$$\|(\boldsymbol{f}^{(0)})^2\|_{L_\Psi} + \|(\boldsymbol{f}^{(1)})^2\|_{L_\Psi} \leqslant B_2 \sum_{i=1}^{m-1} \nu_i (f_{i+1} - f_i)^2 + 4B_1 \sum_{i=m}^{\infty} \nu_i (f_{i+1} - f_i)^2$$

$$\leqslant (B_2 + 4B_1) \sum_{i=1}^{\infty} \nu_i (f_{i+1} - f_i)^2,$$

where we have used Corollary 2.12. We conclude, using Lemma 2.14, that if $B_1 < \infty$, then $\boldsymbol{\nu}$ admits a log-Sobolev inequality with respect to $\boldsymbol{\mu}$ with constant C_{LS} that can be chosen to be $C_{\mathrm{LS}} = 40(B_1 + 4B_2)$. □

We are only left with the proof of Corollary 2.4. The proof relies on the following technical lemma, whose proof is left to Appendix A:

Lemma 2.15. *For any $t \geqslant \frac{3}{2}$, one has*

$$\frac{1}{3} \frac{t}{\log t} \leqslant \Psi^{-1}(t) \leqslant 2 \frac{t}{\log t}.$$

Proof of Corollary 2.4. Due to the choice of m and Lemma 2.15, we know that $\Psi^{-1}(t)$ and $t/\log t$ are equivalent for our choice of

$$t = \frac{1}{\sum_{i=m+1}^{\infty} \mu_i}.$$

This shows the desired equivalence using Theorem 2.3. As for the last estimation, it follows immediately from the fact that

$$B_i \leqslant 3D_i$$

for $i = 1, 2$. □

Now that we have achieved a necessary and sufficient condition for the validity of a discrete log-Sobolev inequality with weight, we will proceed to see how it can be used to prove Theorem 1.1.

3. Energy dissipation inequalities

The log-Sobolev inequality and the Becker–Döring equations. Motivated by our previous section, the first step in trying to show the validity of Cercignani's conjecture would be to relate the energy dissipation, $D(\boldsymbol{c})$, and a term that resembles the right-hand side of (2-2). Recall that, for any nonnegative sequence $\boldsymbol{c} = \{c_i\}$ we defined

$$D(\boldsymbol{c}) = \sum_{i=1}^{\infty} a_i Q_i \, \Theta\left(\frac{c_1 c_i}{Q_i}, \frac{c_{i+1}}{Q_{i+1}}\right)$$

with $\Theta(x, y) := (x - y)(\log x - \log y)$, and

$$\overline{D}(\boldsymbol{c}) = \sum_{i=1}^{\infty} a_i Q_i \left(\sqrt{\frac{c_1 c_i}{Q_i}} - \sqrt{\frac{c_{i+1}}{Q_{i+1}}}\right)^2.$$

We have the following properties:

Lemma 3.1. *For any nonnegative sequence* $\boldsymbol{c}$, *the following holds*:

(i) *We have*

$$4\overline{D}(\boldsymbol{c}) \leqslant D(\boldsymbol{c}). \tag{3-1}$$

(ii) *For any* $z > 0$, *we can rewrite* $D(\boldsymbol{c})$ *as*

$$D(\boldsymbol{c}) = \sum_{i=1}^{\infty} a_i Q_i z^{i+1} \Theta\left(\frac{c_1 c_i}{Q_i z^{i+1}}, \frac{c_{i+1}}{Q_{i+1} z^{i+1}}\right), \tag{3-2}$$

recalling $\Theta(x, y) := (x - y)(\log x - \log y)$, *and*

$$\overline{D}(\boldsymbol{c}) = \sum_{i=1}^{\infty} a_i Q_i z^{i+1} \left(\sqrt{\frac{c_1 c_i}{Q_i z^{i+1}}} - \sqrt{\frac{c_{i+1}}{Q_{i+1} z^{i+1}}}\right)^2. \tag{3-3}$$

Proof. Part (i) is an immediate consequence of the inequality

$$\Theta(x, y) = (x - y)(\log x - \log y) \geqslant 4(\sqrt{x} - \sqrt{y})^2$$

and (ii) is immediate from the homogeneity of the expressions involved. □

Property (ii) of the above lemma gives an indication of how we may be able to find a connection between $\overline{D}(\boldsymbol{c})$ and the relative entropy between $\boldsymbol{c}$ and *some* equilibrium, by appropriately choosing z. Similar to the work of Jabin and Niethammer [2003], another equilibrium state that will play an important role in what is to follow is

$$\widetilde{\mathcal{Q}} = \mathcal{Q}_{c_1} = \{Q_i c_1^i\}_{i \geqslant 1}.$$

Indeed, it is the *only* possible equilibrium under which the right-hand side of (3-3) attains a form that is suitable for the log-Sobolev theory developed in the previous section. From (3-3) we find, after cancelling c_1, that

$$\overline{D}(\boldsymbol{c}) = \sum_{i=1}^{\infty} a_i \widetilde{\mathcal{Q}}_i \widetilde{\mathcal{Q}}_1 \left(\sqrt{\frac{c_i}{\widetilde{\mathcal{Q}}_i}} - \sqrt{\frac{c_{i+1}}{\widetilde{\mathcal{Q}}_{i+1}}}\right)^2. \tag{3-4}$$

This enables us to finally link $\overline{D}(\boldsymbol{c})$ to $H(\boldsymbol{c}|\mathcal{Q})$:

Proposition 3.2. *For given coagulation and detailed balance coefficients* $\{a_i\}_{i\in\mathbb{N}}$ *and* $\{Q_i\}_{i\in\mathbb{N}}$ *and a given positive sequence* $\boldsymbol{c}$ *with finite mass* ϱ *and such that*

$$\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i < +\infty, \quad \sum_{i=1}^{\infty} a_i \widetilde{\mathcal{Q}}_i < +\infty$$

(*recall* $\mathcal{Q}_i := Q_i c_1^i$ *for* $i \geqslant 1$), *we define the measures*

$$\mu_i = \frac{\widetilde{\mathcal{Q}}_i}{\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i}, \quad \nu_i := \frac{a_i \widetilde{\mathcal{Q}}_i}{\sum_{j=1}^{\infty} a_j \widetilde{\mathcal{Q}}_j}, \quad i \in \mathbb{N}. \tag{3-5}$$

Then, if $\boldsymbol{\nu}$ admits a log-Sobolev inequality with respect to $\boldsymbol{\mu}$ with constant C_{LS}, we have

$$\overline{D}(\boldsymbol{c}) \geqslant \frac{c_1^3\big(\sum_{i=1}^{\infty} a_i \widetilde{\mathcal{Q}}_i\big)}{C_{\mathrm{LS}}\big(\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i\big)\big(c_1^2 + 2\big(\sum_{i=1}^{\infty} c_i\big)\big(\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i\big)\big)} H(\boldsymbol{c}|\mathcal{Q}). \tag{3-6}$$

Proof. Let $f_i = \sqrt{c_i/\widetilde{\mathcal{Q}}_i}$. Since $\boldsymbol{\nu}$ admits a log-Sobolev inequality with respect to $\boldsymbol{\mu}$ with constant C_{LS}, we have

$$\overline{D}(\boldsymbol{c}) = \left(\sum_{i=1}^{\infty} a_i \widetilde{\mathcal{Q}}_i \widetilde{\mathcal{Q}}_1\right) \sum_{i=1}^{\infty} \nu_i (f_{i+1} - f_i)^2 \geqslant \frac{c_1\big(\sum_{i=1}^{\infty} a_i \widetilde{\mathcal{Q}}_i\big)}{C_{\mathrm{LS}}} \operatorname{Ent}_{\mu}(\boldsymbol{f}^2). \tag{3-7}$$

Next, we notice that

$$\begin{aligned}
\left(\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i\right) \operatorname{Ent}_{\mu}(\boldsymbol{f}^2) &= \sum_{i=1}^{\infty} c_i \log \frac{c_i}{\widetilde{\mathcal{Q}}_i} - \left(\sum_{i=1}^{\infty} c_i\right)\left(\log \sum_{i=1}^{\infty} c_i - \log \sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i\right) \\
&= H(\boldsymbol{c}|\widetilde{\mathcal{Q}}) + \sum_{i=1}^{\infty} c_i - \sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i - \left(\sum_{i=1}^{\infty} c_i\right)\left(\log \sum_{i=1}^{\infty} c_i - \log \sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i\right) \\
&= H(\boldsymbol{c}|\widetilde{\mathcal{Q}}) - \left(\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i\right) \Lambda\left(\frac{\sum_{i=1}^{\infty} c_i}{\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i}\right),
\end{aligned} \tag{3-8}$$

where $\Lambda(x) = x \log x - x + 1$. We now use the fact that $\mathcal{Q}$ minimises the relative entropy to the set of equilibria to bound the first term,

$$H(\boldsymbol{c}|\widetilde{\mathcal{Q}}) \geqslant H(\boldsymbol{c}|\mathcal{Q}) \tag{3-9}$$

(see Lemma B.1 in Appendix B). The only remaining bound is to show that the term with the negative sign at the end of (3-8) is in fact bounded by $\operatorname{Ent}_{\mu}(\boldsymbol{f}^2)$. For this we will use the following Csiszár–Kullback inequality:

$$\operatorname{Ent}_{\mu}(\boldsymbol{f}^2) \geqslant \frac{1}{2\langle \boldsymbol{f}^2 \rangle}\left(\sum_{i=1}^{\infty} |f_i^2 - \langle \boldsymbol{f}^2 \rangle| \mu_i\right)^2, \tag{3-10}$$

where

$$\langle \boldsymbol{f}^2 \rangle := \sum_{i=1}^{\infty} f_i^2 \mu_i .$$

With (3-10), we find that in our particular setting

$$\operatorname{Ent}_{\mu}(\boldsymbol{f}^2) \geqslant \frac{\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i}{2\sum_{i=1}^{\infty} c_i}\left(\sum_{i=1}^{\infty}\left|\frac{c_i}{\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i} - \frac{\widetilde{\mathcal{Q}}_i\big(\sum_{i=1}^{\infty} c_i\big)}{\big(\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i\big)^2}\right|\right)^2 = \frac{\sum_{i=1}^{\infty} c_i}{2\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i}\left(\sum_{i=1}^{\infty}\left|\frac{c_i}{\sum_{i=1}^{\infty} c_i} - \frac{\widetilde{\mathcal{Q}}_i}{\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i}\right|\right)^2,$$

and keeping only the first term in the last sum we get

$$\operatorname{Ent}_{\mu}(\boldsymbol{f}^2) \geqslant \frac{\sum_{i=1}^{\infty} c_i}{2\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i}\left|\frac{c_1}{\sum_{i=1}^{\infty} c_i} - \frac{\widetilde{\mathcal{Q}}_1}{\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i}\right|^2 = \frac{c_1^2}{2\sum_{i=1}^{\infty} c_i \sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i}\left(1 - \frac{\sum_{i=1}^{\infty} c_i}{\sum_{i=1}^{\infty} \widetilde{\mathcal{Q}}_i}\right)^2.$$

Continuing from (3-8) and using (3-9), the above inequality and the fact that

$$\Lambda(x) \leqslant (x-1)^2$$

show that

$$\begin{aligned}\Big(\sum_{i=1}^{\infty}\widetilde{\mathcal{Q}}_i\Big)\,\mathrm{Ent}_{\mu}(\boldsymbol{f}^2) &\geqslant H(\boldsymbol{c}|\mathcal{Q}) - \Big(\sum_{i=1}^{\infty}\widetilde{\mathcal{Q}}_i\Big)\bigg(\frac{\sum_{i=1}^{\infty}c_i}{\sum_{i=1}^{\infty}\widetilde{\mathcal{Q}}_i}-1\bigg)^2\\ &\geqslant H(\boldsymbol{c}|\mathcal{Q}) - \frac{2}{c_1^2}\Big(\sum_{i=1}^{\infty}\widetilde{\mathcal{Q}}_i\Big)^2\Big(\sum_{i=1}^{\infty}c_i\Big)\,\mathrm{Ent}_{\mu}(\boldsymbol{f}^2).\end{aligned}$$

Thus,

$$H(\boldsymbol{c}|\mathcal{Q}) \leqslant \Big(\sum_{i=1}^{\infty}\widetilde{\mathcal{Q}}_i\Big)\bigg(1+\frac{2}{c_1^2}\Big(\sum_{i=1}^{\infty}\widetilde{\mathcal{Q}}_i\Big)\Big(\sum_{i=1}^{\infty}c_i\Big)\bigg)\,\mathrm{Ent}_{\mu}(\boldsymbol{f}^2).$$

Combining the above with (3-7) completes the proof. □

Main inequality for c_1 "close" to equilibrium. On the basis of Proposition 3.2, one obtains the following.

Proposition 3.3. *Assume the conditions of Proposition 3.2 and the additional condition that $c_1 < z_*$ for some $0 < z_* < z_{\mathrm{s}}$. Setting*

$$\varrho_* := \sum_{i=1}^{\infty} i Q_i z_*^i < \infty,$$

we have

$$\overline{D}(\boldsymbol{c}) \geqslant \frac{a_1 z_*^2 c_1^2}{C_{\mathrm{LS}}(z_*+\varrho_*)(z_*^2+2\varrho(z_*+\varrho_*))} H(\boldsymbol{c}|\mathcal{Q}). \tag{3-11}$$

In particular, if $0 < \delta < c_1 < z_{\mathrm{s}} - \delta$ for some $\delta > 0$,

$$\overline{D}(\boldsymbol{c}) \geqslant \lambda H(\boldsymbol{c}|\mathcal{Q})$$

for some constant $\lambda > 0$ which depends only on δ, ρ, a_1 and $\{Q_i\}_{i\geqslant 1}$.

Proof. This follows immediately from (3-6) and the estimates

$$\sum_{i=1}^{\infty}\widetilde{\mathcal{Q}}_i = \sum_{i=1}^{\infty} Q_i c_1^i \leqslant c_1\Big(1+\frac{1}{z_*}\sum_{i=2}^{\infty}Q_i z_*^i\Big) < c_1\Big(1+\frac{\varrho_*}{z_*}\Big),$$

$$\sum_{i=1}^{\infty} c_i \leqslant \sum_{i=1}^{\infty} i c_i = \varrho,$$

together with $\sum_{i=1}^{\infty} a_i\widetilde{\mathcal{Q}}_i \geqslant a_1 c_1$. □

Proposition 3.3 shows us that as long as c_1 is bounded away from 0 and z_{s}, Cercignani's conjecture will follow immediately from a log-Sobolev inequality for $\boldsymbol{\nu}$ with respect to $\boldsymbol{\mu}$ (which were defined in Proposition 3.2). Our next result shows that this is indeed true for subcritical masses, under reasonable conditions on the coefficients:

Proposition 3.4. *Let $\{a_i\}_{i\in\mathbb{N}}$ and $\{Q_i\}_{i\in\mathbb{N}}$ satisfy Hypotheses 1–3 with $\gamma = 1$ and let $\boldsymbol{c} = \{c_i\}_{i\in\mathbb{N}}$ be an arbitrary positive sequence with finite total density $\varrho < \varrho_{\mathrm{s}} < +\infty$. Assume that there exists $\delta > 0$ such that*

$$c_1 \leqslant z_{\mathrm{s}} - \delta.$$

Then, the measure $\boldsymbol{\nu}$ admits a log-Sobolev inequality with respect to the measure $\boldsymbol{\mu}$ with constant

$$C_{\mathrm{LS}} = \frac{60 z_{\mathrm{s}}^3}{\delta^3} C\left(\frac{z_{\mathrm{s}}-\delta}{z_{\mathrm{s}}}\right)\left(4 + 2e \sup_k \bigl|\log\bigl(\alpha_{k+1}^{\frac{1}{k+1}}\bigr)\bigr| + e \log \frac{z_{\mathrm{s}}}{\delta}\right), \tag{3-12}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ were defined in Proposition 3.2 and

$$C(\eta) = 1 + \sup_{k\geqslant 3}\Bigl(k\bigl(1+\log\bigl(\tfrac{1}{2}k\bigr)\bigr)\eta^{\frac{k}{2}}\Bigr) + \frac{2\eta}{1-\eta}$$

for $\eta < 1$.

Proof. We just need to estimate the constant given in Corollary 2.4. As mentioned in the Introduction, we can assume without loss of generality that $a_i = i$. We define

$$\eta = \frac{c_1}{z_{\mathrm{s}}} \leqslant \frac{z_{\mathrm{s}}-\delta}{z_{\mathrm{s}}} =: \eta_1 < 1.$$

As

$$\widetilde{\mathcal{Q}}_i = \alpha_i z_{\mathrm{s}}^{1-i} c_1^i \leqslant z_{\mathrm{s}} \alpha_i \eta^i,$$

we find that due to the monotonicity of $\{\alpha_i\}_{i\in\mathbb{N}}$,

$$z_{\mathrm{s}}\alpha_{k+1}\eta^{k+1} = \widetilde{\mathcal{Q}}_{k+1} \leqslant \sum_{i=k+1}^{\infty} \widetilde{\mathcal{Q}}_i \leqslant z_{\mathrm{s}}\eta^{k+1}\sum_{i=1}^{\infty}\alpha_{i+k}\eta^{i-1} \leqslant \frac{z_{\mathrm{s}}\alpha_{k+1}\eta^{k+1}}{1-\eta}.$$

As such

$$\alpha_{k+1}(1-\eta)\eta^k \leqslant \sum_{i=k+1}^{\infty}\mu_i \leqslant \alpha_{k+1}\frac{\eta^k}{1-\eta},$$

implying

$$-\sum_{i=k+1}^{\infty}\mu_i \log\left(\sum_{i=k+1}^{\infty}\mu_i\right) \leqslant \frac{\alpha_{k+1}\eta^k}{1-\eta}\left(k\log\left(\frac{1}{\eta}\right) - \log(\alpha_{k+1}(1-\eta))\right). \tag{3-13}$$

Next, we notice that as

$$\sum_{i=1}^{\infty} i y^i = \frac{y}{(1-y)^2},$$

one has

$$z_{\mathrm{s}}\eta \leqslant \sum_{i=1}^{\infty} i\alpha_i z_{\mathrm{s}}\eta^i = \sum_{i=1}^{\infty} a_i \widetilde{\mathcal{Q}}_i \leqslant z_{\mathrm{s}}\frac{\eta}{(1-\eta)^2},$$

from which we find that

$$i\alpha_i(1-\eta)^2\eta^{i-1} \leqslant \nu_i \leqslant i\alpha_i\eta^{i-1}.$$

We notice that for $k \geqslant 3$, the monotonicity of $\{\alpha_i\}_{i\in\mathbb{N}}$ implies

$$\begin{aligned} k\alpha_k\eta^k \sum_{i=1}^{k} \frac{1}{i\alpha_i}\left(\frac{1}{\eta}\right)^i &= 1 + \sum_{i=1}^{k-1} \frac{k\alpha_k}{i\alpha_i}\eta^{k-i} \\ &\leqslant 1 + \sum_{i=1}^{k-1} \frac{k}{i}\eta^{k-i} = 1 + \sum_{i=1}^{[\frac{k}{2}]} \frac{k}{i}\eta^{k-i} + \sum_{i=[\frac{k}{2}]+1}^{k-1} \frac{k}{i}\eta^{k-i} \\ &\leqslant 1 + k\eta_1^{\frac{k}{2}} \sum_{i=1}^{[\frac{k}{2}]} \frac{1}{i} + \frac{k}{[\frac{k}{2}]+1}\sum_{j=1}^{\infty}\eta_1^j \\ &\leqslant 1 + k\big(1+\log(\tfrac{1}{2}k)\big)\eta_1^{\frac{k}{2}} + \frac{2\eta_1}{1-\eta_1}. \end{aligned}$$

Using the definition of $C(\eta)$ and the fact that $C(\eta) > 1+\eta$, we find that for all $k \in \mathbb{N}$,

$$k\alpha_k\eta^k \sum_{i=1}^{k} \frac{1}{i\alpha_i}\left(\frac{1}{\eta}\right)^i \leqslant C(\eta_1),$$

and as such

$$\sum_{i=1}^{k} \frac{1}{\nu_i} \leqslant C(\eta_1)\frac{\eta}{(1-\eta)^2}\frac{1}{k\alpha_k}\left(\frac{1}{\eta}\right)^k. \tag{3-14}$$

Combining the above with (3-13) yields the bound

$$\left(-\sum_{i=k+1}^{\infty}\mu_i \log\left(\sum_{i=k+1}^{\infty}\mu_i\right)\right)\left(\sum_{i=1}^{k}\frac{1}{\nu_i}\right) \leqslant C(\eta_1)\frac{\alpha_{k+1}}{\alpha_k}\frac{\eta}{(1-\eta)^3}\left(\log\left(\frac{1}{\eta}\right) - \frac{1}{k}\log(\alpha_{k+1}(1-\eta))\right).$$

Thus, with the notation of Corollary 2.4,

$$\begin{aligned} D_1 &\leqslant \frac{C(\eta_1)}{(1-\eta_1)^3}\left(\sup_{0\leqslant x\leqslant 1}(-\eta\log(\eta)) + \eta_1\sup_k \frac{k+1}{k}\left|\log\big(\alpha_{k+1}^{\frac{1}{k+1}}\big)\right| + \eta_1\log\left(\frac{1}{1-\eta_1}\right)\right) \\ &\leqslant \frac{C(\eta_1)}{(1-\eta_1)^3}\left(\frac{1}{e} + 2\eta_1\sup_k\left|\log\big(\alpha_{k+1}^{\frac{1}{k+1}}\big)\right| + \eta_1\log\left(\frac{1}{1-\eta_1}\right)\right). \end{aligned}$$

As m, defined in Corollary 2.4, is always finite, we conclude using the same corollary that $\boldsymbol{\nu}$ admits a log-Sobolev inequality with respect to $\boldsymbol{\mu}$. However, in order to estimate the constant C_{LS}, we still need to estimate the constant D_2 in the case where $m > 1$ (otherwise, $D_2 = 0$).

Since

$$\sum_{i=m}^{\infty}\mu_i \leqslant \frac{\alpha_m}{1-\eta}\eta^{m-1},$$

the requirement that $\sum_{i=1}^{m-1}\mu_i < \frac{2}{3}$ implies

$$\frac{1}{\alpha_{m-1}\eta^{m-1}} \leqslant \frac{\alpha_m}{\alpha_{m-1}}\frac{3}{(1-\eta)} \leqslant \frac{3}{(1-\eta)}.$$

Using the above along with the fact that $m > 1$ and inequality (3-14) shows that

$$\sum_{i=1}^{m-1} \frac{1}{\nu_i} \leqslant 3C(\eta_1)\frac{\eta_1}{(1-\eta_1)^3}\frac{1}{m-1} \leqslant 3C(\eta_1)\frac{\eta_1}{(1-\eta_1)^3}.$$

We can conclude that

$$\left(-\sum_{i=m-1}^{\infty} \mu_i \log\left(\sum_{i=m-1}^{\infty} \mu_i\right)\right)\left(\sum_{i=1}^{m-1}\frac{1}{\nu_i}\right) \leqslant 3 \sup_{0\leqslant x\leqslant 1}(-x\log x)C(\eta_1)\frac{\eta_1}{(1-\eta_1)^3}, \tag{3-15}$$

from which we conclude that

$$D_2 \leqslant \frac{3}{e}C(\eta_1)\frac{\eta_1}{(1-\eta_1)^3},$$

which completes the proof, as the result follows directly from Corollary 2.4. □

We finally have all the tools to prove part (i) of Theorem 1.1:

Proof of part (i) of Theorem 1.1. The result follows immediately from Proposition 3.3, Proposition 3.4 and condition (1-16). □

The last part of this section will be devoted to the proof of part (ii) of Theorem 1.1. For that we will need the following lemma:

Lemma 3.5. *For any $\beta \geqslant 0$, any nonnegative sequence $\boldsymbol{c}$ and positive sequence $\{Q_i\}_{i\geqslant 1}$, it holds that*

$$\sum_{i=1}^{\infty} i^\beta Q_i\left(\sqrt{\frac{c_1c_i}{Q_i}} - \sqrt{\frac{c_{i+1}}{Q_{i+1}}}\right)^2 \leqslant 2\left(c_1 + \sup_j \frac{Q_j}{Q_{j+1}}\right)\sum_{i=1}^{\infty} i^\beta c_i. \tag{3-16}$$

Proof. The proof is a direct consequence of the inequality $(a+b)^2 \leqslant 2(a^2+b^2)$:

$$\begin{aligned}\sum_{i=1}^{\infty} i^\beta Q_i\left(\sqrt{\frac{c_1c_i}{Q_i}} - \sqrt{\frac{c_{i+1}}{Q_{i+1}}}\right)^2 &\leqslant 2c_1\sum_{i=1}^{\infty} i^\beta c_i + 2\sum_{i=1}^{\infty} i^\beta \frac{Q_i}{Q_{i+1}}c_{i+i} \\ &\leqslant 2\left(c_1 + \sup_j \frac{Q_j}{Q_{j+1}}\right)\sum_{i=1}^{\infty} i^\beta c_i.\end{aligned}$$ □

Proof of part (ii) of Theorem 1.1. We denote by $\overline{D}_\gamma(\boldsymbol{c})$ the lower free energy dissipation of $\boldsymbol{c}$ associated to the coagulation coefficient $a_i = i^\gamma$. According to part (i) of Theorem 1.1, there exists $K > 0$ that depends only on δ, z_s, ϱ and $\{\alpha_i\}_{i\in\mathbb{N}}$ such that

$$\overline{D}_1(\boldsymbol{c}) \geqslant KH(\boldsymbol{c}|\mathcal{Q}).$$

Using interpolation between γ and β, we find that

$$\overline{D}_1(c) \leqslant \overline{D}_\gamma^{\frac{\beta-1}{\beta-\gamma}}(c)\overline{D}_\beta^{\frac{1-\gamma}{\beta-\gamma}}(c) \leqslant 2^{\frac{1-\gamma}{\beta-\gamma}}\overline{D}_\gamma^{\frac{\beta-1}{\beta-\gamma}}(c)\left(z_s + \frac{1}{z_s}\sup_j\frac{\alpha_j}{\alpha_{j+1}}\right)^{\frac{1-\gamma}{\beta-\gamma}} M_\beta^{\frac{1-\gamma}{\beta-\gamma}}, \tag{3-17}$$

where we have used Lemma 3.5, the upper bound on c_1 and Hypothesis 2. Therefore

$$D(\boldsymbol{c}) \geqslant \overline{D}_\gamma(\boldsymbol{c}) \geqslant \left(\frac{z_{\mathrm{s}} K^{\frac{\beta-\gamma}{1-\gamma}}}{2(z_{\mathrm{s}}^2 + \sup_j \alpha_j/\alpha_{j+1}) M_\beta}\right)^{\frac{1-\gamma}{\beta-1}} H(\boldsymbol{c}|\mathcal{Q})^{\frac{\beta-\gamma}{\beta-1}} \tag{3-18}$$

and the proof is now complete. ☐

This concludes the part of the proof of Theorem 1.1 that relied on the log-Sobolev inequality. In the next subsection we will address the question of what happens when c_1 escapes the "good region" delimited by (1-16).

Energy dissipation estimate when c_1 is "far" from equilibrium. The goal of this subsection is to show that when c_1 is far from equilibrium, in the aforementioned sense, while we may lose our desired inequality between $\overline{D}(\boldsymbol{c})$ and $H(\boldsymbol{c}|\mathcal{Q})$, the energy dissipation becomes *uniformly large* — forcing the free energy to decrease (and as a consequence, the distance between c_1 and $\bar{z}$ decreases as well).

The next proposition, dealing with the case when c_1 is "too large", is an adaptation of a theorem from [Jabin and Niethammer 2003].

Proposition 3.6. *Let $\{a_i\}_{i\in\mathbb{N}}$ and $\{Q_i\}_{i\in\mathbb{N}}$ be the coagulation and detailed balance coefficients for the Becker–Döring equations. Assume that* $\inf_i a_i > 0$ *and*

$$\lim_{i\to\infty} \frac{Q_{i+1}}{Q_i} = \frac{1}{z_{\mathrm{s}}}.$$

Let $\boldsymbol{c} = \{c_i\}$ be a nonnegative sequence with finite total density $\varrho < \varrho_{\mathrm{s}}$. Then, if

$$c_1 > \bar{z} + \delta$$

for any $\delta > 0$, *we have*

$$\overline{D}(\boldsymbol{c}) > \varepsilon_1$$

for a fixed constant ε_1 that depends only on $\{Q_i\}_{i\in\mathbb{N}}$, $\bar{z}$, z_{s} and δ.

Proof. Without loss of generality we may assume that $\bar{z} + \delta < z_{\mathrm{s}}$. Defining $u_i = c_i/Q_i$ we notice that

$$\overline{D}(\boldsymbol{c}) = \sum_{i=1}^{\infty} a_i Q_i (\sqrt{c_1 u_i} - \sqrt{u_{i+1}})^2.$$

Let $\lambda < 1$ be such that $\lambda c_1 = \bar{z} + \frac{1}{2}\delta$ and let $i_0 \in \mathbb{N}$ be the first index such that

$$u_{i+1} < \lambda c_1 u_i.$$

This index exists, else, for any $i \in \mathbb{N}$ we have

$$u_{i+1} \geqslant \lambda c_1 u_i \geqslant (\lambda c_1)^i c_1, \tag{3-19}$$

and thus

$$\varrho = \sum_{i=1}^{\infty} i c_i \geqslant c_1 + c_1 \sum_{i=2}^{\infty} i Q_i (\lambda c_1)^{i-1} \geqslant \sum_{i=1}^{\infty} i Q_i \left(\bar{z} + \tfrac{1}{2}\delta\right)^i,$$

which is a contradiction.

Due to the positivity of each term in the sum consisting of the lower free energy dissipation, we conclude that

$$\overline{D}(\boldsymbol{c}) \geqslant a_{i_0} Q_{i_0} (1-\sqrt{\lambda})^2 c_1 u_{i_0} \geqslant a_{i_0} Q_{i_0} \lambda^{i_0-1} c_1^{i_0+1} (1-\sqrt{\lambda})^2, \tag{3-20}$$

where we have used the fact that up to $i_0 - 1$ we have inequality (3-19).

As we know that there exists $C > 0$, depending only on $\{Q_i\}_{i\in\mathbb{N}}$, $\bar{z}$, z_s and δ such that

$$\sum_{i=i_0+1}^{\infty} i c_1 (\lambda c_1)^{i-1} Q_i \leqslant C Q_{i_0} (\lambda c_1)^{i_0} c_1$$

(see Lemma B.2 in Appendix B), we conclude that, using (3-19) again,

$$C Q_{i_0} (\lambda c_1)^{i_0} c_1 \geqslant \tilde{\varrho} - \sum_{i=1}^{i_0} i Q_i (\lambda c_1)^{i-1} c_1 \geqslant \tilde{\varrho} - \sum_{i=1}^{i_0} i c_i \geqslant \tilde{\varrho} - \varrho,$$

where $\tilde{\varrho} = \sum_{i=1}^{\infty} i Q_i (\lambda c_1)^{i-1} c_1$. We can estimate the difference $\varrho - \tilde{\varrho}$ as

$$\tilde{\varrho} - \varrho \geqslant \sum_{i=1}^{\infty} i Q_i \left(\left(\bar{z} + \tfrac{1}{2}\delta\right)^i - \bar{z}^i \right) \geqslant \left(\sum_{i=1}^{\infty} i^2 Q_i \bar{z}^{i-1} \right) \tfrac{1}{2}\delta.$$

In conclusion, there exists a universal constant $C_1 > 0$, depending only on $\{Q_i\}_{i\in\mathbb{N}}$, $\bar{z}$, z_s and δ, and not on i_0, c_1 or λ, such that

$$Q_{i_0} (\lambda c_1)^{i_0} c_1 > C_1.$$

Recalling (3-20) and using the fact that

$$\lambda = \frac{\bar{z} + \frac{1}{2}\delta}{c_1} < \frac{\bar{z} + \frac{1}{2}\delta}{\bar{z} + \delta}$$

we find that

$$\overline{D}(\boldsymbol{c}) \geqslant C_1\, a_{i_0} \frac{(1-\sqrt{\lambda})^2}{\lambda} \geqslant C_1 \inf_{i\geqslant 1} a_i \frac{\left(\sqrt{\bar{z}+\delta} - \sqrt{\bar{z} + \frac{1}{2}\delta}\right)^2}{\bar{z} + \frac{1}{2}\delta},$$

completing the proof. □

Next, we present a new lower bound estimate for the energy dissipation in the case where c_1 is "too small".

Lemma 3.7. *Let $\{a_i\}_{i\in\mathbb{N}}$ and $\{Q_i\}_{i\in\mathbb{N}}$ be the coagulation and detailed balance coefficients for the Becker–Döring equations. Assume that*

$$\overline{Q} = \sup_i \frac{Q_i}{Q_{i+1}} < +\infty, \quad \underline{Q} = \inf_i \frac{Q_i}{Q_{i+1}} < +\infty,$$
$$\bar{a} = \sup_i \frac{a_i}{a_{i+1}} < +\infty, \quad \underline{a} = \inf_i \frac{a_i}{a_{i+1}} < +\infty,$$

and let $\boldsymbol{c}$ be a nonnegative sequence such that

$$c_1 < \delta$$

for some $\delta > 0$. *Then*,

$$\overline{D}(\boldsymbol{c}) \geqslant \underline{Q}\underline{a}\left(\sum_{i=1}^{\infty} a_i c_i - a_1\delta\right) - 2\sqrt{\delta}\sqrt{\overline{Q}\bar{a}}\left(\sum_{i=1}^{\infty} a_i c_i\right).$$

Proof. Expanding the square, one has

$$\overline{D}(\boldsymbol{c}) = c_1 \sum_{i=1}^{\infty} a_i c_i + \sum_{i=1}^{\infty} a_i \frac{Q_i}{Q_{i+1}} c_{i+1} - 2\sqrt{c_1} \sum_{i=1}^{\infty} a_i \sqrt{\frac{Q_i}{Q_{i+1}}} \sqrt{c_i c_{i+1}}$$

so that

$$\begin{aligned}\overline{D}(\boldsymbol{c}) &\geqslant \underline{Q}\underline{a}\left(\sum_{i=2}^{\infty} a_i c_i\right) - 2\sqrt{c_1}\sqrt{\overline{Q}\bar{a}}\left(\sum_{i=2}^{\infty} a_i c_i\right)^{\frac{1}{2}}\left(\sum_{i=1}^{\infty} a_i c_i\right)^{\frac{1}{2}} \\ &\geqslant \underline{Q}\underline{a}\left(\sum_{i=1}^{\infty} a_i c_i - a_1\delta\right) - 2\sqrt{\delta}\sqrt{\overline{Q}\bar{a}}\left(\sum_{i=1}^{\infty} a_i c_i\right),\end{aligned}$$

which is the desired result. □

Proposition 3.8. *Let* $\{a_i\}_{i\in\mathbb{N}}$ *and* $\{Q_i\}_{i\in\mathbb{N}}$ *be the coagulation and detailed balance coefficients for the Becker–Döring equations. Assume that*

$$\overline{Q} = \sup_i \frac{Q_i}{Q_{i+1}} < +\infty, \quad \underline{Q} = \inf_i \frac{Q_i}{Q_{i+1}} < +\infty.$$

Let $\boldsymbol{c}$ *be a nonnegative sequence with finite total density* ϱ. *Then*:

(i) *If* $a_i = i$ *then there exists a* $\delta_1 > 0$, *depending only on* $\overline{Q}$, $\underline{Q}$ *and* ϱ *such that if* $c_1 < \delta_1$ *then*

$$\overline{D}(\boldsymbol{c}) \geqslant \frac{\underline{Q}\varrho}{4}.$$

(ii) *If* $a_i = i^\gamma$ *for* $\gamma < 1$ *and there exists* $\beta > 1$ *such that* $M_\beta < +\infty$, *then there exists* $\delta_1 > 0$, *depending only on* $\overline{Q}$, $\underline{Q}$, ϱ *and* M_β *such that if* $c_1 < \delta_1$ *then*

$$\overline{D}(\boldsymbol{c}) \geqslant \frac{\underline{Q}\varrho^{\frac{\beta-\gamma}{\beta-1}}}{4M_\beta^{\frac{1-\gamma}{\beta-1}}}.$$

Proof. Both (i) and (ii) will follow immediately from Lemma 3.7 and a suitable choice of δ_1. Indeed, for (i) we notice that

$$\underline{Q}\underline{a}\left(\sum_{i=1}^{\infty} a_i c_i - a_1\delta\right) - 2\sqrt{\delta}\sqrt{\overline{Q}\bar{a}}\left(\sum_{i=1}^{\infty} a_i c_i\right) = \frac{\underline{Q}}{2}(\varrho - \delta) - 2\sqrt{\delta}\sqrt{\overline{Q}}\varrho,$$

where we have used the notations of Lemma 3.7. As the above is less than $\frac{1}{2}\underline{Q}\varrho$ and converges to it as δ goes to zero, we can find δ_1 that satisfies the desired result.

For (ii) we notice that the interpolation estimate

$$\varrho = \sum_{i=1}^{\infty} i c_i \leqslant \left(\sum_{i=1}^{\infty} i^{\gamma} c_i \right)^{\frac{\beta-1}{\beta-\gamma}} (M_\beta)^{\frac{1-\gamma}{\beta-\gamma}}$$

along with the fact that $\sum_{i=1}^{\infty} i^{\gamma} c_i \leqslant \varrho$ implies

$$\underline{Q}\,\underline{a}\left(\sum_{i=1}^{\infty} a_i c_i - a_1 \delta \right) - 2\sqrt{\delta}\sqrt{\overline{Q}\,\overline{a}}\left(\sum_{i=1}^{\infty} a_i c_i \right) \geqslant \frac{\underline{Q}}{2}\left(\frac{\varrho^{\frac{\beta-\gamma}{\beta-1}}}{M_\beta^{\frac{1-\gamma}{\beta-1}}} - \delta \right) - 2\sqrt{\delta}\sqrt{\overline{Q}}\varrho,$$

from which the result follows. □

We are finally ready to complete the proof of Theorem 1.1:

Proof of part (iii) of Theorem 1.1. This follows immediately from Propositions 3.6 and 3.8. □

Now that we have our general functional inequality at hand, one may wonder about the sharpness of this method of using the log-Sobolev inequality. Perhaps we were too coarse in our estimation, and Cercignani's conjecture *is* valid in the case $a_i = i^{\gamma}$ with $\gamma < 1$ under the restrictions of Theorem 1.1. The answer, surprisingly, is that the result is optimal, as we shall see in the next subsection.

Optimality of the results. This subsection is devoted to showing that unlike the case $a_i = i$, the case $a_i = i^{\gamma}$ when $\gamma < 1$ does not satisfy Cercignani's conjecture, even if c_1 is bounded appropriately. This is stated in Theorem 1.2.

Proof of Theorem 1.2. We start by choosing $a_i = i^{\gamma}$, $\gamma < 1$ and $Q_i = e^{-\lambda(i-1)}$ $(i \geqslant 1)$ for some $\lambda \geqslant 0$. We will show the desired result by constructing a family of nonnegative sequences $\{\boldsymbol{c}^{(\varepsilon)}\}_{\varepsilon>0}$ with a fixed mass ϱ such that

$$\lim_{\varepsilon\to 0} \frac{D(\boldsymbol{c}^{(\varepsilon)})}{H(\boldsymbol{c}^{(\varepsilon)}|\mathcal{Q})} = 0.$$

Let $\xi > 0$ be such that

$$\frac{\varrho}{2} = \sum_{i=1}^{\infty} i e^{\lambda} e^{-\xi i} = \frac{e^{\lambda-\xi}}{(1-e^{-\xi})^2}.$$

Consider the sequence $\boldsymbol{c}^{(\varepsilon)} = \{c_i^{(\varepsilon)}\}$ given by

$$c_i^{(\varepsilon)} = e^{\lambda} e^{-\xi i} + A_\varepsilon e^{-\varepsilon i}, \quad i \in \mathbb{N},$$

where $0 < \varepsilon$ is small and A_ε is chosen such that the mass of the sequence $\boldsymbol{c}^{(\varepsilon)}$ is ϱ, i.e., $A_\varepsilon = \frac{1}{2}\varrho e^{\varepsilon}(1-e^{-\varepsilon})^2$. Next, as $Q_i/Q_{i+1} = e^{\lambda}$ for any $i \geqslant 1$, we see that

$$\begin{aligned}\frac{Q_i}{Q_{i+1}} c_{i+1}^{(\varepsilon)} - c_1^{(\varepsilon)} c_i^{(\varepsilon)} &= e^{2\lambda} e^{-\xi(i+1)} + A_\varepsilon e^{\lambda} e^{-\varepsilon(i+1)} - e^{2\lambda} e^{-\xi(i+1)} - A_\varepsilon e^{\lambda}(e^{-\xi i-\varepsilon} + e^{-\varepsilon i-\xi}) - A_\varepsilon^2 e^{-\varepsilon(i+1)} \\ &= A_\varepsilon e^{\lambda} e^{-\varepsilon(i+1)}(1 - e^{-(\xi-\varepsilon)} - e^{-(\xi-\varepsilon)i} - A_\varepsilon e^{-\lambda}) > 0\end{aligned}$$

for ε small enough depending only on λ, ξ and ϱ but not on i. Additionally, one can easily verify that

$$\frac{Q_i c_{i+1}^{(\varepsilon)}}{Q_{i+1} c_1^{(\varepsilon)} c_i^{(\varepsilon)}} \leqslant e^{\lambda}\left(1+\frac{1}{A_\varepsilon}\right).$$

As such, setting $B_{z,\gamma} = \sum_{i=1}^{\infty} i^{\gamma} e^{-zi}$ for any $z > 0$, we find that

$$\begin{aligned} D(\boldsymbol{c}^{(\varepsilon)}) &= \sum_{i=1}^{\infty} i^{\gamma}\left(\frac{Q_i}{Q_{i+1}} c_{i+1}^{(\varepsilon)} - c_i^{(\varepsilon)}\right)\log\left(\frac{Q_i c_{i+1}^{(\varepsilon)}}{Q_{i+1} c_1^{(\varepsilon)} c_i^{(\varepsilon)}}\right) \\ &\leqslant A_\varepsilon e^{\lambda} B_{\varepsilon,\gamma} \log\left(e^{\lambda}\left(1+\frac{1}{A_\varepsilon}\right)\right)\left((1-A_\varepsilon e^{-\lambda})e^{-\varepsilon}-e^{-\xi}\right) - A_\varepsilon e^{\lambda} B_{\xi,\gamma}\log\left(e^{\lambda-\varepsilon}\left(1+\frac{1}{A_\varepsilon}\right)\right). \quad (3\text{-}21) \end{aligned}$$

As $A_\varepsilon \approx \frac{1}{2}\varrho\varepsilon^2$ when ε approaches zero, and $B_{\varepsilon,\gamma}$ is of order $\varepsilon^{-(1+\gamma)}$ (see Lemma B.3 in Appendix B) we conclude that

$$\lim_{\varepsilon\to 0} D(\boldsymbol{c}^{(\varepsilon)}) = 0.$$

Lastly, we turn our attention to the relative free energy. We start by denoting by $\bar{\xi} > 0$ the unique parameter for which

$$\varrho = e^{\lambda} \sum_{i=1}^{\infty} i e^{-\bar{\xi} i}.$$

Clearly, $\bar{\xi} < \xi$ and the associated equilibrium with mass ϱ is $\mathcal{Q}_i = e^{\lambda} e^{-\bar{\xi} i}$. Since, for any fixed $i \geqslant 1$, it holds that

$$\lim_{\varepsilon\to 0} c_i^{(\varepsilon)} = c_i^{(0)} = e^{\lambda} e^{-\xi i},$$

using Fatou's lemma we can conclude that

$$\liminf_{\varepsilon\to 0} H(c^{(\varepsilon)}|\mathcal{Q}) \geqslant H(\boldsymbol{c}^{(0)}|\mathcal{Q}) > 0,$$

as $\boldsymbol{c}^{(0)} \neq \mathcal{Q}$. □

Remark 3.9. We notice the following:

- In the example we provided, $z_s = e^{\lambda} < +\infty$ but $\varrho_s = +\infty$. This, however, is not a great obstacle as all our proofs rely on some *positive* distance from z_s and ϱ_s, and can be reformulated accordingly.
- The constructed sequence $\boldsymbol{c}^{(\varepsilon)}$ satisfies

  $$\sup_{\varepsilon} \sum_{i=1}^{\infty} i^{\beta} c_i^{(\varepsilon)} = +\infty$$

  for any $\beta > 1$. Thus, the conclusion of part (ii) of Theorem 1.1 does not apply to it. Actually, one can easily check that

  $$\lim_{\varepsilon\to 0} \frac{D(\boldsymbol{c}^{(\varepsilon)})}{(H(\boldsymbol{c}^{(\varepsilon)}|\mathcal{Q}))^{s}} = 0$$

  for any $s > 0$.

Inequalities with exponential moments. Up to now, we have avoided using exponential moments in any of our functional inequalities. In this section we will show that when $0 \leqslant \gamma < 1$, under the additional assumption of a bounded exponential moment, one can obtain an improved functional inequality between $\overline{D}(\boldsymbol{c})$ and $H(\boldsymbol{c}|\mathcal{Q})$, extending the result given by Jabin and Niethammer [2003]. The key idea in this section is to avoid using the interpolation inequality (3-17) and replace it with one that involves an exponential weight.

Proposition 3.10. *Let $\boldsymbol{f}$ be a nonnegative sequence and let $0 \leqslant \gamma < 1$. Assume that there exists $\mu \in (0, 4\log 2)$ such that*

$$\sum_{i=1}^{\infty} e^{\mu i} f_i = M_\mu^{\exp}(\boldsymbol{f}) < +\infty.$$

Then,

$$M_\gamma(\boldsymbol{f}) \geqslant \frac{M_1(\boldsymbol{f})}{2}\left(\frac{2}{\mu}\log\left(\frac{4M_\mu^{\exp}(\boldsymbol{f})}{\mu e M_1(\boldsymbol{f})}\right)\right)^{-(1-\gamma)}, \tag{3-22}$$

where $M_\alpha(\boldsymbol{f})$ denotes the α-moment of $\boldsymbol{f}$ and $M_\mu^{\exp}(\boldsymbol{f})$ is the exponential moment defined in (1-14).

Proof. For simplicity, we will use the notation of M_1 and $M_\mu^{\exp}$ instead of $M_1(\boldsymbol{f})$ and $M_\mu^{\exp}(\boldsymbol{f})$. We start with the simple inequality

$$\begin{aligned} M_1 = \sum_{i=1}^{\infty} i f_i &= \sum_{i=1}^{N} i^{1-\gamma} i^{\gamma} f_i + \sum_{i=N+1}^{\infty} i e^{-\frac{\mu i}{2}} e^{-\frac{\mu i}{2}} e^{\mu i} f_i \\ &\leqslant N^{1-\gamma} M_\gamma + \frac{2e^{-\frac{\mu(N+1)}{2}}}{\mu e} M_\mu^{\exp} \quad \forall N \in \mathbb{N}, \end{aligned} \tag{3-23}$$

where we used the fact that $\sup_{x \geqslant 0} x e^{-\lambda x} = 1/(\lambda e)$ for any $\lambda > 0$. Our goal will be to choose a particular N to plug in the inequality above to conclude the desired result. Again, using the supremum of $g(x) = xe^{-\lambda x}$, we conclude that

$$M_1 \leqslant \frac{1}{\mu\, e} M_\mu^{\exp}.$$

As $\mu < 4\log 2$, we find that

$$M_1 < \frac{4M_\mu^{\exp}}{\mu e^{1+\frac{\mu}{2}}},$$

from which we conclude that

$$N = \left[\frac{2}{\mu}\log\left(\frac{4M_\mu^{\exp}}{\mu e M_1}\right)\right] \geqslant 1.$$

Plugging this N into (3-23) we see that

$$e^{-\frac{\mu(N+1)}{2}} \leqslant \frac{\mu e M_1}{4M_\mu^{\exp}},$$

and as such

$$M_\gamma \geqslant N^{\gamma-1}\frac{M_1}{2}$$

and the result follows. ∎

With this proposition at hand, we are prepared to show part (i) of Theorem 1.4.

Proof of part (i) of Theorem 1.4. Without loss of generality we may assume that $\mu \in (0, 4\log 2)$. Introduce the sequence $\boldsymbol{f} = \{f_i\}$, where

$$f_i = Q_i\left(\sqrt{\frac{c_1 c_i}{Q_i}} - \sqrt{\frac{c_{i+1}}{Q_{i+1}}}\right)^2, \quad i \geqslant 1.$$

Following the same proof as presented in Lemma 3.5, we find that

$$M_\mu^{\exp}(\boldsymbol{f}) \leqslant 2\left(c_1 + z_s \sup_j \frac{\alpha_j}{\alpha_{j+1}}\right) M_\mu^{\exp}(\boldsymbol{c}).$$

Thus, using the simple fact that $M_\alpha(\boldsymbol{f}) = \overline{D}_\alpha(\boldsymbol{c})$, for any $\alpha > 0$, together with Proposition 3.10 and parts (i) and (iii) of Theorem 1.1, yields the desired functional inequality. $\square$

4. Rate of convergence to equilibrium

In this section we will use all the information we gathered so far to prove Theorem 1.3 and part (ii) of Theorem 1.4, giving an explicit rate of convergence to equilibrium for the Becker–Döring equations.

The convergence result in Theorem 1.3 is a consequence of Theorem 1.1. To use the functional inequality established there, we need first to invoke uniform (and explicit) upper bounds on moments $M_\beta(\boldsymbol{c}(t))$; see (1-18). This is provided by the following (see [Cañizo et al. 2017]):

Proposition 4.1. *Let $\{a_i\}_{i\in\mathbb{N}}$ and $\{Q_i\}_{i\in\mathbb{N}}$ satisfy Hypotheses 1–3 with $0 \leqslant \gamma \leqslant 1$, and let $\boldsymbol{c}(t) = \{c_i(t)\}_{i\in\mathbb{N}}$ be a solution to the Becker–Döring equations with mass $\varrho \in (0, \varrho_s)$. Let $\beta \geqslant \max\{2-\gamma, 1+\gamma\}$ be such that*

$$M_\beta(\boldsymbol{c}(0)) = \sum_{i=1}^{\infty} i^\beta c_i(0) < \infty.$$

There exists a constant $C > 0$ depending only on β, $M_\beta(0)$, the initial relative free energy $H(\boldsymbol{c}(0)|\mathcal{Q})$, the coefficients $\{a_i\}_{i\geqslant 1}$, $\{b_i\}_{i\geqslant 1}$ and the mass ϱ such that

$$M_\beta(\boldsymbol{c}(t)) = \sum_{i=1}^{\infty} i^\beta c_i(t) \leqslant C \quad \textit{for all } t \geqslant 0.$$

Using such an estimate, the proof is easily derived from Theorem 1.1 and part (i) of Theorem 1.4, yet we provide a proof here for the sake of completeness and to show that we can find all the constants explicitly.

Proof of Theorem 1.3. Combining Theorem 1.1 and Proposition 4.1, we conclude the differential inequality

$$\frac{\mathrm{d}}{\mathrm{d}t} H(\boldsymbol{c}(t)|\mathcal{Q}) \leqslant \begin{cases} -\min\bigl(K H(\boldsymbol{c}(t)|\mathcal{Q}), \varepsilon\bigr), & \gamma = 1, \\ -\min\bigl(K H(\boldsymbol{c}(t)|\mathcal{Q})^{\frac{\beta-\gamma}{\beta-1}}, \varepsilon\bigr), & 0 \leqslant \gamma < 1 \end{cases} \tag{4-1}$$

for appropriate K and ε. We claim that there exists $t_0 \geqslant 0$ such that for all $t \geqslant t_0$,

$$H(c(t)|\mathcal{Q}) \leqslant \begin{cases} \frac{\varepsilon}{K}, & \gamma = 1, \\ \left(\frac{\varepsilon}{K}\right)^{\frac{\beta-1}{\beta-\gamma}}, & 0 \leqslant \gamma < 1. \end{cases} \tag{4-2}$$

Indeed, if $H(\boldsymbol{c}(t))|\mathcal{Q})$ is larger than the appropriate constants in $[0, t]$ then

$$\frac{\mathrm{d}}{\mathrm{d}s} H(\boldsymbol{c}(s)|\mathcal{Q}) \leqslant -\varepsilon \quad \forall s \in (0, t),$$

implying that

$$H(\boldsymbol{c}(t)|\mathcal{Q}) \leqslant H(\boldsymbol{c}(0)|\mathcal{Q}) - \varepsilon t.$$

We define

$$t_0 = \begin{cases} \min\left(0, \left(H(\boldsymbol{c}(0)|\mathcal{Q}) - \frac{\varepsilon}{K}\right)/\varepsilon\right), & \gamma = 1, \\ \min\left(0, \left(H(\boldsymbol{c}(0)|\mathcal{Q}) - \left(\frac{\varepsilon}{K}\right)^{\frac{\beta-1}{\beta-\gamma}}\right)/\varepsilon\right), & 0 \leqslant \gamma < 1, \end{cases}$$

and find that $H(\boldsymbol{c}(t_0)|\mathcal{Q})$ satisfies the appropriate inequality in (4-2). As $H(\boldsymbol{c}(t)|\mathcal{Q})$ is decreasing, we conclude that (4-2) is valid for any $t \geqslant t_0$.

With this in hand, along with (4-1), we have, for all $t \geqslant t_0$,

$$H(\boldsymbol{c}(t)|\mathcal{Q}) \leqslant \begin{cases} H(\boldsymbol{c}(t_0)|\mathcal{Q}) e^{-K(t-t_0)}, & \gamma = 1, \\ \left(H(\boldsymbol{c}(t_0)|\mathcal{Q})^{\gamma-1/\beta-1} + \frac{1-\gamma}{\beta-1} K(t-t_0)\right)^{-\frac{\beta-1}{1-\gamma}}, & 0 \leqslant \gamma < 1. \end{cases}$$

As

$$H(\boldsymbol{c}(t_0)|\mathcal{Q}) = \begin{cases} \min\left(H(\boldsymbol{c}(0)|\mathcal{Q}), \frac{\varepsilon}{K}\right), & \gamma = 1, \\ \min\left(H(\boldsymbol{c}(0)|\mathcal{Q}), \left(\frac{\varepsilon}{K}\right)^{\frac{\beta-1}{\beta-\gamma}}\right), & 0 \leqslant \gamma < 1, \end{cases}$$

and t_0 is given explicitly, we conclude that

$$C(H(\boldsymbol{c}(0)|\mathcal{Q})) = \begin{cases} H(\boldsymbol{c}(0)|\mathcal{Q}), & \gamma = 1, t_0 = 0, \\ \frac{\varepsilon}{K} e^{K\frac{1}{\varepsilon}\left(H(\boldsymbol{c}(0)|\mathcal{Q}) - \frac{\varepsilon}{K}\right)}, & \gamma = 1, t_0 > 0, \\ H(\boldsymbol{c}(0)|\mathcal{Q}), & 0 \leqslant \gamma < 1, t_0 = 0, \\ \left(\frac{\varepsilon}{K}\right)^{\frac{\gamma-1}{\beta-\gamma}} - \frac{1-\gamma}{\beta-1} K \frac{1}{\varepsilon}\left(H(\boldsymbol{c}(0)|\mathcal{Q}) - \left(\frac{\varepsilon}{K}\right)^{\frac{\beta-1}{\beta-\gamma}}\right), & 0 \leqslant \gamma < 1, t_0 > 0, \end{cases}$$

completing the proof. □

Proof of part (ii) of Theorem 1.4. This follows form part (i) of Theorem 1.4 by the same methods used in the above proof and the fact that

$$\sup_{t \geqslant 0} M_{\mu'}^{\exp}(\boldsymbol{c}(t)) < +\infty$$

for some $0 < \mu' < \mu$ (a known result from [Jabin and Niethammer 2003]). □

5. Consequences for general coagulation and fragmentation models

In this final section we illustrate how the functional inequalities investigated in Section 3 provide new insights on the behaviour of solutions to general discrete coagulation-fragmentation models.

General discrete coagulation-fragmentation equation. The Becker–Döring equations (1-1) are derived under the assumption that the only relevant reactions taking place are those between monomers and clusters of any size. One can obtain a more general model by taking into account reactions between clusters of any size. Keeping the notation of the Introduction, this means that we consider reactions of the type

$$\{i\} + \{j\} \rightleftharpoons \{i + j\}$$

for any positive integer sizes i and j. We assume their coagulation rate (i.e., the reaction from left to right) is determined by a coefficient we call $a_{i,j}$, and their fragmentation rate (the reaction from right to left) by a coefficient called $b_{i,j}$. These coefficients are always assumed to be nonnegative (as before) and symmetric in i, j (that is, $a_{i,j} = a_{j,i}$ and $b_{i,j} = b_{j,i}$ for all i, j). The equation corresponding to (1-1) is then

$$\frac{\mathrm{d}}{\mathrm{d}t} c_i(t) = \frac{1}{2} \sum_{j=1}^{i-1} W_{j,i-j}(t) - \sum_{j=1}^{\infty} W_{i,j}(t), \quad i \in \mathbb{N}, \tag{5-1}$$

where

$$W_{i,j}(t) := a_{i,j}\, c_i(t) c_j(t) - b_{i,j}\, c_{i+j}(t), \quad i \in \mathbb{N}. \tag{5-2}$$

The system (1-1) is then a particular case of (5-1) obtained by choosing $a_{i,j}$, $b_{i,j}$ as

$$a_{i,j} = b_{i,j} = 0 \quad \text{when } \min\{i, j\} \geqslant 2, \tag{5-3}$$

$$a_{1,1} := 2a_1, \qquad a_{i,1} = a_{1,i} = a_i \qquad \text{for } i \geqslant 2, \tag{5-4}$$

$$b_{1,1} := 2b_2, \qquad b_{i,1} = b_{1,i} = b_{i+1} \quad \text{for } i \geqslant 2. \tag{5-5}$$

The mathematical theory of this full system is much less complete than that of (1-1). Well-posedness of mass-conserving solutions has been studied in [Ball and Carr 1990], and there are a number of works on asymptotic behaviour, for instance [Cañizo 2005; 2007; Carr and da Costa 1994; Carr 1992], but it is still not fully understood. To start with, it is unclear whether equilibria of (5-1) are unique or not (when they exist). A common physical condition imposed on the coefficients $a_{i,j}$, $b_{i,j}$ which avoids this problem is that of *detailed balance*: we say it holds when there exists a sequence $\{Q_i\}_{i \geqslant 1}$ of strictly positive numbers such that

$$a_{i,j} Q_i Q_j = b_{i,j} Q_{i+j} \quad \text{for any } i, j, \tag{5-6}$$

where we always further assume without loss of generality that $Q_1 = 1$. This is the analogue of (1-4), but in this case it needs to be imposed as a condition since numbers Q_i satisfying (5-6) cannot always be found (unlike in the Becker–Döring case). If we assume (5-6) then equilibria (5-1) exist and have the same form (1-5) as in the Becker–Döring case, and a similar phase transition in the long-time behaviour has been rigorously proved in some cases (see [Cañizo 2005; 2007; Carr and da Costa 1994; Carr 1992] for more details). However, even with detailed balance, the long-time behaviour is in general not understood except in particular cases. If clusters larger than a given size N do not react among themselves (that is, if $a_{i,j} = b_{i,j} = 0$ whenever $\min\{i, j\} > N$) the system is known as the *generalised Becker–Döring system*, and has been studied in [Cañizo 2005; da Costa 1998]. For coefficients $a_{i,j}$ given by

$$a_{i,j} = i^{\gamma} j^{\eta} + i^{\eta} j^{\gamma} \quad \text{for any } i, j, \tag{5-7}$$

with $\eta \leqslant 0 \leqslant \gamma$ and $\gamma+\eta \leqslant 1$, the asymptotic behaviour was identified in [Cañizo 2007] and a constructive (though probably far from optimal) rate of convergence to equilibrium was given. Very little is known about the asymptotic behaviour for coefficients of the type (5-7) with $\gamma, \eta > 0$ and $\gamma+\eta \leqslant 1$. In this case the size of $a_{i,i}$ is larger than that of $a_{i,1}$ and the system (5-1) may behave quite differently from (1-1).

A natural question is whether any of the functional inequalities investigated in this paper can shed new light on the behaviour of solutions to (5-1). Assuming the detailed balance condition (5-6), along a solution $\boldsymbol{c}(t) = \{c_i(t)\}_{i\geqslant 1}$ to (5-1) we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} H(\boldsymbol{c}(t)) &= -D_{\mathrm{CF}}(\boldsymbol{c}(t)) \\ &:= -\frac{1}{2}\sum_{i,j=1}^{\infty} a_{i,j} Q_i Q_j \left(\frac{c_i c_j}{Q_i Q_j} - \frac{c_{i+j}}{Q_{i+j}}\right)\left(\log\frac{c_i c_j}{Q_i Q_j} - \log\frac{c_{i+j}}{Q_{i+j}}\right) \\ &\leqslant -\sum_{i=1}^{\infty} a_i Q_i \left(\frac{c_i c_1}{Q_i} - \frac{c_{i+1}}{Q_{i+1}}\right)\left(\log\frac{c_i c_1}{Q_i} - \log\frac{c_{i+1}}{Q_{i+1}}\right) = D(\boldsymbol{c}(t)) \leqslant 0 \end{aligned} \tag{5-8}$$

(see [Cañizo 2007] for a rigorous proof), where the a_i are defined by (5-3) for any $i \geqslant 1$. Hence the free energy is also a Lyapunov functional for (5-1), and it dissipates at a *faster* rate than for the Becker–Döring equations (since more types of reactions are allowed). As such, it is reasonable to think that the inequalities from Section 3 can be useful also in this case. This turns out to be true, and some improvements can be made on existing results. However, it also turns out that our results are not able to extend the range of possible coefficients for which convergence to a particular subcritical equilibrium can be proved; we cannot give any new results for coefficients such as (5-7) with $\gamma, \eta > 0$ and $\gamma+\eta \leqslant 1$.

Proof of Theorem 1.5. We now give the proof of our main result concerning the above model (5-1). One of the main obstacles in applying directly our results to equation (5-1) is that, unlike for the Becker–Döring equations, the moments of solutions to the general coagulation and fragmentation system are not known to be bounded; i.e., Proposition 4.1 is not available for (5-1). One can for example say the following about integer moments (this result can easily be extended to noninteger powers by interpolation, and was known from the early works in the topic [Carr and da Costa 1994; Carr 1992]). From this point onward we will assume that

$$a_{i,j} = i^\gamma j^\eta + i^\eta j^\gamma \quad \text{for } i, j \in \mathbb{N}, \tag{5-9}$$

with $\eta \leqslant \gamma$ and $0 \leqslant \lambda := \gamma+\eta \leqslant 1$.

Lemma 5.1. *Let $k \in \mathbb{N}$ and let $\boldsymbol{c} = \boldsymbol{c}(t) = \{c_i(t)\}_{i\in\mathbb{N}}$ be a solution with mass ϱ to the coagulation and fragmentation system (5-1) with coefficients satisfying (5-9). Then*

$$M_k(\boldsymbol{c}(t)) \leqslant \begin{cases} \left(M_k(\boldsymbol{c}(0)) + \frac{1-\lambda}{k-1}(2^k-2)\varrho^{\frac{1-\gamma}{k-1}} t\right)^{\frac{k-1}{1-\lambda}} & \text{if } 0<\lambda<1, \\ M_k(\boldsymbol{c}(0)) \exp(2(2^k-2)\varrho t) & \text{if } \lambda = 1, \end{cases} \tag{5-10}$$

where $M_p(\boldsymbol{c}(t)) := \sum_{i=1}^{\infty} i^p c_i(t)$ for any $p \geqslant 0$, $t \geqslant 0$.

Proof. We give a formal proof for completeness; a rigorous one can be obtained by standard approximation methods, and can be found in [Ball and Carr 1990]. To simplify the notation and since $\boldsymbol{c}(t)$ is fixed, we define $M_j(t) = M_j(\boldsymbol{c}(t))$ for any $j \geqslant 1$, $t \geqslant 0$. One can check the following weak formula for the integral of the right-hand side of (5-1) against a test sequence $\{\phi(i)\}_i$:

$$\sum_{i=1}^{\infty} \phi(i)\Big(\frac{1}{2}\sum_{j=1}^{i-1} W_{i-j,j} - \sum_{j=1}^{\infty} W_{i,j}\Big) = \frac{1}{2}\sum_{i,j}\big(\phi(i+j)-\phi(i)-\phi(j)\big)W_{i,j}.$$

Applying this to $\phi(i) := i^k$, neglecting the negative contribution of the fragmentation terms and using the binomial formula, one obtains

$$\frac{\mathrm{d}}{\mathrm{d}t} M_k(t) \leqslant \sum_{l=1}^{k-1}\binom{k}{l} M_{l+\gamma}(t) M_{k-l+\eta}(t) \quad \forall t \geqslant 0.$$

Next, we use the interpolation

$$M_\delta(t) \leqslant M_1^{\frac{k-\delta}{k-1}}(t) M_k^{\frac{\delta-1}{k-1}}(t),$$

where $1 < \delta < k$, to find that

$$M_{l+\gamma}(t) M_{k-l+\eta}(t) \leqslant M_1(t)^{\frac{k-\lambda}{k-1}} M_k(t)^{\frac{k+\lambda-2}{k-1}}.$$

Thus,

$$\frac{\mathrm{d}}{\mathrm{d}t} M_k(t) \leqslant (2^k-2)\varrho^{\frac{k-\lambda}{k-1}} M_k^{\frac{k+\lambda-2}{k-1}}(t) \quad \forall t \geqslant 0$$

and the result follows from this differential inequality. □

With the above at hand, we are now able to prove our main result about the rate of convergence to equilibrium in the general setting of coagulation and fragmentation equations:

Proof of Theorem 1.5. Assume for the moment that $a_{i,j}$ is of the form (5-7), in order to see why the proof only works for coefficients of the form (1-27).

Fix $\delta > 0$ such that $0 < \delta < \bar{z} < z_{\mathrm{s}} - \delta$. We use the observation (5-8) that $D_{\mathrm{CF}}(\boldsymbol{c}(t)) \geqslant D(\boldsymbol{c}(t))$ at all times $t \geqslant 0$, defining $\{a_i\}_{i\in\mathbb{N}}$ by (5-3). Using Theorem 1.1 (actually, its more detailed forms in equation (3-18) and Proposition 3.8), we obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} H(\boldsymbol{c}(t)|\mathcal{Q}) &= -D_{\mathrm{CF}}(\boldsymbol{c}(t)) \leqslant -D(\boldsymbol{c}(t)) \\
&\leqslant \begin{cases} -C M_k(\boldsymbol{c}(t))^{\frac{\gamma-1}{k-1}} H(\boldsymbol{c}(t)|\mathcal{Q})^{\frac{k-\gamma}{k-1}} & \text{if } \delta < c_1(t) < z_{\mathrm{s}} - \delta, \\ -C M_k(\boldsymbol{c}(t))^{\frac{\gamma-1}{k-1}} & \text{if } c_1(t) < \delta \text{ or } c_1(t) \geqslant z_{\mathrm{s}} - \delta \end{cases} \\
&\leqslant -C_0 M_k(\boldsymbol{c}(t))^{\frac{\gamma-1}{k-1}} H(\boldsymbol{c}(t)|\mathcal{Q})^{\frac{k-\gamma}{k-1}}
\end{aligned}$$

for some constant $C_0 > 0$ that depends also on $H(\boldsymbol{c}(0)|\mathcal{Q})$. Using Lemma 5.1 this implies

$$\frac{\mathrm{d}}{\mathrm{d}t} H(\boldsymbol{c}(t)|\mathcal{Q}) \leqslant -\frac{C_0}{\big(M_k(\boldsymbol{c}(0)) + \frac{1-\lambda}{k-1}(2^k-2)\varrho^{\frac{k-\lambda}{k-1}} t\big)^{\frac{1-\gamma}{1-\lambda}}} H(\boldsymbol{c}(t)|\mathcal{Q})^{\frac{k-\gamma}{k-1}}, \quad t \geqslant 0.$$

This implies decay of $H(\boldsymbol{c}(t))$ only when $\lambda = \gamma$, that is, when $\eta = 0$ (since $\lambda = \gamma + \eta$). Solving the differential inequality yields the result. □

Remark 5.2. The same decay rate was obtained in [Cañizo 2007] by means of the particular case of inequality (1-19) for $\beta = 2 - \gamma$. Here we obtain slightly different decay rates by assuming higher moments of the initial data $\boldsymbol{c}(0)$ are finite, but the method does not seem to give a better decay than a power of $\log t$ in any case.

Remark 5.3. It seems to the authors that the inequality we use in the proof of Theorem 1.5 is not optimal, and could be improved to deal with the case

$$a_{i,j} = i^\gamma j^\eta + i^\eta j^\gamma,$$

with a resulting convergence rate that would depend on $\lambda = \gamma + \eta$.

Appendix A: Additional computations for the theory of the discrete log-Sobolev inequality with weights

We have collected here technical lemmas from Section 2 that we felt would have encumbered it.

Lemma A.1. *For any sequence $\boldsymbol{f}$, we have*

$$\mathrm{Ent}_\mu(\boldsymbol{f}^2) \leqslant \mathcal{L}(\boldsymbol{f}) \leqslant \mathrm{Ent}_\mu(\boldsymbol{f}^2) + 2\sum_{i=1}^\infty \mu_i f_i^2.$$

Proof. From the definition of $\mathcal{L}$, the inequality

$$\mathrm{Ent}_\mu(\boldsymbol{f}^2) \leqslant \mathcal{L}(\boldsymbol{f})$$

is trivial. We thus consider the right-hand side inequality. For a given sequence $\boldsymbol{f}$ and any $\alpha \in \mathbb{R}$ we define

$$\begin{aligned} G_\alpha(t) &= \sum_{i=1}^\infty \mu_i (t f_i + \alpha)^2 \log\left(\frac{(t f_i + \alpha)^2}{\sum_{i=1}^\infty \mu_i (t f_i + \alpha)^2}\right) \\ &= 2\sum_{i=1}^\infty \mu_i (t f_i + \alpha)^2 \log|t f_i + \alpha| - \left(\sum_{i=1}^\infty \mu_i (t f_i + \alpha)^2\right) \log\left(\sum_{i=1}^\infty \mu_i (t f_i + \alpha)^2\right), \end{aligned}$$

and notice that

$$G_0(t) = t^2\,\mathrm{Ent}_\mu(\boldsymbol{f}^2).$$

Next, we define $g(t) = G_0(t) + 2t^2 \sum_{i=1}^\infty \mu_i f_i^2$ and notice that the inequality we want to prove is equivalent to

$$G_\alpha(1) \leqslant g(1)$$

for any $\alpha \in \mathbb{R}$. Clearly $G_\alpha(t) \leqslant g(t)$ when $t = 0$. Differentiating G we find that

$$\begin{aligned} G'_\alpha(t) &= 4\sum_{i=1}^{\infty} \mu_i f_i |tf_i+\alpha| \log(tf_i+\alpha) + 2\sum_{i=1}^{\infty} \mu_i f_i (tf_i+\alpha) \\ &\qquad - 2\Big(\sum_{i=1}^{\infty} \mu_i f_i (tf_i+\alpha)\Big)\log\Big(\sum_{i=1}^{\infty}\mu_i (tf_i+\alpha)^2\Big) - 2\sum_{i=1}^{\infty}\mu_i f_i(tf_i+\alpha) \\ &= 4\sum_{i=1}^{\infty}\mu_i f_i (tf_i+\alpha)\log|tf_i+\alpha| - 2\Big(\sum_{i=1}^{\infty}\mu_i f_i(tf_i+\alpha)\Big)\log\Big(\sum_{i=1}^{\infty}\mu_i(tf_i+\alpha)^2\Big), \end{aligned}$$

which satisfies $G'_\alpha(0) = 0$ for any $\boldsymbol{f}$ and α, implying that $G'_\alpha(0) = g'(0) = 0$. As G is defined for any $t \in [0, 1]$ we see that it is enough to show that when defined,

$$G''_\alpha(t) \leqslant g''(t)$$

for any α. Indeed,

$$\begin{aligned} G''_\alpha(t) &= 4\sum_{i=1}^{\infty}\mu_i f_i^2 \log|tf_i+\alpha| + 4\sum_{i=1}^{\infty}\mu_i f_i^2 - 2\sum_{i=1}^{\infty}\mu_i f_i^2 \log\Big(\sum_{i=1}^{\infty}\mu_i(tf_i+\alpha)^2\Big) \\ &\qquad - 4\frac{\big(\sum_{i=1}^{\infty}\mu_i f_i(tf_i+\alpha)\big)^2}{\sum_{i=1}^{\infty}\mu_i(tf_i+\alpha)^2} \\ &= 2\sum_{i=1}^{\infty}\mu_i f_i^2\log\Big(\frac{(tf_i+\alpha)^2}{\sum_{i=1}^{\infty}\mu_i(tf_i+\alpha)^2}\Big) + 4\sum_{i=1}^{\infty}\mu_i f_i^2 - 4\frac{\big(\sum_{i=1}^{\infty}\mu_i f_i(tf_i+\alpha)\big)^2}{\sum_{i=1}^{\infty}\mu_i(tf_i+\alpha)^2}. \end{aligned}$$

As

$$\mathrm{Ent}_\mu(\boldsymbol{f}^2) = \sup\Big\{\sum_{i=1}^{\infty}\mu_i f_i^2 \log h_i : \sum_{i=1}^{\infty}\mu_i h_i = 1\Big\},$$

we see that by choosing

$$h_i = \frac{(tf_i+\alpha)^2}{\sum_{i=1}^{\infty}\mu_i(tf_i+\alpha)^2}$$

we get

$$G''_\alpha(t) \leqslant 2\,\mathrm{Ent}_\mu(\boldsymbol{f}^2) + 4\sum_{i=1}^{\infty}\mu_i f_i^2 = g''(t). \qquad \square$$

Lemma A.2. *For all $\boldsymbol{f} \in L_\Phi$, we have*

$$\|\boldsymbol{f}\|_{L^1_\mu} \leqslant \|\boldsymbol{f}\|_{L^2_\mu} \leqslant \sqrt{\tfrac{3}{2}}\|\boldsymbol{f}\|_{L_\Phi}. \tag{A-1}$$

Proof. The inequality

$$\|\boldsymbol{f}\|_{L^1_\mu} \leqslant \|\boldsymbol{f}\|_{L^2_\mu}$$

is immediate as $\boldsymbol{\mu}$ is a probability measure. To show the last inequality we may assume that $\|\boldsymbol{f}\|_{L_\Phi} = 1$. Due to Fatou's lemma we know that if $k_n \xrightarrow{n\to\infty} k > 0$ then

$$\sum_{i=1}^{\infty} \mu_i \Phi\left(\frac{|f_i|}{k}\right) \leqslant \liminf_{n\to\infty} \sum_{i=1}^{\infty} \mu_i \Phi\left(\frac{|f_i|}{k_n}\right),$$

implying that if $\|f\|_{L_\Phi} > 0$ then

$$\sum_{i=1}^{\infty} \mu_i \Phi\left(\frac{|f_i|}{\|f\|_{L_\Phi}}\right) \leqslant 1.$$

In our case, since $\Psi(x)$ is convex we find that

$$1 \geqslant \sum_{i=1}^{\infty} \mu_i \Phi(f_i) = \sum_{i=1}^{\infty} \mu_i \Psi(f_i^2) \geqslant \Psi\left(\sum_{i=1}^{\infty} \mu_i f_i^2\right) = \Psi(\|\boldsymbol{f}\|_{L^2_\mu}^2).$$

As Ψ is increasing and $\Psi\left(\frac{3}{2}\right) > 1$ we conclude that

$$\|\boldsymbol{f}\|_{L^2_\mu}^2 < \tfrac{3}{2},$$

yielding the desired result. □

Lemma A.3. *Let $\boldsymbol{f} \in L_\Phi$. Then*

$$\|\boldsymbol{f} - \langle \boldsymbol{f}\rangle\|_{L^2_\mu}^2 = \tfrac{1}{2} \lim_{|a|\to\infty} \mathrm{Ent}_\mu((\boldsymbol{f} + a)^2). \tag{A-2}$$

Proof. We start by noticing that

$$\mathrm{Ent}_\mu((\boldsymbol{f} + a)^2) = \sum_{i=1}^{\infty} \mu_i (f_i^2 + 2af_i + a^2) \log\left(\frac{(1 + f_i/a)^2}{\sum_{i=1}^{\infty} \mu_i (1 + f_i/a)^2}\right),$$

and continue by assuming that f_i is uniformly bounded, from which the result will follow with an application of an appropriate convergence theorem. There exists a_0 such that if $|a| > |a_0|$ we have that $|f_i/a| < \frac{1}{2}$ uniformly in i. As on $\left[-\frac{1}{2}, \frac{1}{2}\right]$, we have that there exists $C > 0$ such that

$$\left|\log(1+x) - x + \tfrac{1}{2}x^2\right| \leqslant Cx^3.$$

We conclude that

$$\log\left(1 + 2\frac{f_i}{a} + \frac{f_i^2}{a^2}\right) = \left(2\frac{f_i}{a} + \frac{f_i^2}{a^2}\right) - 2\frac{f_i^2}{a^2} + \frac{E_{1,i}}{a^3} = 2\frac{f_i}{a} - \frac{f_i^2}{a^2} + \frac{E_{1,i}}{a^3}$$

and

$$\log\left(1 + 2\frac{\langle \boldsymbol{f}\rangle}{a} + \frac{\|\boldsymbol{f}\|_{L^2_\mu}^2}{a^2}\right) = 2\frac{\langle \boldsymbol{f}\rangle}{a} + \frac{\|\boldsymbol{f}\|_{L^2_\mu}^2}{a^2} - 2\frac{\langle \boldsymbol{f}\rangle^2}{a^2} + \frac{E_{2,i}}{a^3},$$

where $E_{1,i}$, $E_{2,i}$ are uniformly bounded in i. This implies

$$\operatorname{Ent}_\mu((\boldsymbol{f}+a)^2)=\sum_{i=1}^\infty \mu_i(f_i^2+2af_i+a^2)\left(2\frac{f_i}{a}-2\frac{\langle \boldsymbol{f}\rangle}{a}-\frac{f_i^2}{a^2}-\frac{\|\boldsymbol{f}\|^2_{L^2_\mu}}{a^2}+2\frac{\langle \boldsymbol{f}\rangle^2}{a^2}\right) + \frac{1}{a}\sum_{i=1}^\infty \mu_i\left(1+2\frac{f_i}{a}+\frac{f_i^2}{a^2}\right)(E_{1,i}-E_{2,i}).$$

The last term clearly goes to zero as $|a|$ goes to infinity, so we are only left to deal with the first expression.

$$\begin{aligned}\sum_{i=1}^\infty \mu_i(f_i^2+2af_i+a^2)&\left(2\frac{f_i}{a}-2\frac{\langle \boldsymbol{f}\rangle}{a}-\frac{f_i^2}{a^2}-\frac{\|\boldsymbol{f}\|^2_{L^2_\mu}}{a^2}+2\frac{\langle \boldsymbol{f}\rangle^2}{a^2}\right)\\ &=4\|\boldsymbol{f}\|^2_{L^2_\mu}-4\langle \boldsymbol{f}\rangle^2+2a\langle \boldsymbol{f}\rangle-2a\langle \boldsymbol{f}\rangle-\|\boldsymbol{f}\|^2_{L^2_\mu}-\|\boldsymbol{f}\|^2_{L^2_\mu}+2\langle \boldsymbol{f}\rangle^2+\frac{E_3}{a}\\ &=2(\|\boldsymbol{f}\|^2_{L^2_\mu}-\langle \boldsymbol{f}\rangle^2)+\frac{E_3}{a}.\end{aligned}$$

This completes the proof as $\|\boldsymbol{f}-\langle \boldsymbol{f}\rangle\|^2_{L^2_\mu}=\|\boldsymbol{f}\|^2_{L^2_\mu}-\langle \boldsymbol{f}\rangle^2$. □

Lemma A.4. *Let $\boldsymbol{f}$ be a sequence such that $f_m=0$ for some $m\in\mathbb{N}$. Set by $\boldsymbol{f}^{(0)}=\boldsymbol{f}\mathbb{1}_{i<m}$ and $\boldsymbol{f}^{(1)}=\boldsymbol{f}\mathbb{1}_{i>m}$. Then*

$$\begin{aligned}\|\langle \boldsymbol{f}^{(0)}\rangle\|_{L_\Phi}\leqslant|\langle \boldsymbol{f}^{(0)}\rangle|\leqslant\|\boldsymbol{f}^{(0)}\|_{L^2_\mu}\left(\sum_{i=1}^{m-1}\mu_i\right)^{\frac12},\\ \|\langle \boldsymbol{f}^{(1)}\rangle\|_{L_\Phi}\leqslant|\langle \boldsymbol{f}^{(1)}\rangle|\leqslant\|\boldsymbol{f}^{(1)}\|_{L^2_\mu}\left(\sum_{i=m+1}^{\infty}\mu_i\right)^{\frac12}.\end{aligned}\tag{A-3}$$

Proof. We start by noticing that for any constant sequence $\boldsymbol{f}=\alpha$ one has

$$\|\alpha\|_{L_\Phi}=\inf_{k>0}\left\{\sum_{i=1}^\infty\mu_i\Phi\left(\frac{|\alpha|}{k}\right)\leqslant 1\right\}=\inf_{k>0}\left\{\Phi\left(\frac{|\alpha|}{k}\right)\leqslant 1\right\}=\frac{|\alpha|}{\Phi^{-1}(1)}\leqslant|\alpha|,$$

as long as $\Phi(1)<1$, which is valid in our case. Next we notice that

$$|\langle \boldsymbol{f}^{(0)}\rangle|\leqslant\sum_{i=1}^{m-1}\mu_i|f_i|\leqslant\left(\sum_{i=1}^{m-1}\mu_i f_i^2\right)^{\frac12}\left(\sum_{i=1}^{m-1}\mu_i\right)^{\frac12}=\|\boldsymbol{f}^{(0)}\|_{L^2_\mu}\left(\sum_{i=1}^{m-1}\mu_i\right)^{\frac12}.$$

This yields the first inequality and similar arguments yield the second inequality. □

Remark A.5. As was shown in the proof of Lemma A.4, one can actually improve the bounds in (A-3) by a factor of $\Psi^{-1}(1)$.

Lemma A.6. *For any $t\geqslant\frac32$ one has that*

$$\frac13\frac{t}{\log t}\leqslant\Psi^{-1}(t)\leqslant 2\frac{t}{\log t}.\tag{A-4}$$

Proof. We start by noticing that

$$\Psi\left(\frac{1}{3}\frac{t}{\log t}\right)=\frac{1}{3}\frac{t}{\log t}\log\left(1+\frac{1}{3}\frac{t}{\log t}\right)\leqslant\frac{1}{3}\frac{t}{\log t}\log\left(1+\frac{t}{\log(\frac{27}{8})}\right)\leqslant\frac{1}{3}\frac{t}{\log t}\log(1+t).$$

Thus, one notices that if

$$1+t\leqslant t^3$$

when $t\geqslant\frac{3}{2}$, we have

$$\Psi\left(\frac{1}{3}\frac{t}{\log t}\right)\leqslant t,$$

yielding the left-hand side of (A-4). This is indeed the case as $g(t)=t^3-t-1$ is increasing on $[1/\sqrt{3},\infty)$ and $g\left(\frac{3}{2}\right)>0$.

For the converse we notice that

$$\Psi\left(2\frac{t}{\log t}\right)=2\frac{t}{\log t}\log\left(1+2\frac{t}{\log t}\right)\geqslant t$$

if and only if

$$1+2\frac{t}{\log t}\geqslant\sqrt{t}.$$

Considering the function $g(x)=x/\log x$ for $x>1$, we see that it obtains a minimum at $x=e$. Thus, for any $x>1$, we have $g(x)\geqslant e>1$. We conclude that for $t>\frac{3}{2}$,

$$2\frac{t}{\log t}=\sqrt{t}\,g(\sqrt{t})\geqslant\sqrt{t},$$

showing the desired result. □

Appendix B: Additional useful computations

Lemma B.1. *For given coagulation and detailed balance coefficients $\{a_i\}_{i\in\mathbb{N}}$ and $\{Q_i\}_{i\in\mathbb{N}}$, and a given positive sequence $\boldsymbol{c}$ with finite mass ϱ, we have, for any $z>0$,*

$$H(\boldsymbol{c}|\mathcal{Q})\leqslant H(\boldsymbol{c}|\mathcal{Q}_z),$$

where $\mathcal{Q}=\mathcal{Q}_{\bar{z}}$.

Proof. We have

$$H(\boldsymbol{c}|\mathcal{Q}_z)=\sum_{i=1}^{\infty}c_i\left(\log\left(\frac{c_1}{Q_iz^i}\right)-1\right)+\sum_{i=1}^{\infty}Q_iz^i$$

implying

$$H(\boldsymbol{c}|\mathcal{Q}_{z_1})-H(\boldsymbol{c}|\mathcal{Q}_{z_2})=\sum_{i=1}^{\infty}ic_i\log\left(\frac{z_2}{z_1}\right)+\sum_{i=1}^{\infty}Q_i(z_1^i-z_2^i).$$

In particular, if $z_2=\bar{z}$ we have, for any $z>0$,

$$H(\boldsymbol{c}|\mathcal{Q}_z)=H(\boldsymbol{c}|\mathcal{Q})+\varrho\log\left(\frac{\bar{z}}{z}\right)+\sum_{i=1}^{\infty}Q_i(z^i-\bar{z}^i)$$

$$= H(\boldsymbol{c}|\mathcal{Q}) + \sum_{i=1}^{\infty} i Q_i \bar{z}^i \log\left(\frac{\bar{z}}{z}\right) + \sum_{i=1}^{\infty} Q_i z^i \left(1-\left(\frac{\bar{z}}{z}\right)^i\right)$$
$$= H(\boldsymbol{c}|\mathcal{Q}) + \sum_{i=1}^{\infty} Q_i z^i \left(\left(\frac{\bar{z}}{z}\right)^i \log\left(\left(\frac{\bar{z}}{z}\right)^i\right) - \left(\frac{\bar{z}}{z}\right)^i + 1\right)$$
$$= H(\boldsymbol{c}|\mathcal{Q}) + \sum_{i=1}^{\infty} Q_i z^i \Lambda\left(\frac{(\mathcal{Q}_z)_i}{\mathcal{Q}_i}\right),$$

where $\Lambda(x) = x\log x - x + 1 > 0$ when $x > 0$. □

Lemma B.2. *Let $\{Q_i\}_{i\in\mathbb{N}}$ be a nonnegative sequence such that $\lim_{i\to\infty} Q_{i+1}/Q_i = 1/r$ for some $r > 0$. Assume that $0 < x < r_1 < r$. Then*

$$\sum_{i=i_0+1}^{\infty} i Q_i x^{i-1} \leqslant C Q_{i_0} x^{i_0},$$

where C is a constant depending only on $\{Q_i\}_{i\in\mathbb{N}}$ and r_1.

Proof. Define $\beta_i = Q_{i+1}/Q_i$. We have that $\lim_{i\to\infty} \beta_i = 1/r$, and as such we can find $l \in \mathbb{N}$ such that for all $i > l$

$$\Lambda_1 = \sup_{i>l} \beta_i < \frac{1}{r_1}.$$

Let $\Lambda_2 = \sup_{i\leqslant l} \beta_i$. Since for any $i > i_0$

$$Q_i = \left(\prod_{j=i_0}^{i-1} \beta_j\right) Q_{i_0},$$

we see that

$$\sum_{i=i_0+1}^{\infty} i Q_i x^{i-1} = Q_{i_0} x^{i_0} \sum_{i=i_0+1}^{\infty} i \left(\prod_{j=i_0}^{i-1} \beta_j\right) x^{i-i_0-1}$$
$$\leqslant Q_{i_0} x^{i_0} \left(\Lambda_2 \sum_{j=0}^{l-i_0} i(\Lambda_2 r_1)^j + \Lambda_1 \sum_{j=l+1-i_0}^{\infty} i(\Lambda_1 r_1)^j\right)$$
$$\leqslant Q_{i_0} x^{i_0} \left(\Lambda_2 \sum_{j=0}^{l} j(\Lambda_2 r_1)^j + \Lambda_1 \sum_{j=0}^{\infty} j(\Lambda_1 r_1)^j\right),$$

completing the proof as l, Λ_1 and Λ_2 depend solely on $\{Q_i\}_{i\in\mathbb{N}}$. □

Lemma B.3. *Let $\varepsilon > 0$ and $\gamma > 0$. Define*

$$B_{\varepsilon,\gamma} = \sum_{i=1}^{\infty} i^{\gamma} e^{-\varepsilon i}.$$

Then $\varepsilon^{1+\gamma} B_{\varepsilon,\gamma}$ is of order 1 when ε goes to zero.

Proof. We start by noticing that the function $g_{\varepsilon,\gamma}(x) = x^\gamma e^{-\varepsilon x}$ is increasing in $\left[0, \frac{\gamma}{\varepsilon}\right]$ and decreasing in $\left[\frac{\gamma}{\varepsilon}, \infty\right)$. As such

$$B_{\varepsilon,\gamma} \geqslant \sum_{i=\left[\frac{\gamma}{\varepsilon}\right]+1}^{\infty} i^\gamma e^{-\varepsilon i} \geqslant \int_{\left[\frac{\gamma}{\varepsilon}\right]+1}^{\infty} x^\gamma e^{-\varepsilon x}\,\mathrm{d}x = \varepsilon^{-(1+\gamma)} \int_{\varepsilon\left(\left[\frac{\gamma}{\varepsilon}\right]+1\right)}^{\infty} y^\gamma e^{-y}\,\mathrm{d}y \geqslant \varepsilon^{-(1+\gamma)} \int_{\varepsilon}^{\infty} y^\gamma e^{-y}\,\mathrm{d}y,$$

showing the lower bound. For the upper bound we notice that

$$B_{\varepsilon,\gamma} \leqslant \sup_{x \geqslant 0} g_{\frac{\varepsilon}{e},\gamma}(x) \sum_{i=1}^{\infty} e^{-\frac{\varepsilon}{2} i} = \left(\frac{2\gamma}{\varepsilon}\right)^\gamma e^{-\gamma} \frac{e^{-\frac{\varepsilon}{2}}}{1-e^{-\frac{\varepsilon}{2}}},$$

which completes the proof since

$$\sup_{\varepsilon > 0} \frac{\varepsilon\, e^{-\frac{\varepsilon}{2}}}{1-e^{-\frac{\varepsilon}{2}}} < +\infty. \qquad \square$$

References

[Ball and Carr 1988] J. M. Ball and J. Carr, "Asymptotic behaviour of solutions to the Becker–Döring equations for arbitrary initial data", *Proc. Roy. Soc. Edinburgh Sect. A* **108**:1-2 (1988), 109–116. MR Zbl

[Ball and Carr 1990] J. M. Ball and J. Carr, "The discrete coagulation-fragmentation equations: existence, uniqueness, and density conservation", *J. Statist. Phys.* **61**:1-2 (1990), 203–234. MR Zbl

[Ball et al. 1986] J. M. Ball, J. Carr, and O. Penrose, "The Becker–Döring cluster equations: basic properties and asymptotic behaviour of solutions", *Comm. Math. Phys.* **104**:4 (1986), 657–692. MR Zbl

[Barthe and Roberto 2003] F. Barthe and C. Roberto, "Sobolev inequalities for probability measures on the real line", *Studia Math.* **159**:3 (2003), 481–497. MR Zbl

[Becker and Döring 1935] R. Becker and W. Döring, "Kinetische Behandlung der Keimbildung in übersättigten Dämpfen", *Ann. Phys.* **416**:8 (1935), 719–752. Zbl

[Bobkov and Götze 1999] S. G. Bobkov and F. Götze, "Exponential integrability and transportation cost related to logarithmic Sobolev inequalities", *J. Funct. Anal.* **163**:1 (1999), 1–28. MR Zbl

[Burton 1977] J. J. Burton, "Nucleation theory", pp. 195–234 in *Statistical mechanics, Part A: Equilibrium techniques*, edited by B. J. Berne, Modern Theoretical Chemistry **5**, Springer, 1977.

[Cañizo 2005] J. A. Cañizo, "Asymptotic behaviour of solutions to the generalized Becker–Döring equations for general initial data", *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **461**:2064 (2005), 3731–3745. MR Zbl

[Cañizo 2007] J. A. Cañizo, "Convergence to equilibrium for the discrete coagulation-fragmentation equations with detailed balance", *J. Stat. Phys.* **129**:1 (2007), 1–26. MR Zbl

[Cañizo and Lods 2013] J. A. Cañizo and B. Lods, "Exponential convergence to equilibrium for subcritical solutions of the Becker–Döring equations", *J. Differential Equations* **255**:5 (2013), 905–950. MR Zbl

[Cañizo et al. 2017] J. A. Cañizo, A. Einav, and B. Lods, "Uniform moment propagation for the Becker–Döring equation", preprint, 2017. arXiv

[Capasso 2003] V. Capasso (editor), *Mathematical modelling for polymer processing: polymerization, crystallization, manufacturing*, Mathematics in Industry **2**, Springer, 2003. MR Zbl

[Carr 1992] J. Carr, "Asymptotic behaviour of solutions to the coagulation-fragmentation equations, I: The strong fragmentation case", *Proc. Roy. Soc. Edinburgh Sect. A* **121**:3-4 (1992), 231–244. MR Zbl

[Carr and da Costa 1994] J. Carr and F. P. da Costa, "Asymptotic behavior of solutions to the coagulation-fragmentation equations, II: Weak fragmentation", *J. Statist. Phys.* **77**:1-2 (1994), 89–123. MR Zbl

[Collet et al. 2002] J.-F. Collet, T. Goudon, F. Poupaud, and A. Vasseur, "The Beker–Döring system and its Lifshitz–Slyozov limit", *SIAM J. Appl. Math.* **62**:5 (2002), 1488–1500. MR Zbl [da Costa 1998] F. P. da Costa, "Asymptotic behaviour of low density solutions to the generalized Becker–Döring equations", *NoDEA Nonlinear Differential Equations Appl.* **5**:1 (1998), 23–37. MR Zbl

[Desvillettes et al. 2011] L. Desvillettes, C. Mouhot, and C. Villani, "Celebrating Cercignani's conjecture for the Boltzmann equation", *Kinet. Relat. Models* **4**:1 (2011), 277–294. MR Zbl

[Farjoun and Neu 2008] Y. Farjoun and J. C. Neu, "Exhaustion of nucleation in a closed system", *Phys. Rev. E* **78**:5 (2008), art. id. 051402.

[Farjoun and Neu 2011] Y. Farjoun and J. C. Neu, "Aggregation according to classical kinetics: from nucleation to coarsening", *Phys. Rev. E* **83**:5 (2011), art. id. 051607.

[Jabin and Niethammer 2003] P.-E. Jabin and B. Niethammer, "On the rate of convergence to equilibrium in the Becker–Döring equations", *J. Differential Equations* **191**:2 (2003), 518–543. MR Zbl

[Kelton et al. 1983] K. F. Kelton, A. L. Greer, and C. V. Thompson, "Transient nucleation in condensed systems", *J. Chem. Phys.* **79**:12 (1983), 6261–6276.

[Murray and Pego 2015] R. W. Murray and R. L. Pego, "Polynomial decay to equilibrium for the Becker–Döring equations", preprint, 2015. arXiv

[Neu et al. 2002] J. C. Neu, J. A. Cañizo, and L. L. Bonilla, "Three eras of micellization", *Phys. Rev. E* **66**:6 (2002), art. id. 061406.

[Neu et al. 2005] J. C. Neu, L. L. Bonilla, and A. Carpio, "Igniting homogeneous nucleation", *Phys. Rev. E* (3) **71**:2 (2005), art. id. 021601. MR

[Niethammer 2003] B. Niethammer, "On the evolution of large clusters in the Becker–Döring model", *J. Nonlinear Sci.* **13**:1 (2003), 115–155. MR Zbl

[Oxtoby 1992] D. W. Oxtoby, "Homogeneous nucleation: theory and experiment", *J. Phys. Cond. Matter* **4**:38 (1992), 7620–7650.

[Penrose 1989] O. Penrose, "Metastable states for the Becker–Döring cluster equations", *Comm. Math. Phys.* **124**:4 (1989), 515–541. MR Zbl

[Penrose 1997] O. Penrose, "The Becker–Döring equations at large times and their connection with the LSW theory of coarsening", *J. Statist. Phys.* **89**:1-2 (1997), 305–320. MR Zbl

[Penrose and Lebowitz 1979] O. Penrose and J. L. Lebowitz, "Towards a rigorous molecular theory of metastability", pp. 293–340 in *Fluctuation phenomena*, edited by E. Montroll and J. L. Lebowitz, Studies in Statisical Mechanics **7**, North-Holland, New York, 1979.

[Rothaus 1985] O. S. Rothaus, "Analytic inequalities, isoperimetric inequalities and logarithmic Sobolev inequalities", *J. Funct. Anal.* **64**:2 (1985), 296–313. MR Zbl

[Schmelzer 2005] J. Schmelzer (editor), *Nucleation theory and applications*, Wiley-VCH, Weinheim, 2005.

[Velázquez 1998] J. J. L. Velázquez, "The Becker–Döring equations and the Lifshitz–Slyozov theory of coarsening", *J. Statist. Phys.* **92**:1-2 (1998), 195–236. MR Zbl

[Villani 2003] C. Villani, "Cercignani's conjecture is sometimes true and always almost true", *Comm. Math. Phys.* **234**:3 (2003), 455–490. MR Zbl

[Xiao and Haasen 1991] S. Q. Xiao and P. Haasen, "HREM investigation of homogeneous decomposition in a ni-12 at. % a1 alloy", *Acta. Metall. Mater.* **39**:4 (1991), 651–659.

Received 24 Dec 2016. Revised 28 Mar 2017. Accepted 29 May 2017.

JOSÉ A. CAÑIZO: canizo@ugr.es
Departamento de Matemática Aplicada, Universidad de Granada, Av. Fuentenueva S/N, 18071 Granada, Spain

AMIT EINAV: a.einav@dpmms.cam.ac.uk
Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, CB3 0WB, United Kingdom

BERTRAND LODS: bertrand.lods@unito.it
Department of Economics and Statistics & Collegio Carlo Alberto, Università degli Studi di Torino,
Corso Unione Sovietica, 218/bis, 10134 Torino, Italy

ANALYSIS AND PDE
Vol. 10, No. 7, 2017
dx.doi.org/10.2140/apde.2017.10.1709

msp

THE A_∞-PROPERTY OF THE KOLMOGOROV MEASURE

KAJ NYSTRÖM

We consider the Kolmogorov–Fokker–Planck operator

$$\mathcal{K} := \sum_{i=1}^{m} \partial_{x_i x_i} + \sum_{i=1}^{m} x_i \partial_{y_i} - \partial_t$$

in unbounded domains of the form

$$\Omega = \{(x, x_m, y, y_m, t) \in \mathbb{R}^{N+1} \mid x_m > \psi(x, y, t)\}.$$

Concerning ψ and Ω, we assume that Ω is what we call an (unbounded) admissible Lip_K-domain: ψ satisfies a uniform Lipschitz condition, adapted to the dilation structure and the (non-Euclidean) Lie group underlying the operator $\mathcal{K}$, as well as an additional regularity condition formulated in terms of a Carleson measure. We prove that in admissible Lip_K-domains the associated parabolic measure is absolutely continuous with respect to a surface measure and that the associated Radon–Nikodym derivative defines an A_∞ weight with respect to this surface measure. Our result is sharp.

1. Introduction

A major breakthrough in the study of boundary value problems for the heat operator

$$\mathcal{H} := \sum_{i=1}^{m} \partial_{x_i x_i} - \partial_t, \tag{1-1}$$

in $\mathbb{R}^{m+1}$, $m \geq 1$, in (unbounded) Lipschitz-type domains

$$\Omega = \{(x, x_m, t) \in \mathbb{R}^{m+1} \mid x_m > \psi(x, t)\}, \tag{1-2}$$

was achieved in [Lewis and Silver 1988; Lewis and Murray 1995; Hofmann and Lewis 1996; Hofmann 1997]; see also [Hofmann and Lewis 2001b]. In these papers the correct notion of time-dependent Lipschitz-type cylinders, correct from the perspective of parabolic measure, parabolic singular integral operators, parabolic layer potentials, as well as from the perspective of the Dirichlet, Neumann and regularity problems with data in L^p for the heat operator, was found. In particular, in [Lewis and Silver 1988; Lewis and Murray 1995] the mutual absolute continuity of the parabolic measure, with respect to surface measure, and the A_∞-property was studied/established and in [Hofmann and Lewis 1996]

The author was supported by a grant from the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine.
MSC2010: primary 35K65, 35K70; secondary 35H20, 35R03.
Keywords: Kolmogorov equation, ultraparabolic, hypoelliptic, Lipschitz domain, doubling measure, parabolic measure, Kolmogorov measure, A_∞.

the authors solved the Dirichlet, Neumann and regularity problems with data in L^2. The Neumann and regularity problems with data in L^p were considered in [Hofmann and Lewis 1999; 2005]. For further and related results concerning the fine properties of parabolic measures we refer to the impressive and influential work [Hofmann and Lewis 2001a].

The assumptions on the time-dependent function ψ underlying the analysis in all of the papers mentioned can be formulated as follows: there exist constants $0 < M_1, M_2 < \infty$ such that

$$|\psi(x,t)-\psi(\tilde x,\tilde t)| \le M_1(|x-\tilde x|+|t-\tilde t|^{1/2}) \tag{1-3}$$

whenever $(x,t), (\tilde x,\tilde t) \in \mathbb{R}^m$ and such that

$$\sup_{(x,t)\in\mathbb{R}^m,\, r>0} r^{-(m+1)}\int_0^r\int_{B_\lambda(x,t)}(\gamma_\psi(\tilde x,\tilde t,\lambda))^2\,\frac{d\tilde x\,d\tilde t\,d\lambda}{\lambda} \le M_2. \tag{1-4}$$

In (1-4), $B_\lambda(x,t)$ is the parabolic ball centered at $(x,t)\in\mathbb{R}^m$, with radius λ, and

$$\gamma_\psi(\tilde x,\tilde t,\lambda) := \left(\lambda^{-(m+1)}\int_{B_\lambda(\tilde x,\tilde t)}\left|\frac{\psi(\bar x,\bar t)-\psi(\tilde x,\tilde t)-\mathcal{P}_\lambda(\nabla_x\psi)(\tilde x,\tilde t)(\bar x-\tilde x)}{\lambda}\right|^2 d\bar x\,d\bar t\right)^{1/2}, \tag{1-5}$$

where $\mathcal{P}\in C_0^\infty(B_1(0,0))$ is a standard approximation of the identity, $\mathcal{P}_\lambda(x,t)=\lambda^{-(m+1)}\mathcal{P}(\lambda^{-1}x,\lambda^{-2}t)$, for $\lambda>0$, and $\mathcal{P}_\lambda(\nabla_x\psi)$ denotes the convolution of $\nabla_x\psi$ with $\mathcal{P}_\lambda$. Inequality (1-3) is sufficient for the validity of the doubling property of the caloric/parabolic measure, while the additional regularity imposed through (1-4) is necessary and sufficient for the A_∞-property of caloric measure, with respect to the surface measure $d\sigma_t\,dt$, to hold: this is a consequence of [Lewis and Silver 1988; Lewis and Murray 1995; Hofmann 1997; Hofmann et al. 2003; 2004].

In this paper we initiate the corresponding developments for the Kolmogorov–Fokker–Planck operator

$$\mathcal{K} := \sum_{i=1}^m \partial_{x_ix_i} + \sum_{i=1}^m x_i\partial_{y_i} - \partial_t \tag{1-6}$$

in $\mathbb{R}^{N+1}$, $N=2m$, $m\ge 1$, equipped with coordinates $(X,Y,t) := (x_1,\dots,x_m,y_1,\dots,y_m,t)\in\mathbb{R}^m\times\mathbb{R}^m\times\mathbb{R}$, in unbounded domains of the form

$$\Omega = \{(x,x_m,y,y_m,t)\in\mathbb{R}^{N+1} \mid x_m > \psi(x,y,t)\}, \tag{1-7}$$

The function $\psi:\mathbb{R}^{m-1}\times\mathbb{R}^{m-1}\times\mathbb{R}\to\mathbb{R}$ is, for reasons to be explained, assumed to be independent of the variable y_m.

The operator $\mathcal{K}$, referred to as the Kolmogorov or Kolmogorov–Fokker–Planck operator plays a central role in many application in analysis, physics and finance. $\mathcal{K}$ was introduced and studied by Kolmogorov [1934] as an example of a degenerate parabolic operator having strong regularity properties. Kolmogorov proved that $\mathcal{K}$ has a fundamental solution $\Gamma=\Gamma(X,Y,t,\widetilde X,\widetilde Y,\tilde t)$ which is smooth on the set $\{(X,Y,t)\neq(\widetilde X,\widetilde Y,\tilde t)\}$. As a consequence,

$$\mathcal{K}u := f\in C^\infty \;\Rightarrow\; u\in C^\infty \tag{1-8}$$

for every distributional solution of $\mathcal{K}u = f$. The property in (1-8) can also be stated as

$$\mathcal{K} \text{ is hypoelliptic;} \tag{1-9}$$

see (2-3) below.

Kolmogorov was originally motivated by statistical physics and he studied $\mathcal{K}$ in the context of stochastic processes. Indeed, the fundamental solution $\Gamma(\,\cdot\,,\cdot\,,\cdot, \widetilde{X}, \widetilde{Y}, \tilde{t})$ describes the density of the stochastic process (X_t, Y_t), which solves the Langevin system

$$\begin{cases} dX_t = \sqrt{2}dW_t, & X_{\tilde{t}} = \widetilde{X}, \\ dY_t = X_t dt, & Y_{\tilde{t}} = \widetilde{Y}, \end{cases} \tag{1-10}$$

where W_t is an m-dimensional Wiener process. The system in (1-10) describes the density of a system with $2m$ degrees of freedom. Given $Z := (X, Y) \in \mathbb{R}^{2m}$, the variables $X = (x_1, \dots, x_m)$ and $Y = (y_1, \dots, y_m)$ are, respectively, the velocity and the position of the system. The model in (1-10) and the equation in (1-6) are of fundamental importance in kinetic theory as they form the basis for Langevin-type models for particle dispersion, see [Bernardin et al. 2009; 2010; Chauvin et al. 2010; Bossy et al. 2011; Pope 2000], but they also appear in many other applied areas including finance [Barucci et al. 2001; Pascucci 2011], and vision [Citti and Sarti 2006; 2014].

In this paper we are concerned with the solvability of the Dirichlet problem for the operator $\mathcal{K}$ in unbounded domains of the form (1-7), and throughout the paper we will assume that Ω is a Lip_K-domain in the sense of Definition 1.1 below. Given $\varphi \in C(\partial\Omega)$ with compact support, we consider the boundary value problem

$$\begin{cases} \mathcal{K}u = 0 & \text{in } \Omega, \\ \quad u = \varphi & \text{on } \partial\Omega. \end{cases} \tag{1-11}$$

Using the Perron–Wiener–Brelot method one can prove the existence of a solution to this problem and, in the sequel, $u = u_\varphi$ will denote this solution to (1-11). Using the results of [Manfredini 1997], and assuming that Ω is a Lip_K-domain in the sense of Definition 1.1, it follows that all points on $\partial\Omega$ are regular for the Dirichlet problem for $\mathcal{K}$, i.e.,

$$\lim_{\substack{(Z,t)\to(Z_0,t_0)\\(Z,t)\in\Omega}} u_\varphi(Z,t) = \varphi(Z_0,t_0) \quad \text{for any } \varphi \in C(\partial\Omega) \tag{1-12}$$

whenever $(Z_0, t_0) \in \partial\Omega$. Moreover, there exists, for every $(Z, t) \in \Omega$, a unique probability measure $\omega(Z, t, \cdot)$ on $\partial\Omega$ such that

$$u(Z,t) = \int_{\partial\Omega} \varphi(\widetilde{Z}, \tilde{t})\, d\omega(Z, t, \widetilde{Z}, \tilde{t}). \tag{1-13}$$

We refer to $\omega(Z, t, \cdot)$ as the Kolmogorov measure, or parabolic measure, associated to $\mathcal{K}$, relative to (Z, t) and Ω. In this paper we are particularly interested in scale- and translation-invariant estimates of $\omega(Z, t, \cdot)$ in terms of a (physical) surface measure, $d\sigma$, on $\partial\Omega$. In particular, assuming that Ω is an admissible Lip_K-domain in the sense of Definition 1.1 below, we establish a scale-invariant form of mutual absolute continuity of $\omega(Z, t, \cdot)$ with respect to $d\sigma$ on $\partial\Omega$.

Despite the relevance of the operator $\mathcal{K}$ to analysis, stochastics, physics, and in the applied sciences, the analysis of its properties is in several respects fundamentally underdeveloped. Indeed, geometry, the fine properties of the Dirichlet problem in (1-11) and the Kolmogorov measure, the boundary behavior of nonnegative solutions and the Green function, are currently only modestly studied and explored in the literature. One reason for this may be the intrinsic and intricate complexity built into the operator $\mathcal{K}$ through the lack of diffusion in the coordinates $(y_1, \dots, y_m)$ and through the presence of the lower-order drift term $\sum_{i=1}^{m} x_i \partial_{y_i} - \partial_t$. These two features of $\mathcal{K}$, which make this operator decisively different from the heat operator $\mathcal{H}$, have the consequence that the Lie group of translations $(\mathbb{R}^{N+1}, \circ)$ underlying $\mathcal{K}$ is different from the standard group of Euclidean translations and that already fundamental principles like the Harnack inequality and the construction of appropriate Harnack chains under geometrical restrictions become issues; see [Nyström and Polidoro 2016].

To briefly outline the current state of the literature, in our context, it is fair to mention that the first proof of the scale-invariant Harnack inequality, which constitutes one of the building blocks for our paper, can be found in [Garofalo and Lanconelli 1990]. In that paper the Harnack inequality is expressed in terms of level sets of the fundamental solution; hence it depends implicitly on the underlying Lie group structure. This fact was used in [Lanconelli and Polidoro 1994], where the group law, see (1-15) below, was used explicitly and the Harnack inequality, in the form we use it, was proved for the first time. The Perron–Wiener–Brelot method in the context of the Dirichlet problem in (1-11), as well as criteria based on which boundary points can be proved to be regular, were developed in the important work [Manfredini 1997]. In [Cinti et al. 2010; 2012; 2013], the author, together with Chiara Cinti and Sergio Polidoro, developed a number of important preliminary estimates concerning the boundary behavior of nonnegative solutions to equations of Kolmogorov–Fokker–Planck type in Lipschitz-type domains. These papers were the result of our ambition to establish scale- and translation-invariant boundary comparison principles, boundary Harnack inequalities and doubling properties of associated parabolic measures, results previously established for uniformly parabolic equations with bounded measurable coefficients in Lipschitz-type domains, see [Fabes and Safonov 1997; Fabes et al. 1986; 1999; Safonov and Yuan 1999; Nyström 1997; Salsa 1981], for nonnegative solutions to the equation $\mathcal{K}u = 0$ and for more general equations of Kolmogorov–Fokker–Planck type. In [Nyström and Polidoro 2016], the author together with Sergio Polidoro took the program started in [Cinti et al. 2010; 2012; 2013] a large step forward by establishing several new results concerning the boundary behavior of nonnegative solutions to the equation $\mathcal{K}u = 0$ near the noncharacteristic part of the boundary of local versions of the Lip_K-domains defined in Definition 1.1 below. Generalizations to more general operators of Kolmogorov–Fokker–Planck type were also discussed. In particular, in [Nyström and Polidoro 2016] scale- and translation-invariant quantitative estimates concerning the behavior, at the boundary, for nonnegative solutions vanishing on a portion of the boundary were proved as well as a scale- and translation-invariant doubling property of the Kolmogorov measure. The results in that paper are developed under the regularity condition stated below in (1-25) in Definition 1.1; in particular, for reasons that they explain in detail, the results, including the translation-invariant doubling property of the Kolmogorov measure, were derived using the assumption that the defining function for Ω in (1-7), ψ, was assumed to be independent of the variable y_m. This assumption gave us a crucial additional degree of freedom at our disposal when building Harnack chains to connect points: we could freely connect points in the x_m-variable, taking geometric restriction into account, accepting that the path in the y_m-variable will most likely not end up in "the right spot". This is one reason why we also in this paper consider domains which are constant in the y_m-direction.

The main achievement of this paper is that we take the analysis in [Nyström and Polidoro 2016] one step further by proving, see Theorem 1.6 below, that if Ω is an admissible Lip_K-domain with constants (M_1, M_2) in the sense of Definition 1.1 below, then ω is mutually absolutely continuous with respect to a (physical) surface measure σ on $\partial\Omega$ and $\omega \in A_\infty(\partial\Omega, d\sigma)$ with constants depending only on N, M_1 and M_2. This gives a generalization of [Lewis and Silver 1988; Lewis and Murray 1995] to the operator $\mathcal{K}$ and in the case of graphs which are independent of all y-variables, our assumptions coincide with the geometrical conditions underlying [Lewis and Silver 1988; Lewis and Murray 1995; Hofmann and Lewis 1996; Hofmann 1997]; see (1-3) and (1-4) above.

1A. *Notation.* The natural family of dilations for $\mathcal{K}$, denoted by $(\delta_r)_{r>0}$, on $\mathbb{R}^{N+1}$ is defined by

$$\delta_r(X, Y, t) = (rX, r^3Y, r^2t) \tag{1-14}$$

for $(X, Y, t) \in \mathbb{R}^{N+1}$, $r > 0$. Due to the presence of nonconstant coefficients in the drift term of $\mathcal{K}$, the usual Euclidean change of variable does not preserve the Kolmogorov equation. Instead the Lie group on $\mathbb{R}^{N+1}$ preserving $\mathcal{K}u = 0$ is defined by the group law

$$(\widetilde{Z}, \tilde{t}) \circ (Z, t) = (\widetilde{X}, \widetilde{Y}, \tilde{t}) \circ (X, Y, t) = (\widetilde{X} + X, \widetilde{Y} + Y - t\widetilde{X}, \tilde{t} + t) \tag{1-15}$$

whenever $(Z, t), (\widetilde{Z}, \tilde{t}) \in \mathbb{R}^{N+1}$. Note that

$$(Z, t)^{-1} = (X, Y, t)^{-1} = (-X, -Y - tX, -t), \tag{1-16}$$

and hence

$$(\widetilde{Z}, \tilde{t})^{-1} \circ (Z, t) = (\widetilde{X}, \widetilde{Y}, \tilde{t})^{-1} \circ (X, Y, t) = (X - \widetilde{X}, Y - \widetilde{Y} + (t - \tilde{t})\widetilde{X}, t - \tilde{t}) \tag{1-17}$$

whenever $(Z, t), (\widetilde{Z}, \tilde{t}) \in \mathbb{R}^{N+1}$. Given $(Z, t) = (X, Y, t) \in \mathbb{R}^{N+1}$ we let

$$\|(Z, t)\| = \|(X, Y, t)\| := |(X, Y)| + |t|^{1/2}, \quad |(X, Y)| = |X| + |Y|^{1/3}. \tag{1-18}$$

Note that $\|\delta_r(X, Y, t)\| = r\|(X, Y, t)\|$ when $(X, Y, t) \in \mathbb{R}^{N+1}$, $r > 0$. We define

$$d((Z, t), (\widetilde{Z}, \tilde{t})) := \tfrac{1}{2}\big(\|(\widetilde{Z}, \tilde{t})^{-1} \circ (Z, t)\| + \|(Z, t)^{-1} \circ (\widetilde{Z}, \tilde{t})\|\big). \tag{1-19}$$

Then, as discussed in the bulk of the paper, d is a symmetric quasidistance on $\mathbb{R}^{N+1}$. Based on d we introduce the balls

$$\mathcal{B}_r(Z, t) := \{(\widetilde{Z}, \tilde{t}) \in \mathbb{R}^{N+1} \mid d((\widetilde{Z}, \tilde{t}), (Z, t)) < r\} \tag{1-20}$$

for $(Z, t) \in \mathbb{R}^{N+1}$ and $r > 0$. The measure of the ball $\mathcal{B}_r(Z, t)$, denoted by $|\mathcal{B}_r(Z, t)|$, is approximately $r^{\boldsymbol{q}}$, where $\boldsymbol{q} := 4m + 2$, independent of (Z, t). Similarly, given $(z, t) = (x, y, t) \in \mathbb{R}^{N-1} = \mathbb{R}^{m-1} \times \mathbb{R}^{m-1} \times \mathbb{R}$ we let

$$\mathcal{B}_r(z, t) := \{(\tilde{z}, \tilde{t}) \in \mathbb{R}^{N-1} \mid d((\tilde{x}, 0, \tilde{y}, 0, \tilde{t}), (x, 0, y, 0, t)) < r\}. \tag{1-21}$$

The measure of the ball $\mathcal{B}_r(z,t)$, denoted by $|\mathcal{B}_r(z,t)|$, is approximately $r^{\boldsymbol{q}-4}$, independent of (z,t). With a slight abuse of notation we will by $\mathcal{B}_r(Z,t)$, note the capital Z, always denote a ball in $\mathbb{R}^{N+1}$, and by $\mathcal{B}_r(z,t)$, note the lowercase z, we will always denote a ball in $\mathbb{R}^{N-1}$.

1B. *Geometry.* Our geometrical setting is that of unbounded domains of the form

$$\Omega = \{(x, x_m, y, y_m, t) \in \mathbb{R}^{N+1} \mid x_m > \psi(x,y,t)\}, \tag{1-22}$$

and here we define the restrictions that we impose on the function $\psi : \mathbb{R}^{m-1}\times\mathbb{R}^{m-1}\times\mathbb{R}\to\mathbb{R}$. Let $\mathcal{P}\in C_0^\infty(\mathcal{B}_1(0,0))$, where $\mathcal{B}_1(0,0)\subset\mathbb{R}^{N-1}$, be a standard approximation of the identity. Let $\mathcal{P}_\lambda(x,y,t) = \lambda^{-(\boldsymbol{q}-4)}\mathcal{P}(\lambda^{-1}x, \lambda^{-3}y, \lambda^{-2}t)$ for $\lambda>0$. Given a function f defined on $\mathbb{R}^{N-1}$ we let

$$\begin{aligned} \mathcal{P}_\lambda f(x,y,t) &:= \int_{\mathbb{R}^{N-1}} f(\bar{x},\bar{y},\bar{t})\,\mathcal{P}_\lambda\big((\bar{x},\bar{y},\bar{t})^{-1}\circ(x,y,t)\big)\,d\bar{x}\,d\bar{y}\,d\bar{t} \\ &= \int_{\mathbb{R}^{N-1}} f(\bar{x},\bar{y},\bar{t})\,\mathcal{P}_\lambda\big(x-\bar{x},\, y-\bar{y}+(t-\bar{t})\bar{x},\, t-\bar{t}\big)\,d\bar{x}\,d\bar{y}\,d\bar{t}. \end{aligned} \tag{1-23}$$

$\mathcal{P}_\lambda f$ represents a regularization of f. Given $(\tilde{z},\tilde{t})\in\mathbb{R}^{N-1}$, $\lambda>0$, we introduce

$$\gamma_\psi(\tilde{z},\tilde{t},\lambda) := \left(\lambda^{-(\boldsymbol{q}-4)}\int_{\mathcal{B}_\lambda(\tilde{z},\tilde{t})}\left|\frac{\psi(\bar{x},\bar{y},\bar{t})-\psi(\tilde{x},\tilde{y},\tilde{t})-\mathcal{P}_\lambda(\nabla_x\psi)(\tilde{x},\tilde{y},\tilde{t})(\bar{x}-\tilde{x})}{\lambda}\right|^2 d\bar{x}\,d\bar{y}\,d\bar{t}\right)^{1/2}. \tag{1-24}$$

We are now ready to formulate our conditions on $\psi : \mathbb{R}^{m-1}\times\mathbb{R}^{m-1}\times\mathbb{R}\to\mathbb{R}$ and Ω.

Definition 1.1. Assume that there exist constants $0 < M_1, M_2 < \infty$ such that

$$|\psi(z,t)-\psi(\tilde{z},\tilde{t})| \le M_1\|(\tilde{z},\tilde{t})^{-1}\circ(z,t)\| \tag{1-25}$$

whenever $(z,t), (\tilde{z},\tilde{t})\in\mathbb{R}^{N-1}$ and such that

$$\sup_{(z,t)\in\mathbb{R}^{N-1},\, r>0} r^{-(\boldsymbol{q}-4)}\int_0^r\int_{\mathcal{B}_\lambda(z,t)} (\gamma_\psi(\tilde{z},\tilde{t},\lambda))^2\,\frac{d\tilde{z}\,d\tilde{t}\,d\lambda}{\lambda} \le M_2. \tag{1-26}$$

Let $\Omega=\Omega_\psi$ be defined as in (1-22). We say that Ω, defined by a function ψ satisfying (1-25), is a Lip_K-domain with constant M_1. We say that Ω, defined by a function ψ satisfying (1-25) and (1-26), is an admissible Lip_K-domain with constants (M_1, M_2).

Remark 1.2. Inequality (1-25) implies

$$\begin{aligned} |\psi(x,y,t)-\psi(\tilde{x},y,t)| &\le M_1\|(x-\tilde{x},0,0)\| = M_1|x-\tilde{x}|, \\ |\psi(x,y,t)-\psi(x,\tilde{y},t)| &\le M_1\|(0,y-\tilde{y},0)\| = M_1|y-\tilde{y}|^{1/3}, \\ |\psi(x,y,t)-\psi(x,y,\tilde{t})| &\le M_1\|(0,(t-\tilde{t})x,(t-\tilde{t}))\| = M_1(|(t-\tilde{t})x|^{1/3}+|t-\tilde{t}|^{1/2}) \end{aligned} \tag{1-27}$$

uniformly with respect to the remaining variables. From the perspective of dilations and translations, Lip_K-domains are, assuming y_m-independence, the natural replacement in the context of the operator $\mathcal{K}$ of the $\mathrm{Lip}\big(1,\frac{1}{2}\big)$-domains considered in the context of the heat operator.

Remark 1.3. Inequality (1-26) states that the measure

$$(\gamma_\psi(\tilde z,\tilde t,\lambda))^2\,\frac{d\tilde z\,d\tilde t\,d\lambda}{\lambda}$$

is a Carleson measure on $\mathbb{R}^{N-1}\times\mathbb{R}_+$. In this paper we prove, from the perspective of the finer properties of the Kolmogorov measure, that admissible Lip_K-domains are, assuming y_m-independence, the natural replacements in the context of the operator $\mathcal{K}$ of the admissible time-varying domains discovered and explored in [Lewis and Murray 1995; Hofmann 1997; Hofmann and Lewis 1996; 2001b] in the context of the heat operator.

Remark 1.4. Assume that $\Omega=\Omega_\psi\subset\mathbb{R}^{N+1}$ is a Lip_K-domain, with constant M_1. We define a (physical) measure σ on $\partial\Omega$ as

$$d\sigma(X,Y,t):=\sqrt{1+|\nabla_x\psi(x,y,t)|^2}\,dx\,dY\,dt,\qquad (X,Y,t)\in\partial\Omega. \tag{1-28}$$

We will refer to σ as the surface measure on $\partial\Omega$.

1C. ***Statement of the main result.*** Given $\varrho>0$ and $\Lambda>0$, we let

$$A^+_{\varrho,\Lambda}=\big(0,\Lambda\varrho,0,-\tfrac{2}{3}\Lambda\varrho^3,\varrho^2\big)\in\mathbb{R}^{m-1}\times\mathbb{R}\times\mathbb{R}^{m-1}\times\mathbb{R}\times\mathbb{R}. \tag{1-29}$$

We let

$$A^+_{\varrho,\Lambda}(Z_0,t_0)=(Z_0,t_0)\circ A^+_{\varrho,\Lambda} \tag{1-30}$$

whenever $(Z_0,t_0)\in\mathbb{R}^{N+1}$. Using the main result of [Nyström and Polidoro 2016], see Lemma 4.12 below, one can prove the following theorem.

Theorem 1.5. *Assume that $\Omega=\Omega_\psi\subset\mathbb{R}^{N+1}$ is a (unbounded) Lip_K-domain with constant M_1. Then there exist $\Lambda=\Lambda(N,M_1)$, $1\le\Lambda<\infty$, and $c=c(N,M_1)$, $1\le c<\infty$, such that the following is true. Let $(Z_0,t_0)\in\partial\Omega$, $0<\varrho_0<\infty$. Then*

$$\omega\big(A^+_{c\varrho_0,\Lambda}(Z_0,t_0),\partial\Omega\cap\mathcal{B}_{2\varrho}(\widetilde Z_0,\tilde t_0)\big)\le c\,\omega\big(A^+_{c\varrho_0,\Lambda}(Z_0,t_0),\partial\Omega\cap\mathcal{B}_{\varrho}(\widetilde Z_0,\tilde t_0)\big)$$

for all balls $\mathcal{B}_\varrho(\widetilde Z_0,\tilde t_0)$, $(\widetilde Z_0,\tilde t_0)\in\partial\Omega$ such that $\mathcal{B}_\varrho(\widetilde Z_0,\tilde t_0)\subset\mathcal{B}_{4\varrho_0}(Z_0,t_0)$.

The following is the main new result proved in this paper.

Theorem 1.6. *Assume that $\Omega\subset\mathbb{R}^{N+1}$ is an (unbounded) admissible Lip_K-domain with constants (M_1,M_2) in the sense of Definition 1.1. Then there exist $\Lambda=\Lambda(N,M_1)$, $1\le\Lambda<\infty$, and $c=c(N,M_1)$, $1\le c<\infty$, and $\tilde c=\tilde c(N,M_1,M_2)$, $1\le\tilde c<\infty$, and $\eta=\eta(N,M_1,M_2)$, $0<\eta<1$, such that the following is true. Let $(Z_0,t_0)\in\partial\Omega$, $0<\varrho_0<\infty$. Then*

$$\tilde c^{-1}\bigg(\frac{\sigma(E)}{\sigma(\partial\Omega\cap\mathcal{B}_\varrho(\widetilde Z_0,\tilde t_0))}\bigg)^{1/\eta}\le\frac{\omega(A^+_{c\varrho_0,\Lambda}(Z_0,t_0),E)}{\omega\big(A^+_{c\varrho_0,\Lambda}(Z_0,t_0),\partial\Omega\cap\mathcal{B}_\varrho(\widetilde Z_0,\tilde t_0)\big)}\le\tilde c\bigg(\frac{\sigma(E)}{\sigma(\partial\Omega\cap\mathcal{B}_\varrho(\widetilde Z_0,\tilde t_0))}\bigg)^{\eta}$$

whenever $E\subset\partial\Omega\cap\mathcal{B}_\varrho(\widetilde Z_0,\tilde t_0)$ for some ball $\mathcal{B}_\varrho(\widetilde Z_0,\tilde t_0)$, $(\widetilde Z_0,\tilde t_0)\in\partial\Omega$ such that $\mathcal{B}_\varrho(\widetilde Z_0,\tilde t_0)\subset\mathcal{B}_{\varrho_0}(Z_0,t_0)$.

Remark 1.7. A short formulation of the conclusion of Theorem 1.6 is that

$$\omega(A^+_{c\varrho_0,\Lambda}(Z_0,t_0),\cdot)\in A_\infty(\partial\Omega\cap\mathcal{B}_{\varrho_0}(Z_0,t_0),d\sigma)$$

for all $(Z_0,t_0)\in\partial\Omega,\ 0<\varrho_0<\infty$, and with constants independent of (Z_0,t_0) and ϱ_0.

Remark 1.8. Theorem 1.6 states that a sufficient condition for the conclusion that $\omega(A^+_{c\varrho_0,\Lambda}(Z_0,t_0),\cdot)\in A_\infty(\partial\Omega\cap\mathcal{B}_{\varrho_0}(Z_0,t_0),d\sigma)$ uniformly is that $\Omega\subset\mathbb{R}^{N+1}$ is an (unbounded) admissible Lip_K-domain with constants (M_1,M_2) in the sense of Definition 1.1. In fact, the condition in (1-26) in Definition 1.1 is also necessary in the following sense. Using [Lewis and Silver 1988; Hofmann et al. 2003] one can conclude that there exists a function $\psi:\mathbb{R}^{m-1}\times\mathbb{R}\to\mathbb{R}$ which satisfies (1-3) for some M_1, but violates (1-4) for all $M_2<\infty$, and such that the parabolic measure associated to the heat operator in

$$\{(x,x_m,t)\in\mathbb{R}^{m+1}\mid x_m>\psi(x,t)\},\tag{1-31}$$

denoted by $\omega_{\mathcal{H}}$, is singular with respect to the surface measure $d\sigma_t\,dt$. Obviously this ψ also satisfies (1-25) with constant M_1, but violates (1-26) for all $M_2<\infty$. Consider now the domain

$$\Omega:=\{(x,x_m,y,y_m,t)\in\mathbb{R}^{2m+1}\mid x_m>\psi(x,t)\},\tag{1-32}$$

which is constant as a function of (y,y_m). Using that solutions to $\mathcal{H}u=0$ also satisfy $\mathcal{K}u=0$, estimates for nonnegative solutions to $\mathcal{H}u=0$, see [Hofmann et al. 2004] for example, and Lemma 4.11, Theorem 4.8 and Theorem 4.9 stated below, it can then be proved that the Kolmogorov measure in Ω must be singular with respect to the surface measure $d\sigma$ defined in Remark 1.4.

1D. *Discussion of the proof.* To prove Theorem 1.6 it suffices to prove Theorem 5.1 below. To prove Theorem 5.1 we use, and expand on, results from [Nyström and Polidoro 2016] and we implement ideas similar to the ideas in the recent paper [Kenig et al. 2016], where similar types of results are established but in the context of elliptic measure and second-order elliptic operators in divergence form. The final part of our proof of Theorem 1.6 is based on a crucial square function estimate, Lemma 5.3 below. The lemma states that if $u(Z,t):=\omega(Z,t,S)$, where $S\subset\partial\Omega$ is a Borel set, and if $c=c(N,M_1)\geq 1$, then there exists $\tilde c=\tilde c(N,M_1,M_2),\ 1\leq\tilde c<\infty$, such that

$$\iint_{T_{cQ_0}}\big(|\nabla_X u|^2\delta+|\nabla_Y u|^2\delta^5+|(X\cdot\nabla_Y-\partial_t)(u)|^2\delta^3\big)\,dZ\,dt\leq\tilde c\sigma(Q_0),$$

where T_{cQ_0} is a Carleson box associated with cQ_0, where $Q_0\subset\Omega$ is a (dyadic) surface cube, see Section 5A and (5-13), and where $\delta=\delta(Z,t)$ is the relevant distance from $(Z,t)\in\Omega$ to $\partial\Omega$. To prove Lemma 5.3 and to enable partial integration, we use a Dahlberg–Kenig–Stein-type of mapping adapted to the underlying group law,

$$(w,w_m,y,y_m,t)\in U\to(w,w_m+P_{\gamma w_m}\psi(w,y,t),y,y_m,t),\tag{1-33}$$

where

$$U=\big\{(W,Y,t)=(w,w_m,y,y_m,t)\in\mathbb{R}^{m-1}\times\mathbb{R}\times\mathbb{R}^{m-1}\times\mathbb{R}\times\mathbb{R}\;\big|\;w_m>0\big\}.\tag{1-34}$$

Then u satisfies $\mathcal{K}u = 0$ in Ω if and only if $v(W, Y, t) = u(w, w_m + P_{\gamma w_m}\psi(w, y, t), y, y_m, t)$ satisfies

$$\nabla_W \cdot (A\nabla_W v) + B \cdot \nabla_W v + \big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)v = 0 \quad \text{in } U. \tag{1-35}$$

Using this change of coordinates, it turns out to be sufficient to prove Lemma 5.3 below for solutions to the equation in (1-35) and our proof explores, as a consequence of our assumptions on ψ and as discussed in Section 2, that the coefficients A and B are independent of the variable y_m and that A and B define certain Carleson measures on U; see (6-10) below.

1E. *Organization of the paper.* In Section 2 we give additional preliminaries and we discuss implications of the assumptions in (1-25) and (1-26). In particular, considering a Dahlberg–Kenig–Stein-type of mapping as in (1-33), (1-34), we prove, as a consequence of the assumptions on ψ, that certain measures defined based on ψ are Carleson measures; see Lemma 2.2. In Section 3 we discuss the Dirichlet problem (1-11). In Section 4 we state, and elaborate on, some crucial estimates from [Nyström and Polidoro 2016]. In Section 5 we prove Theorem 1.6, assuming the square function estimate referred to above. The proof of the square function estimate is then given in Section 6.

2. Preliminaries

As discussed in Section 1A, see (1-14)–(1-17), the natural family of dilations for $\mathcal{K}$, denoted by $(\delta_r)_{r>0}$, on $\mathbb{R}^{N+1}$, and the Lie group on $\mathbb{R}^{N+1}$ preserving $\mathcal{K}u = 0$ are different from standard parabolic dilations and Euclidean translations applicable in the context of the heat operator. Using the notation of Section 1A, the operator $\mathcal{K}$ is δ_r-homogeneous of degree two, i.e., $\mathcal{K} \circ \delta_r = r^2(\delta_r \circ \mathcal{K})$, for all $r > 0$, and the operator $\mathcal{K}$ can be expressed as

$$\mathcal{K} = \sum_{i=1}^{m} X_i^2 + X_0,$$

where

$$X_i := \partial_{x_i}, \quad i = 1, \dots, m, \qquad X_0 := \sum_{i=1}^{m} x_i \partial_{y_i} - \partial_t, \tag{2-1}$$

and the vector fields $X_1, \dots, X_m$ and X_0 are left-invariant with respect to the group law (1-15) in the sense that

$$X_i(u((\widetilde{Z}, \tilde{t}) \circ \cdot)) = (X_i u)((\widetilde{Z}, \tilde{t}) \circ \cdot), \quad i = 0, \dots, m, \tag{2-2}$$

for every $(\widetilde{Z}, \tilde{t}) \in \mathbb{R}^{N+1}$. Consequently,

$$\mathcal{K}(u((\widetilde{Z}, \tilde{t}) \circ \cdot)) = (\mathcal{K}u)((\widetilde{Z}, \tilde{t}) \circ \cdot).$$

Taking commutators we see that $[X_i, X_0] = \partial_{y_i}$ for $i \in \{1, \dots, m\}$ and that the vector fields $\{X_1, \dots, X_m, X_0\}$ generate the Lie algebra associated to the Lie group $(\mathbb{R}^{N+1}, \circ)$. In particular, (1-9) is equivalent to the Hörmander condition,

$$\operatorname{rank} \operatorname{Lie}(X_1, \dots, X_m, X_0)(Z, t) = N + 1 \quad \text{for all } (Z, t) \in \mathbb{R}^{N+1}; \tag{2-3}$$

see [Hörmander 1967]. Furthermore, while X_i represents a differential operator of order one, ∂_{y_i} acts as a third-order operator. This fact is also reflected in the dilations group $(\delta_r)_{r>0}$ defined above.

2A. ***A symmetric quasidistance.*** Recall the notation $\|(Z,t)\| = \|(X,Y,t)\|$ for $(Z,t)=(X,Y,t)\in\mathbb{R}^{N+1}$, introduced in (1-18). We recall the following pseudotriangular inequality: there exists a positive constant c such that

$$\|(Z,t)^{-1}\| \le c\|(Z,t)\|, \quad \|(Z,t)\circ(\widetilde{Z},\tilde{t})\| \le c(\|(Z,t)\| + \|(\widetilde{Z},\tilde{t})\|) \tag{2-4}$$

whenever $(Z,t), (\widetilde{Z},\tilde{t}) \in \mathbb{R}^{N+1}$. Using (2-4) it follows directly that

$$\|(\widetilde{Z},\tilde{t})^{-1}\circ(Z,t)\| \le c\,\|(Z,t)^{-1}\circ(\widetilde{Z},\tilde{t})\| \tag{2-5}$$

whenever $(Z,t), (\widetilde{Z},\tilde{t}) \in \mathbb{R}^{N+1}$. Furthermore, defining $d((Z,t),(\widetilde{Z},\tilde{t}))$ as in (1-19), and using (2-5), it follows that

$$\|(\widetilde{Z},\tilde{t})^{-1}\circ(Z,t)\| \sim d((Z,t),(\widetilde{Z},\tilde{t})) \sim \|(Z,t)^{-1}\circ(\widetilde{Z},\tilde{t})\| \tag{2-6}$$

with constants of comparison independent of $(Z,t), (\widetilde{Z},\tilde{t}) \in \mathbb{R}^{N+1}$. Again using (2-4) we also see that

$$d((Z,t),(\widetilde{Z},\tilde{t})) \le c\big(d((Z,t),(\widehat{Z},\hat{t})) + d((\widehat{Z},\hat{t}),(\widetilde{Z},\tilde{t}))\big) \tag{2-7}$$

whenever $(Z,t), (\widehat{Z},\hat{t}), (\widetilde{Z},\tilde{t}) \in \mathbb{R}^{N+1}$, and hence that d is a symmetric quasidistance. Based on d, in (1-20) we introduced the balls $\mathcal{B}_r(Z,t)$ for $(Z,t)\in\mathbb{R}^{N+1}$ and $r>0$, and in (1-21) we introduced the balls $\mathcal{B}_r(z,t)$ for $(z,t)\in\mathbb{R}^{N-1}$ and $r>0$. Note that

$$\begin{aligned}\mathcal{B}_r(Z,t) &= (Z,t)\circ\big\{(\widetilde{Z},\tilde{t})\in\mathbb{R}^{N+1} \bigm| \|(\widetilde{Z},\tilde{t})\| + \|(\widetilde{Z},\tilde{t})^{-1}\| < r\big\},\\ \mathcal{B}_r(z,t) &= (z,t)\circ\big\{(\tilde{z},\tilde{t})\in\mathbb{R}^{N-1} \bigm| \|(\tilde{z},\tilde{t})\| + \|(\tilde{z},\tilde{t})^{-1}\| < r\big\}.\end{aligned} \tag{2-8}$$

We emphasize that throughout the paper we will stick to the convention that $\mathcal{B}_r(Z,t)$, with a capital Z, always denotes a ball in $\mathbb{R}^{N+1}$, and that $\mathcal{B}_r(z,t)$, with a lowercase z, always denotes a ball in $\mathbb{R}^{N-1}$.

2B. ***Geometry and Carleson measures.*** Assume ψ satisfies (1-25) and (1-26) for some constants $0 < M_1, M_2 < \infty$. Let $\gamma\in(0,1)$ and consider the change of coordinates/mapping

$$(W,Y,t) = (w,w_m,y,y_m,t)\in U \to (w, w_m + P_{\gamma w_m}\psi(w,y,t), y, y_m, t)$$

defined in (1-33), where $P_{\gamma w_m}\psi(x,y,t)$ is defined in (1-23), and where U is defined in (1-34). This mapping is a version of the Dahlberg–Kenig–Stein mapping used in elliptic and parabolic problems. The purpose of this section is to prove properties of this change of coordinates assuming that ψ satisfies (1-25) and (1-26). In particular, we prove that if ψ satisfies (1-25) and (1-26), then certain measures, naturally associated to $P_{\gamma w_m}\psi$, are Carleson measures. Throughout the section and the paper $\mathcal{P}$ will denote a parabolic approximation of the identity chosen based on a finite stock of functions and fixed throughout the paper. Let $\mathcal{P}\in C_0^\infty(\mathcal{B}_1(0,0))$, where $\mathcal{B}_1(0,0)\subset\mathbb{R}^{N-1}$, $\mathcal{P}\ge 0$ be real-valued, and $\int\mathcal{P}\,dz\,dt = 1$. We will assume, as we may by imposing a product structure on $\mathcal{P}$, that $\mathcal{P}$ is even in the sense that

$$\int x_i\mathcal{P}(z,t)\,dz\,dt = \int y_i\mathcal{P}(z,t)\,dz\,dt = \int t\mathcal{P}(z,t)\,dz\,dt = 0 \tag{2-9}$$

for $i \in \{1, \dots, m-1\}$. We set $\mathcal{P}_\lambda(z,t) = \mathcal{P}_\lambda(x,y,t) = \lambda^{-(q-4)}\mathcal{P}(\lambda^{-1}x, \lambda^{-3}y, \lambda^{-2}t)$ whenever $\lambda > 0$. Given $\mathcal{P}$, we let $\mathcal{P}_\lambda$ define a convolution operator as introduced in (1-23). Similarly, we will by $\mathcal{Q}_\lambda$ denote a generic approximation to the zero operator, not necessarily the same at each instance, but chosen from a finite set of such operators depending only on our original choice of $\mathcal{P}_\lambda$. In particular, $\mathcal{Q}_\lambda(z,t) = \mathcal{Q}_\lambda(x,y,t) = \lambda^{-(q-4)}\mathcal{Q}(\lambda^{-1}x, \lambda^{-3}y, \lambda^{-2}t)$, where $\mathcal{Q} \in C_0^\infty(\mathcal{B}_1(0,0))$, $\int \mathcal{Q}\, dz\, dt = 0$. We first prove the following lemma.

Lemma 2.1. *Let ψ be a function satisfying (1-25) for some constant $0 < M_1 < \infty$, let $\gamma \in (0,1)$ and let $P_{\gamma w_m}\psi$ be defined as above for $w_m > 0$. Let $\theta, \tilde\theta \geq 0$ be integers and let $(\phi_1, \dots, \phi_{m-1})$ and $(\tilde\phi_1, \dots, \tilde\phi_{m-1})$ denote multi-indices. Let $\ell := (\theta + |\phi| + 3|\tilde\phi| + 2\tilde\theta)$. Then*

$$\left|\frac{\partial^{\theta+|\phi|+|\tilde\phi|}}{\partial w_m^\theta \partial w^\phi \partial y^{\tilde\phi}}\big((w\cdot\nabla_y - \partial_t)^{\tilde\theta}(\mathcal{P}_{\gamma w_m}\psi(w,y,t))\big)\right| \leq c(m,l)\gamma^{1-(l-\theta)} w_m^{1-l} M_1 \tag{2-10}$$

whenever $(W,Y,t) \in U$.

Proof. We first consider the case $\theta = 1,\ \phi = 0,\ \tilde\phi = 0,\ \tilde\theta = 0$. In this case, simply using that $\mathcal{P}_{\gamma w_m}$ is an approximation of the identity operator, we see that (1-25) immediately implies

$$\left|\frac{\partial}{\partial w_m}(\mathcal{P}_{\gamma w_m}\psi(w,y,t))\right| \leq c(m)\gamma M_1. \tag{2-11}$$

By similar considerations we have

$$\begin{aligned} w_m^{\theta-1}\left|\frac{\partial^\theta}{\partial w_m^\theta}(\mathcal{P}_{\gamma w_m}\psi(w,y,t))\right| &\leq c(m,l)\gamma M_1, \\ w_m^{|\phi|-1}\left|\frac{\partial^{|\phi|}}{\partial w^\phi}(\mathcal{P}_{\gamma w_m}\psi(w,y,t))\right| &\leq c(m,l)\gamma^{1-|\phi|} M_1, \\ w_m^{3|\tilde\phi|-1}\left|\frac{\partial^{|\tilde\phi|}}{\partial y^{\tilde\phi}}(\mathcal{P}_{\gamma w_m}\psi(w,y,t))\right| &\leq c(m,l)\gamma^{1-3|\tilde\phi|} M_1 \end{aligned} \tag{2-12}$$

whenever $\theta \geq 1,\ |\phi| \geq 1,\ |\tilde\phi| \geq 1$. Furthermore,

$$(\gamma w_m)(w\cdot\nabla_y - \partial_t)(\mathcal{P}_{\gamma w_m}\psi(w,y,t)) = (\gamma w_m)^{-1}(w\cdot\nabla_y\mathcal{P} - \partial_t\mathcal{P})_{\gamma w_m}\psi(w,y,t), \tag{2-13}$$

and hence, again arguing as above, we can conclude that

$$w_m^{2|\tilde\theta|-1}\left|(w\cdot\nabla_y - \partial_t)^{\tilde\theta}(\mathcal{P}_{\gamma w_m}\psi(w,y,t))\right| \leq c(m,l)\gamma^{1-2\tilde\theta} M_1. \tag{2-14}$$

Combining the above, the lemma follows. □

Lemma 2.2. *Let ψ be a function satisfying (1-25) and (1-26) for some constants $0 < M_1, M_2 < \infty$, let $\gamma \in (0,1)$ and let $P_{\gamma w_m}\psi$ be defined as above for $w_m > 0$. Let $\theta, \tilde\theta \geq 0$ be integers and let $(\phi_1, \dots, \phi_{m-1})$ and $(\tilde\phi_1, \dots, \tilde\phi_{m-1})$ denote multi-indices. Let $\ell := (\theta + |\phi| + 3|\tilde\phi| + 2\tilde\theta)$. Let*

$$d\mu = d\mu(W,Y,t) := \left|\frac{\partial^{\theta+|\phi|+|\tilde\phi|}}{\partial w_m^\theta \partial w^\phi \partial y^{\tilde\phi}}\big((w\cdot\nabla_y - \partial_t)^{\tilde\theta}(\mathcal{P}_{\gamma w_m}\psi(w,y,t))\big)\right|^2 w_m^{2l-3}\, dW\, dy\, dt, \tag{2-15}$$

defined on U. Then

$$\mu(U \cap \mathcal{B}_r) \leq c(m, l, M_1, M_2)\gamma^{2-2(l-\theta)} r^{\boldsymbol{q}-1}$$

for all balls $\mathcal{B}_r = \mathcal{B}_r(Z_0, t_0) \subset \mathbb{R}^{N+1}$ centered on ∂U, $r > 0$.

Proof. As in the proof of Lemma 2.1, we first consider the case $\theta = 1$, $\phi = 0$, $\tilde{\phi} = 0$, $\tilde{\theta} = 0$. Then

$$\frac{\partial}{\partial w_m}(\mathcal{P}_{\gamma w_m})(w, y, t) = \frac{1}{w_m}(\mathcal{Q}_{\gamma w_m})(w, y, t), \tag{2-16}$$

where $\mathcal{Q}$ is such that $\int w_i \mathcal{Q}_{\gamma w_m}(w, y, t)\, dw\, dy\, dt = 0$ for all $i \in \{1, \dots, m-1\}$. Let

$$l^{\psi}_{(w,y,t)}(\bar{w}, \bar{y}, \bar{t}) = \psi(\bar{w}, \bar{y}, \bar{t}) - \psi(w, y, t) - \mathcal{P}_{w_m}(\nabla_w \psi)(w, y, t)(\bar{w} - w). \tag{2-17}$$

Then,

$$\begin{aligned}
\left|\frac{\partial}{\partial w_m}(\mathcal{P}_{\gamma w_m}\psi)(w, y, t)\right| &= \frac{1}{w_m}\left|\int_{\mathbb{R}^{N-1}} \psi(\bar{w}, \bar{y}, \bar{t})\, \mathcal{Q}_{\gamma w_m}((\bar{w}, \bar{y}, \bar{t})^{-1} \circ (w, y, t))\, d\bar{w}\, d\bar{y}\, d\bar{t}\right| \\
&\leq \frac{1}{w_m}\left|\int_{\mathbb{R}^{N-1}} (l^{\psi}_{(w,y,t)}(\bar{w}, \bar{y}, \bar{t}))\, \mathcal{Q}_{\gamma w_m}((\bar{w}, \bar{y}, \bar{t})^{-1} \circ (w, y, t))\, d\bar{w}\, d\bar{y}\, d\bar{t}\right| \\
&\leq c\gamma\gamma_\psi(z, t, c w_m)
\end{aligned} \tag{2-18}$$

for some $c = c(m)$, $1 \leq c < \infty$. Hence, using (1-26) we have

$$\begin{aligned}
\iint_{U \cap \mathcal{B}_r} \left|\frac{\partial}{\partial w_m}(\mathcal{P}_{\gamma w_m}\psi)(w, y, t)\right|^2 w_m^{-1}\, dW\, dy\, dt &\leq c\gamma^2 \iint_{U \cap \mathcal{B}_r} (\gamma_\psi(w, y, t, c w_m))^2 w_m^{-1}\, dW\, dy\, dt \\
&\leq c M_2 \gamma^2 r^{\boldsymbol{q}-1}
\end{aligned} \tag{2-19}$$

for all balls $\mathcal{B}_r \subset \mathbb{R}^{N+1}$ centered on ∂U, $r > 0$. By similar considerations, using (1-26), we have

$$\begin{aligned}
\iint_{U \cap \mathcal{B}_r} \left|\frac{\partial^\theta}{\partial w_m^\theta}(\mathcal{P}_{\gamma w_m}\psi(w, y, t))\right|^2 w_m^{2\theta-3}\, dW\, dy\, dt &\leq c(m, l)\gamma^2 r^{\boldsymbol{q}-1}, \\
\iint_{U \cap \mathcal{B}_r} \left|\frac{\partial^{|\phi|}}{\partial w^\phi}(\mathcal{P}_{\gamma w_m}\psi(w, y, t))\right|^2 w_m^{2|\phi|-3}\, dW\, dy\, dt &\leq c(m, l)\gamma^{2-2|\phi|} r^{\boldsymbol{q}-1}, \\
\iint_{U \cap \mathcal{B}_r} \left|\frac{\partial^{|\tilde{\phi}|}}{\partial y^{\tilde{\phi}}}(\mathcal{P}_{\gamma w_m}\psi(w, y, t))\right|^2 w_m^{6|\tilde{\phi}|-3}\, dW\, dy\, dt &\leq c(m, l)\gamma^{2-6|\tilde{\phi}|} r^{\boldsymbol{q}-1}, \\
\iint_{U \cap \mathcal{B}_r} \left|(w \cdot \nabla_y - \partial_t)^{\tilde{\theta}}(\mathcal{P}_{\gamma w_m}\psi(w, y, t))\right|^2 w_m^{4|\tilde{\theta}|-3}\, dW\, dy\, dt &\leq c(m, l)\gamma^{2-4\tilde{\theta}} r^{\boldsymbol{q}-1}
\end{aligned} \tag{2-20}$$

for all balls $\mathcal{B}_r \subset \mathbb{R}^{N+1}$ centered on ∂U, $r > 0$, whenever $\theta \geq 1$, $|\phi| \geq 2$, $|\tilde{\phi}| \geq 1$, $\tilde{\theta} \geq 1$. Combining the above, the lemma follows. □

Remark 2.3. Using Lemma 2.1 we see that there exists $\hat{\gamma} = \hat{\gamma}(m, M_1) \in (0, 1)$ such that if $\gamma \in (0, \hat{\gamma})$ then

$$\tfrac{1}{2} \leq 1 + \frac{\partial}{\partial w_m}(\mathcal{P}_{\gamma w_m}\psi)(w, y, t) \leq \tfrac{3}{2}$$

whenever $(w, w_m, y, y_m, t) \in U$. This implies, in particular, that the map $(w, w_m, y, y_m, t) \in U \to (w, w_m + P_{\gamma w_m}\psi(w, y, t), y, y_m, t)$ is one-to-one.

2C. *A Poincaré inequality.* We introduce the open cube

$$\mathcal{Q}_r(0,0) = \big\{(Z,t)=(X,Y,t) \in \mathbb{R}^{N+1} \;\big|\; |x_i| < r,\ |y_i| < r^3,\ |t| < r^2\big\}, \tag{2-21}$$

where $i \in \{1, \dots, m\}$. Given $(Z_0, t_0) \in \mathbb{R}^{N+1}$, we let $\mathcal{Q}_r(Z_0, t_0) = (Z_0, t_0) \circ \mathcal{Q}_r(0, 0)$. We will need the following Poincaré inequality.

Lemma 2.4. *Consider $\mathcal{Q}_r := \mathcal{Q}_r(Z_0, t_0) \subset \mathbb{R}^{N+1}$ and let p, $1 < p < \infty$, be given. Let u be a (smooth) function defined on $\mathcal{Q}_r$ and let E denote the mean value of u on $\mathcal{Q}_r$. Then there exists a constant $c = c(N, p)$, $1 \le c < \infty$, such that*

$$\iint_{\mathcal{Q}_r} |u - E|^p \, dZ\, dt \le c \iint_{\mathcal{Q}_r} \big(r^p |\nabla_X u|^p + r^{3p}|\nabla_Y u|^p + r^{2p}|X_0(u)|^p\big)\, dZ\, dt.$$

Proof. Assume first that $(Z_0, t_0) = (0, 0)$ and let $\tilde{u}$ be a (smooth) function defined on $\mathcal{Q}_r(0, 0)$. Then, using the mean value theorem and arguing, for example, as in the proof of Lemma 6.12 in [Lieberman 1996], we see that

$$\begin{aligned}\iint_{\mathcal{Q}_r(0,0)} &|\tilde{u} - E(\tilde{u}, \mathcal{Q}_r(0,0))|^p \, dZ\, dt \\ &\le c(N, p) \iint_{\mathcal{Q}_r(0,0)} \big(r^p|\nabla_X \tilde{u}|^p + r^{3p}|\nabla_Y \tilde{u}|^p + r^{2p}|\partial_t \tilde{u}|^p\big)\, dZ\, dt,\end{aligned} \tag{2-22}$$

where $E(\tilde{u}, \mathcal{Q}_r(0, 0))$ denotes the mean value of $\tilde{u}$ on $\mathcal{Q}_r(0, 0)$. Next, consider a function u defined on $\mathcal{Q}_r(Z_0, t_0)$ for some $(Z_0, t_0) \neq (0, 0)$. Let $\tilde{u}(Z, t) = u((Z_0, t_0) \circ (Z, t))$. Then $\tilde{u}$ is a function defined on $\mathcal{Q}_r(Z_0, t_0)$, $E(\tilde{u}, \mathcal{Q}_r(0, 0)) = E(u, \mathcal{Q}_r(Z_0, t_0))$ and (2-22) applies to $\tilde{u}$. Applying (2-22) to $\tilde{u}$ and expressing the result in terms of u the conclusion of the lemma follows. □

2D. *Interior regularity.*

Lemma 2.5. *Assume that $\mathcal{K}u = 0$ in $\mathcal{B}_{2r} = \mathcal{B}_{2r}(Z_0, t_0) \subset \mathbb{R}^{N+1}$. Then there exists a constant $c = c(N)$, $1 \le c < \infty$, such that*

(i) $r^{\mathbf{q}}(\sup_{\mathcal{B}_r} |u|)^2 \le c \iint_{\mathcal{B}_{2r}} |u|^2 \, dZ\, dt,$

(ii) $\iint_{\mathcal{B}_r} |\nabla_X u|^2 \, dZ\, dt \le \dfrac{c}{r^2} \iint_{\mathcal{B}_{2r}} u^2 \, dZ\, dt,$

(iii) $\sup_{\mathcal{B}_r} \big(r|\nabla_X u| + r^3|\nabla_Y u| + r^2|X_0(u)|\big) \le c \sup_{\mathcal{B}_{2r}} |u|.$

Proof. For (i) and (iii) we refer to [Lanconelli and Polidoro 1994]; (ii) is an energy estimate which can be proved by standard arguments. □

Lemma 2.6. *Assume that $\mathcal{K}u = 0$ in $\mathcal{B}_{2r} = \mathcal{B}_{2r}(Z_0, t_0) \subset \mathbb{R}^{N+1}$. Let $\zeta \in C_0^\infty(\mathcal{B}_{2r})$ be such that $0 \le \zeta \le 1$, $\zeta \equiv 1$ on $\mathcal{B}_r$, and such that $r|\nabla_X \zeta| + r^3|\nabla_Y \zeta| + r^2|X_0(\zeta)| \le c(N)$. Let $i \in \{1, \dots, m\}$. Then there exists a constant $c = c(N)$, $1 \le c < \infty$, such that*

(i) $\displaystyle\iint_{\mathcal{B}_{2r}} |\nabla_X(\partial_{y_i}u)|^2\zeta^6\,dZ\,dt \le \frac{c}{r^2}\iint_{\mathcal{B}_{2r}} |\partial_{y_i}u|^2\zeta^4\,dZ\,dt,$

(ii) $\displaystyle\iint_{\mathcal{B}_{2r}} |\partial_{y_i}u|^2\zeta^4\,dZ\,dt \le \frac{c}{r^2}\iint_{\mathcal{B}_{2r}} |X_0(u)|^2\zeta^2\,dZ\,dt + \frac{c}{r^4}\iint_{\mathcal{B}_{2r}} |\nabla_X u|^2\,dZ\,dt,$

(iii) $\displaystyle\iint_{\mathcal{B}_{2r}} |X_0(u)|^2\zeta^2\,dZ\,dt \le \frac{c}{r^2}\iint_{\mathcal{B}_{2r}} |\nabla_X u|^2\,dZ\,dt.$

Proof. Let

$$A_1=\iint_{\mathcal{B}_{2r}} |\nabla_X(\partial_{y_i}u)|^2\zeta^6\,dZ\,dt,\quad A_2=\iint_{\mathcal{B}_{2r}} |\partial_{y_i}u|^2\zeta^4\,dZ\,dt,\quad A_3=\iint_{\mathcal{B}_{2r}} |X_0(u)|^2\zeta^2\,dZ\,dt. \tag{2-23}$$

As $\tilde u := \partial_{y_i}u$ solves $\mathcal{K}\tilde u=0$, we see that (i) follows immediately from Lemma 2.5(ii) and its proof. To prove (ii) we first note, integrating by parts,

$$\begin{aligned}
A_2 &= \iint_{\mathcal{B}_{2r}} (\partial_{y_i}u)\big(X_0(\partial_{x_i}u)-\partial_{x_i}(X_0(u)\big)\zeta^4\,dZ\,dt\\
&= -\iint_{\mathcal{B}_{2r}} X_0(\partial_{y_i}u)(\partial_{x_i}u)\zeta^4\,dZ\,dt - 4\iint_{\mathcal{B}_{2r}} (\partial_{y_i}u)(\partial_{x_i}u)\zeta^3 X_0(\zeta)\,dZ\,dt\\
&\qquad + \iint_{\mathcal{B}_{2r}} (\partial_{x_i y_i}u)X_0(u)\zeta^4\,dZ\,dt + 4\iint_{\mathcal{B}_{2r}} (\partial_{y_i}u)X_0(u)\zeta^3\partial_{x_i}(\zeta)\,dZ\,dt.
\end{aligned} \tag{2-24}$$

Next, writing $X_0(\partial_{y_i}u)=\partial_{y_i}(X_0(u))$ and integrating by parts in the first term, we see that

$$\begin{aligned}
A_2 &= \iint_{\mathcal{B}_{2r}} X_0(u)\partial_{y_i x_i}u\zeta^4\,dZ\,dt + 4\iint_{\mathcal{B}_{2r}} X_0(u)\partial_{x_i}u\zeta^3\partial_{y_i}\zeta\,dZ\,dt - 4\iint_{\mathcal{B}_{2r}} (\partial_{y_i}u)(\partial_{x_i}u)\zeta^3 X_0(\zeta)\,dZ\,dt\\
&\qquad + \iint_{\mathcal{B}_{2r}} (\partial_{x_i y_i}u)X_0(u)\zeta^4\,dZ\,dt + 4\iint_{\mathcal{B}_{2r}} (\partial_{y_i}u)X_0(u)\zeta^3\partial_{x_i}(\zeta)\,dZ\,dt.
\end{aligned} \tag{2-25}$$

Using this we see that

$$\begin{aligned}
A_2 &\le \iint_{\mathcal{B}_{2r}} |\nabla_X(\partial_{y_i}u)||X_0(u)|\zeta^4\,dZ\,dt + \frac{c}{r^3}\iint_{\mathcal{B}_{2r}} |X_0(u)||\nabla_X u|\zeta^3\,dZ\,dt\\
&\qquad + \frac{c}{r^2}\iint_{\mathcal{B}_{2r}} |\partial_{y_i}u||\nabla_X u|\zeta^3\,dZ\,dt + \frac{c}{r}\iint_{\mathcal{B}_{2r}} |\partial_{y_i}u||X_0(u)|\zeta^3\,dZ\,dt.
\end{aligned} \tag{2-26}$$

Hence, using Cauchy–Schwarz we see that

$$A_2 \le \epsilon r^2 A_1 + \tilde\epsilon A_2 + c(\epsilon,\tilde\epsilon,n)\bigg(\frac{1}{r^2}A_3 + \frac{1}{r^4}\iint_{\mathcal{B}_{2r}} |\nabla_X u|^2\,dZ\,dt\bigg), \tag{2-27}$$

where $\epsilon>0$ and $\tilde\epsilon>0$ are degrees of freedom. Furthermore, using the conclusion established in (i) we see that

$$A_2 \le c\epsilon A_2 + \tilde\epsilon A_2 + \tilde c(\epsilon,\tilde\epsilon,n)\bigg(\frac{1}{r^2}A_3 + \frac{1}{r^4}\iint_{\mathcal{B}_{2r}} |\nabla_X u|^2\,dZ\,dt\bigg). \tag{2-28}$$

Part (ii) now follows by elementary manipulations. To prove (iii) we use the equation $\mathcal{K}u=0$ and write

$$A_3 = -\sum_{i=1}^m \iint_{\mathcal{B}_{2r}} X_0(u)(\partial_{x_i x_i}u)\zeta^2\,dZ\,dt = A_{31}+A_{32}+A_{33}, \tag{2-29}$$

where

$$
\begin{aligned}
A_{31} &:= 2\sum_{i=1}^{m}\iint_{\mathcal{B}_{2r}} X_0(u)(\partial_{x_i}u)\zeta\,\partial_{x_i}(\zeta)\,dZ\,dt,\\
A_{32} &:= \sum_{i=1}^{m}\iint_{\mathcal{B}_{2r}} X_0(\partial_{x_i}u)(\partial_{x_i}u)\zeta^2\,dZ\,dt,\\
A_{33} &:= \sum_{i=1}^{m}\iint_{\mathcal{B}_{2r}} (\partial_{y_i}u)(\partial_{x_i}u)\zeta^2\,dZ\,dt.
\end{aligned}
\tag{2-30}
$$

Then

$$
|A_{31}|+|A_{33}| \le \epsilon A_3 + \tilde{\epsilon} r^2 A_1 + \frac{c(\epsilon,\tilde{\epsilon})}{r^2}\iint_{\mathcal{B}_{2r}} |\nabla_X u|^2\,dZ\,dt, \tag{2-31}
$$

where $\epsilon > 0$ and $\tilde{\epsilon} > 0$ are degrees of freedom. To handle A_{32} we simply note, lifting the vector field X_0 by partial integration, that

$$
2A_{32} = -2\sum_{i=1}^{m}\iint_{\mathcal{B}_{2r}} |\partial_{x_i}u|^2\zeta X_0(\zeta)\,dZ\,dt. \tag{2-32}
$$

Hence,

$$
A_3 \le \epsilon A_3 + \tilde{\epsilon} r^2 A_1 + \frac{c(\epsilon,\tilde{\epsilon})}{r^2}\iint_{\mathcal{B}_{2r}} |\nabla_X u|^2\,dZ\,dt. \tag{2-33}
$$

Combining (2-33) and (i), (ii) of the lemma, we see that (iii) follows. □

Remark 2.7. To construct ζ as in the statement of Lemma 2.6, simply choose $\zeta(Z,t) := \tilde{\zeta}((Z_0,t_0)\circ(Z,t))$, where $\tilde{\zeta} \in C_0^\infty(\mathcal{B}_{2r}(0,0))$ is such that $0 \le \tilde{\zeta} \le 1$, $\tilde{\zeta} \equiv 1$ on $\mathcal{B}_r(0,0)$, and such that

$$
r|\nabla_X\tilde{\zeta}| + r^3|\nabla_Y\tilde{\zeta}| + r^2|\partial_t\tilde{\zeta}| \le c(N).
$$

We can construct $\tilde{\zeta}$ in a standard manner by smoothing out the indicator function of say $\mathcal{B}_{3r/2}(0,0)$.

Lemma 2.8. *Assume that $\mathcal{K}u = 0$ in $\mathcal{B}_{2r} = \mathcal{B}_{2r}(Z_0,t_0) \subset \mathbb{R}^{N+1}$. Let $i \in \{1,\dots,m\}$. Then there exists a constant $c = c(N)$, $1 \le c < \infty$, such that*

$$
\iint_{\mathcal{B}_r} \big(r^4|\nabla_X(\partial_{y_i}u)|^2 + r^2|\nabla_Y u|^2 + |X_0(u)|^2\big)\,dZ\,dt \le \frac{c}{r^2}\iint_{\mathcal{B}_{2r}} |\nabla_X u|^2\,dZ\,dt.
$$

Proof. The lemma is an immediate consequence of Lemma 2.6. □

3. The Dirichlet problem

Let $\Omega = \Omega_\psi \subset \mathbb{R}^{N+1}$ be an unbounded Lip_K-domain in the sense of Definition 1.1. We consider here the well-posedness of the boundary value problem

$$
\begin{cases} \mathcal{K}u = 0 & \text{in } \Omega,\\ \;\;u = \varphi & \text{on } \partial\Omega. \end{cases} \tag{3-1}
$$

Note that we can without loss of generality assume that $\psi(0,0,0)=0$ and hence that $(0,0,0,0,0)\in\partial\Omega$. To conform with the notation used in [Nyström and Polidoro 2016] we let

$$\Omega_r := \Omega_{\psi,r} := \bigl\{(X,Y,t)\bigm| |x_i|<r^2,\ |y_i|<r^3,\ |t|<2r^2,\ |y_m|<r^3,\ \psi(x,y,t)<x_m<4M_1r\bigr\} \tag{3-2}$$

for $r>0$ and where $i\in\{1,\dots,m-1\}$. As outlined in Subsection 2.4 of that paper, using the Perron–Wiener–Brelot method, the existence of a solution to the problem in (3-1) with Ω replaced by Ω_r can be established. In Definition 3 of the same paper, we introduced what we here refer to as the Kolmogorov boundary of Ω_r, denoted by $\partial_K\Omega_r$. The notion of the Kolmogorov boundary replaces the notion of the parabolic boundary used in the context of uniformly parabolic equations and by definition $\partial_K\Omega_r\subset\partial\Omega_r$ is the set of all points on the topological boundary of Ω_r, which is contained in the closure of the propagation of at least one interior point in Ω_r. The importance of the Kolmogorov boundary of Ω_r is highlighted in the following lemma; see Lemma 2.2 in [Nyström and Polidoro 2016].

Lemma 3.1. *Consider the Dirichlet problem in (3-1), with Ω replaced by Ω_r, with boundary data $\varphi\in C(\partial\Omega_r)$ and let $u=u_\varphi$ be the corresponding Perron–Wiener–Brelot solution. Then*

$$\sup_{\Omega_r}|u|\le\sup_{\partial_K\Omega_r}|\varphi|.$$

In particular, if $\varphi\equiv0$ on $\partial_K\Omega_r$ then $u\equiv0$ in Ω_r.

The set $\partial_K\Omega_r$ is the largest subset of the topological boundary of Ω_r on which we can attempt to impose boundary data if we want to construct nontrivial solutions to the Dirichlet problem in (3-1), with Ω replaced by Ω_r. The notion of regular points on $\partial\Omega_r$ for the Dirichlet problem only makes sense for points on the Kolmogorov boundary and we let $\partial_R\Omega_r$ be the set of all $(z_0,t_0)\in\partial_K\Omega_r$ such that

$$\lim_{(Z,t)\to(Z_0,t_0)} u_\varphi(Z,t)=\varphi(Z_0,t_0)\quad\text{for any }\varphi\in C(\partial\Omega_r). \tag{3-3}$$

We refer to $\partial_R\Omega_r$ as the regular boundary of Ω_r with respect to the operator $\mathcal{K}$. By definition, $\partial_R\Omega_r\subseteq\partial_K\Omega_r$.

Lemma 3.2 [Nyström and Polidoro 2016, Lemma 2.2]. *Let $\Omega\subset\mathbb{R}^{N+1}$ be a Lip_K-domain with constant M_1 and let Ω_r be as defined in (3-2). Then*

$$\partial_R\Omega_r=\partial_K\Omega_r;$$

i.e., all points on the Kolmogorov boundary of Ω_r are regular for the operator $\mathcal{K}$.

Lemma 3.3. *Let $\Omega\subset\mathbb{R}^{N+1}$ be a Lip_K-domain with constant M_1, consider the Dirichlet problem in (3-1) and assume that $\varphi\in C(\partial\Omega)\cap L^\infty(\partial\Omega)$ is such that $\varphi(Z,t)\to0$ as $\|(Z,t)\|\to\infty$. Then there exists a unique solution to the Dirichlet problem in (3-1) in Ω such that $u\in C(\bar\Omega)$, $u=\varphi$ on $\partial\Omega$. Furthermore, $\|u\|_{L^\infty(\Omega)}\le\|\varphi\|_{L^\infty(\Omega)}$.*

Proof. This can be proved by exhausting Ω with the bounded domains Ω_{r_j}, $j\in\mathbb{Z}_+$, $r_j=j$, for example, and by constructing u as the limit of $\{u_j\}$, where $\mathcal{K}u_j=0$ in Ω_{r_j} and with u_j having appropriate data on $\partial\Omega_{r_j}$. We here omit the routine details. □

Remark 3.4. The operator adjoint to $\mathcal{K}$ is

$$\mathcal{K}^* = \sum_{i=1}^m \partial_{x_i x_i} - \sum_{i=1}^m x_i \partial_{y_i} + \partial_t. \tag{3-4}$$

In the case of the adjoint operator $\mathcal{K}^*$ we denote the associated Kolmogorov boundary of Ω_r by $\partial_K^* \Omega_r$. The above discussion, lemmas and Lemma 3.2 then apply to $\mathcal{K}^*$ subject to the natural modifications.

Lemma 3.5. *Let $\Omega \subset \mathbb{R}^{N+1}$ be a* Lip_K*-domain with constant M_1. Let $\varphi \in C(\partial\Omega) \cap L^\infty(\partial\Omega)$ be such that $\varphi(Z,t) \to 0$ as $\|(Z,t)\| \to \infty$. Then there exist unique solutions $u = u_\varphi$, $u \in C(\overline{\Omega})$, and $u^* = u_{\varphi^*}$, $u^* \in C(\overline{\Omega})$, to the Dirichlet problem in (3-1) and to the corresponding Dirichlet problem for $\mathcal{K}^*$, respectively. Moreover, there exist, for every $(Z,t) \in \Omega$, unique probability measures $\omega(Z,t,\cdot)$ and $\omega^*(Z,t,\cdot)$ on $\partial\Omega$ such that*

$$u(Z,t) = \int_{\partial\Omega} \varphi(\widetilde{Z},\tilde{t})\, d\omega(Z,t,\widetilde{Z},\tilde{t}), \quad u^*(Z,t) = \int_{\partial\Omega} \varphi^*(\widetilde{Z},\tilde{t})\, d\omega^*(Z,t,\widetilde{Z},\tilde{t}). \tag{3-5}$$

Proof. The lemma is an immediate consequence of Lemma 3.2. □

Definition 3.6. Let $(Z,t) \in \Omega$. Then $\omega(Z,t,\cdot)$ is referred to as the Kolmogorov measure relative to (Z,t) and Ω, and $\omega^*(Z,t,\cdot)$ is referred to as the adjoint Kolmogorov measure relative to (Z,t) and Ω.

3A. ***The fundamental solution and the Green function.*** Following [Kolmogorov 1934] and [Lanconelli and Polidoro 1994], it is well known that an explicit fundamental solution, Γ, associated to $\mathcal{K}$ can be constructed. Indeed, let

$$B := \begin{pmatrix} 0 & I_m \\ 0 & 0 \end{pmatrix}, \qquad E(s) = \exp(-sB^*)$$

for $s \in \mathbb{R}$, where I_m, 0, represent the identity matrix and the zero matrix in $\mathbb{R}^m$, respectively. Here * denotes the transpose. Furthermore, let

$$\mathcal{C}(t) := \int_0^t E(s) \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix} E^*(s)\, ds = \begin{pmatrix} tI_m & -\frac{1}{2}t^2 I_m \\ -\frac{1}{2}t^2 I_m & \frac{1}{3}t^3 I_m \end{pmatrix}$$

whenever $t \in \mathbb{R}$. Note that $\det \mathcal{C}(t) = \frac{1}{12}t^{4m}$ and that

$$(\mathcal{C}(t))^{-1} = 12 \begin{pmatrix} \frac{1}{3}t^{-1} I_m & \frac{1}{2}t^{-2} I_m \\ \frac{1}{2}t^{-2} I_m & t^{-3} I_m \end{pmatrix}.$$

Using this notation, a fundamental solution, with pole at $(\widetilde{Z},\tilde{t})$, $\Gamma(\cdot,\cdot,\widetilde{Z},\tilde{t})$, can be defined by

$$\Gamma(Z,t,\widetilde{Z},\tilde{t}) = \Gamma\big(Z - E(t-\tilde{t})\widetilde{Z},\, t-\tilde{t},\, 0,\, 0\big), \tag{3-6}$$

where $\Gamma(Z,t,0,0) = 0$ if $t \le 0$, $Z \ne 0$, and

$$\Gamma(Z,t,0,0) = \frac{(4\pi)^{-N/2}}{\sqrt{\det \mathcal{C}(t)}} \exp\big(-\tfrac{1}{4}\langle \mathcal{C}(t)^{-1} Z, Z\rangle\big) \quad \text{if } t > 0. \tag{3-7}$$

Here $\langle\cdot,\cdot\rangle$ denotes the standard inner product on $\mathbb{R}^N$. We also note that

$$\Gamma(Z,t,\widetilde{Z},\tilde{t}) \le \frac{c(N)}{\|(\widetilde{Z},\tilde{t})^{-1}\circ(Z,t)\|^{\boldsymbol{q}-2}} \quad \text{for all } (Z,t), (\widetilde{Z},\tilde{t}) \in \mathbb{R}^{N+1},\ t > \tilde{t}. \tag{3-8}$$

We define the Green function for Ω, with pole at $(\widehat{Z},\hat{t}) \in \Omega$, as

$$G(Z,t,\widehat{Z},\hat{t}) = \Gamma(Z,t,\widehat{Z},\hat{t}) - \int_{\partial\Omega} \Gamma(\widetilde{Z},\tilde{t},\widehat{Z},\hat{t})\,d\omega(Z,t,\widetilde{Z},\tilde{t}), \tag{3-9}$$

where Γ is the fundamental solution to the operator $\mathcal{K}$ introduced in (3-6). If we instead consider $(Z,t)\in\Omega$ as fixed, then, for $(\widehat{Z},\hat{t}) \in \Omega$,

$$G(Z,t,\widehat{Z},\hat{t}) = \Gamma(Z,t,\widehat{Z},\hat{t}) - \int_{\partial\Omega} \Gamma(Z,t,\widetilde{Z},\tilde{t})\,d\omega^*(\widehat{Z},\hat{t},\widetilde{Z},\tilde{t}), \tag{3-10}$$

where $\omega^*(\widehat{Z},\hat{t},\cdot)$ is the associated adjoint Kolmogorov measure relative to $(\widehat{Z},\hat{t})$ and Ω. Given $\theta \in C_0^\infty(\mathbb{R}^{N+1})$, we have the representation formulas

$$\begin{aligned} \theta(Z,t) &= \int_{\partial\Omega} \theta(\widetilde{Z},\tilde{t})\,d\omega(Z,t,\widetilde{Z},\tilde{t}) + \int G(Z,t,\widehat{Z},\hat{t})\,\mathcal{K}\theta(\widehat{Z},\hat{t})\,d\widehat{Z}\,d\hat{t},\\ \theta(\widehat{Z},\hat{t}) &= \int_{\partial\Omega} \theta(\widetilde{Z},\tilde{t})\,d\omega^*(\widehat{Z},\hat{t},\widetilde{Z},\tilde{t}) + \int G(Z,t,\widehat{Z},\hat{t})\,\mathcal{K}^*\theta(Z,t)\,dZ\,dt \end{aligned} \tag{3-11}$$

whenever $(Z,t), (\widehat{Z},\hat{t}) \in \Omega$. In particular,

$$\begin{aligned} \int G(Z,t,\widehat{Z},\hat{t})\,\mathcal{K}\theta(\widehat{Z},\hat{t})\,d\widehat{Z}\,d\hat{t} &= -\int \theta(\widetilde{Z},\tilde{t})\,d\omega(Z,t,\widetilde{Z},\tilde{t}),\\ \int G(Z,t,\widehat{Z},\hat{t})\,\mathcal{K}^*\theta(Z,t)\,dZ\,dt &= -\int \theta(\widetilde{Z},\tilde{t})\,d\omega^*(\hat{x},\hat{t},\widetilde{Z},\tilde{t}) \end{aligned} \tag{3-12}$$

whenever $\theta \in C_0^\infty(\mathbb{R}^{N+1}\setminus\{(Z,t)\})$ and $\theta \in C_0^\infty(\mathbb{R}^{N+1}\setminus\{(\widehat{Z},\hat{t})\})$, respectively.

Remark 3.7. Recall that $\boldsymbol{q} = 4m+2$. However, we note that in [Nyström and Polidoro 2016] a different definition of $\boldsymbol{q}$ ($\boldsymbol{q} = 4m$) was used. Hence, in this paper some statements containing $\boldsymbol{q}$ differ slightly compared to the corresponding statements in that paper.

4. Estimates for nonnegative solutions

In this section we develop and state a number of estimates concerning nonnegative solutions, the Kolmogorov measure as well as the kernel function. Throughout this section we assume that $\Omega = \Omega_\psi \subset \mathbb{R}^{N+1}$ is a Lip_K-domain, with constant M_1, in the sense of Definition 1.1. Given $\varrho > 0$ and $\Lambda > 0$ we let

$$\begin{aligned} A^+_{\varrho,\Lambda} &= \big(\Lambda\varrho, 0, -\tfrac{2}{3}\Lambda\varrho^3, 0, \varrho^2\big) \in \mathbb{R}\times\mathbb{R}^{m-1}\times\mathbb{R}\times\mathbb{R}^{m-1}\times\mathbb{R},\\ A_{\varrho,\Lambda} &= (\Lambda\varrho, 0, 0, 0, 0) \in \mathbb{R}\times\mathbb{R}^{m-1}\times\mathbb{R}\times\mathbb{R}^{m-1}\times\mathbb{R},\\ A^-_{\varrho,\Lambda} &= \big(\Lambda\varrho, 0, \tfrac{2}{3}\Lambda\varrho^3, 0, -\varrho^2\big) \in \mathbb{R}\times\mathbb{R}^{m-1}\times\mathbb{R}\times\mathbb{R}^{m-1}\times\mathbb{R}. \end{aligned} \tag{4-1}$$

Given $(Z_0, t_0) \in \mathbb{R}^{N+1}$ we let

$$A^{\pm}_{\varrho,\Lambda}(Z_0, t_0) = (Z_0, t_0) \circ A^{\pm}_{\varrho,\Lambda}, \quad A_{\varrho,\Lambda}(Z_0, t_0) = (Z_0, t_0) \circ A_{\varrho,\Lambda}.$$

Furthermore, given $(Z_0, t_0) = (X_0, Y_0, t_0) = (x_0, \psi(x_0, y_0, t_0), Y_0, t_0) \in \partial\Omega_\psi$ and $r > 0$ we let $\Omega_r(Z_0, t_0) = \Omega_{\psi,r}(Z_0, t_0)$ be the set of all points $(X, Y, t) = (x, x_m, y, y_m, t)$ which satisfy the conditions

$$\begin{gathered} |x_i - x_{0,i}| < r, \quad |y_i - y_{0,i} + (t - t_0)x_{0,i}| < r^3 \quad \text{for } i \in \{1, \dots, m-1\}, \\ |t - t_0| < 2r^2, \quad \big|y_m - y_{0,m} + (t - t_0)\psi(x_0, y_0, t_0)\big| < r^3, \\ \psi(x, y, t) < x_m < 4M_1 r + \psi(x_0, y_0, t_0). \end{gathered} \tag{4-2}$$

Note that if we let $(\widetilde{X}, \widetilde{Y}, \tilde{t}) := (X_0, Y_0, t_0)^{-1} \circ (X, Y, t)$, and if we define

$$\tilde{\psi}(\tilde{x}, \tilde{y}, \tilde{t}) := \psi((x_0, y_0, t_0) \circ (\tilde{x}, \tilde{y}, \tilde{t})) - \psi(x_0, y_0, t_0),$$

then

$$\Omega_r(Z_0, t_0) = \Omega_{\psi,r}(Z_0, t_0) = \Omega_{\tilde{\psi},r}, \tag{4-3}$$

with $\Omega_{\tilde{\psi},r}$ defined as in (3-2). To be consistent with the notation used and the estimates proved in [Nyström and Polidoro 2016], we here simply note that there exists $c = c(N)$, $1 \le c < \infty$, such that

$$\Omega \cap \mathcal{B}_{r/c}(Z_0, t_0) \subset \Omega_r(Z_0, t_0) \subset \Omega \cap \mathcal{B}_{cr}(Z_0, t_0) \tag{4-4}$$

for all $(Z_0, t_0) \in \partial\Omega_\psi$, $r > 0$.

4A. ***The Harnack inequality.*** To formulate the Harnack inequality we first need to introduce some additional notation. We let, for $r > 0$ and $(Z_0, t_0) \in \mathbb{R}^{N+1}$,

$$Q^- = \big(B\big(\tfrac{1}{2}e_1, 1\big) \cap B\big(-\tfrac{1}{2}e_1, 1\big)\big) \times [-1, 0], \quad Q^-_r(Z_0, t_0) = (Z_0, t_0) \circ \delta_r(Q^-), \tag{4-5}$$

where e_1 is the unit vector pointing in the direction of x_m and $B\big(\tfrac{1}{2}e_1, 1\big)$ and $B\big(-\tfrac{1}{2}e_1, 1\big)$ are standard Euclidean balls of radius 1 in $\mathbb{R}^N$, centered at $\tfrac{1}{2}e_1$ and $-\tfrac{1}{2}e_1$, respectively. Similarly, we let

$$Q = \big(B\big(\tfrac{1}{2}e_1, 1\big) \cap B\big(-\tfrac{1}{2}e_1, 1\big)\big) \times [-1, 1], \quad Q_r(Z_0, t_0) = (Z_0, t_0) \circ \delta_r(Q). \tag{4-6}$$

Given $\alpha, \beta, \gamma, \theta \in \mathbb{R}$ such that $0 < \alpha < \beta < \gamma < \theta^2$, we set

$$\begin{aligned} \widetilde{Q}^+_r(Z_0, t_0) &= \big\{(x, t) \in Q^-_{\theta r}(Z_0, t_0) \;\big|\; t_0 - \alpha r^2 \le t \le t_0\big\}, \\ \widetilde{Q}^-_r(Z_0, t_0) &= \big\{(x, t) \in Q^-_{\theta r}(Z_0, t_0) \;\big|\; t_0 - \gamma r^2 \le t \le t_0 - \beta r^2\big\}. \end{aligned}$$

In the following we will formulate two versions of the Harnack inequality. The first version reads as follows and we refer to [Lanconelli and Polidoro 1994] for details and proofs.

Lemma 4.1. *There exist constants $c > 1$ and $\alpha, \beta, \gamma, \theta \in (0, 1)$, with $0 < \alpha < \beta < \gamma < \theta^2$, such that the following is true. Assume u is a nonnegative solution to $\mathcal{K}u = 0$ in $Q^-_r(Z_0, t_0)$ for some $r > 0$, $(Z_0, t_0) \in \mathbb{R}^{N+1}$. Then,*

$$\sup_{\widetilde{Q}^-_r(Z_0, t_0)} u \le c \inf_{\widetilde{Q}^+_r(Z_0, t_0)} u.$$

To formulate another version of the Harnack inequality we recall that the tool used to build Harnack chains is that of $\mathcal{K}$-admissible paths. A path $\gamma : [0, T] \to \mathbb{R}^{N+1}$ is called $\mathcal{K}$-admissible if it is absolutely continuous and satisfies

$$\frac{d}{d\tau}\gamma(\tau) = \sum_{j=1}^{m} \omega_j(\tau) X_j(\gamma(\tau)) + \lambda(\tau) X_0(\gamma(\tau)) \quad \text{for a.e. } \tau \in [0, T], \tag{4-7}$$

where $\omega_j \in L^2([0, T])$ for $j = 1, \dots, m$, and λ are nonnegative measurable functions. We say that γ connects $(Z, t) = (X, Y, t) \in \mathbb{R}^{N+1}$ to $(\widetilde{Z}, \tilde{t}) = (\widetilde{X}, \widetilde{Y}, \tilde{t}) \in \mathbb{R}^{N+1}$, $\tilde{t} < t$, if $\gamma(0) = (Z, t)$ and $\gamma(T) = (\widetilde{Z}, \tilde{t})$. When considering Kolmogorov operators in the domain $\mathbb{R}^N \times (T_0, T_1)$, it is well known that (2-3) implies the existence of a $\mathcal{K}$-admissible path γ for any points (Z, t), $(\widetilde{Z}, \tilde{t}) \in \mathbb{R}^{N+1}$ with $T_0 < \tilde{t} < t < T_1$. Given a domain $\Omega \subset \mathbb{R}^{N+1}$, and a point $(Z, t) \in \Omega$, we let $A_{(Z,t)} = A_{(Z,t)}(\Omega)$ denote the set

$$\left\{(\widetilde{Z}, \tilde{t}) \in \Omega \,\middle|\, \exists \text{ a } \mathcal{K}\text{-admissible } \gamma : [0, T] \to \Omega \text{ connecting } (Z, t) \text{ to } (\widetilde{Z}, \tilde{t})\right\},$$

and we define $\mathcal{A}_{(Z,t)} = \mathcal{A}_{(Z,t)}(\Omega) = \overline{A_{(Z,t)}(\Omega)}$. Here and in the sequel, $\mathcal{A}_{(Z,t)}(\Omega)$ is referred to as the propagation set of the point (Z, t) with respect to Ω. The presence of the drift term in $\mathcal{K}$ considerably changes the geometric structure of $A_{(Z,t)}(\Omega)$ and $\mathcal{A}_{(Z,t)}(\Omega)$ compared to the case of uniformly parabolic equations. The second version of the Harnack inequality reads as follows and we refer to [Cinti et al. 2010] for details and proofs.

Lemma 4.2. *Let $\Omega \subset \mathbb{R}^{N+1}$ be a domain and let $(Z_0, t_0) \in \Omega$. Let K be a compact set contained in the interior of $A_{(Z_0,t_0)}(\Omega)$. Then there exists a positive constant c_K, depending only on Ω and K, such that*

$$\sup_K u \le c_K\, u(Z_0, t_0)$$

for every nonnegative solution u of $\mathcal{K}u = 0$ in Ω.

Remark 4.3. We emphasize, and this is different compared to the case of uniform parabolic equations, that the constants $\alpha, \beta, \gamma, \theta$ in Lemma 4.1 cannot be arbitrarily chosen. In particular, according to Lemma 4.2, the cylinder $\widetilde{Q}_r^-(Z_0, t_0)$ has to be contained in the interior of the propagation set $\mathcal{A}_{(Z_0,t_0)}(Q_r^-(Z_0, t_0))$.

Several arguments in [Nyström and Polidoro 2016] involving the Harnack inequality explore that, by construction,

$$\gamma^+(\tau) = A^+_{(1-\tau)\varrho,\Lambda}(Z_0, t_0), \quad \gamma^-(\tau) = A^-_{(1-\tau)\varrho,\Lambda}(Z_0, t_0), \quad \tau \in [0, 1], \tag{4-8}$$

are $\mathcal{K}$-admissible paths; see Lemma 3.5–Lemma 3.8 in [Nyström and Polidoro 2016]. Here we state one of the results established in the same paper, which will be used in the forthcoming sections.

Lemma 4.4 [Nyström and Polidoro 2016, Lemma 3.9]. *Let $\Omega \subset \mathbb{R}^{N+1}$ be a Lip_K-domain with constant M_1. Then there exist $\Lambda = \Lambda(N, M_1)$, $1 \le \Lambda < \infty$, and $c = c(N, M_1)$, $1 \le c < \infty$, and $\gamma = \gamma(N, M_1)$, $0 < \gamma < \infty$, such that the following is true. Let $(Z_0, t_0) \in \partial\Omega$ and $r > 0$. Assume that u is a nonnegative solution to $\mathcal{K}u = 0$ in $\Omega \cap \mathcal{B}_{2r}(Z_0, t_0)$ and consider $\varrho, \tilde{\varrho}$, with $0 < \tilde{\varrho} \le \varrho < r/c$. Then*

$$\begin{aligned} u(A^+_{\tilde{\varrho},\Lambda}(Z_0, t_0)) &\le c(\varrho/\tilde{\varrho})^{\gamma} u(A^+_{\varrho,\Lambda}(Z_0, t_0)), \\ u(A^-_{\tilde{\varrho},\Lambda}(Z_0, t_0)) &\ge c^{-1}(\tilde{\varrho}/\varrho)^{\gamma} u(A^-_{\varrho,\Lambda}(Z_0, t_0)). \end{aligned} \tag{4-9}$$

Proof. Note that the lemma follows from the construction of Harnack chains along the paths in (4-8) and from Lemma 3.8 in [Nyström and Polidoro 2016]. For the details we refer to Lemma 3.9 in that paper and to Lemma 4.3 in [Cinti et al. 2013]. □

Remark 4.5. Let $\Omega \subset \mathbb{R}^{N+1}$ be a Lip_K-domain with constant M_1. The constants $\Lambda = \Lambda(N, M_1)$, $1 \le \Lambda < \infty$, and $c = c(N, M_1)$, $1 \le c < \infty$, referred to in Lemma 4.4, are fixed in Remark 3.7 in [Nyström and Polidoro 2016]. In particular, these constants are fixed so that the validity of Lemmas 3.5–Lemma 3.7 in that paper are ensured. In the following we also let Λ and c be determined accordingly.

4B. *Hölder continuity estimates and boundary comparison principles.*

Lemma 4.6. *Let $\Omega \subset \mathbb{R}^{N+1}$ be a Lip_K-domain with constant M_1. Let $(Z_0, t_0) \in \partial\Omega$ and $r > 0$. Let $\epsilon \in (0, 1)$ be given. Then there exists $c = c(N, M_1, \epsilon)$, $1 \le c < \infty$, such that following holds. Assume that u is a nonnegative solution to $\mathcal{K}u = 0$ in $\Omega \cap \mathcal{B}_{2r}(Z_0, t_0)$, vanishing continuously on $\partial\Omega \cap \mathcal{B}_{2r}(Z_0, t_0)$. Then*

$$\sup_{\Omega\cap\mathcal{B}_{r/c}(Z_0,t_0)} u \le \epsilon \sup_{\Omega\cap\mathcal{B}_{2r}(Z_0,t_0)} u. \tag{4-10}$$

Proof. This follows from Lemma 3.11 in [Nyström and Polidoro 2016]. □

Lemma 4.7. *Let $\Omega \subset \mathbb{R}^{N+1}$ be a Lip_K-domain with constant M_1. Let $\Lambda = \Lambda(N, M_1)$ and $c = c(N, M_1)$ be in accordance with Remark 4.5. Let $(Z_0, t_0) \in \partial\Omega$ and $r > 0$. Assume that u is a nonnegative solution to $\mathcal{K}u = 0$ in $\Omega \cap \mathcal{B}_{2r}(Z_0, t_0)$, vanishing continuously on $\partial\Omega \cap \mathcal{B}_{2r}(Z_0, t_0)$. Then*

$$u(Z, t) \le cu(A^+_{\varrho,\Lambda}(Z_0, t_0))$$

whenever $(Z, t) \in \Omega \cap \mathcal{B}_{2\varrho/c}(Z_0, t_0)$, $0 < \varrho < r/c$.

Proof. This is essentially Theorem 1.1 in [Cinti et al. 2013]. □

Theorem 4.8. *Let $\Omega \subset \mathbb{R}^{N+1}$ be a Lip_K-domain with constant M_1. Let $\Lambda = \Lambda(N, M_1)$ and $c = c(N, M_1)$ be in accordance with Remark 4.5. Let $(Z_0, t_0) \in \partial\Omega$ and $r > 0$. Assume that u is a nonnegative solution to $\mathcal{K}u = 0$ in $\Omega \cap \mathcal{B}_{2r}(Z_0, t_0)$, vanishing continuously on $\partial\Omega \cap \mathcal{B}_{2r}(Z_0, t_0)$. Let $\varrho_0 = r/c$,*

$$m^+ = u(A^+_{\varrho_0,\Lambda}(Z_0, t_0)), \quad m^- = u(A^-_{\varrho_0,\Lambda}(Z_0, t_0)), \tag{4-11}$$

and assume $m^- > 0$. Then there exist constants $c_1 = c_1(N, M_1)$, $1 \le c_1 < \infty$, and $c_2 = c_2(N, M_1, m^+/m^-)$, $1 \le c_2 < \infty$, such that if we let $\varrho_1 = \varrho_0/c_1$, then

$$u(Z, t) \le c_2 u(A_{\varrho,\Lambda}(\tilde{Z}_0, \tilde{t}_0)),$$

whenever $(Z, t) \in \Omega \cap \mathcal{B}_{\varrho/c_1}(\tilde{Z}_0, \tilde{t}_0)$, for some $0 < \varrho < \varrho_1$ and $(\tilde{Z}_0, \tilde{t}_0) \in \partial\Omega \cap \mathcal{B}_{\varrho_1}(Z_0, t_0)$.

Proof. Using (4-3) and (4-4), it is easily seen that the theorem is a consequence of Theorem 1.1 in [Nyström and Polidoro 2016]. □

Theorem 4.9. *Let $\Omega \subset \mathbb{R}^{N+1}$ be a Lip_K-domain with constant M_1. Let $\Lambda = \Lambda(N, M_1)$ and $c = c(N, M_1)$ be in accordance with Remark 4.5. Let $(Z_0, t_0) \in \partial\Omega$ and $r > 0$. Assume that u and v are nonnegative* *solutions to* $\mathcal{K}u = 0$ *in* Ω*, vanishing continuously on* $\partial\Omega \cap \mathcal{B}_{2r}(Z_0, t_0)$*. Let* $\varrho_0 = r/c$*,*

$$\begin{aligned} m_1^+ &= v(A^+_{\varrho_0,\Lambda}(Z_0, t_0)), \quad m_1^- = v(A^-_{\varrho_0,\Lambda}(Z_0, t_0)), \\ m_2^+ &= u(A^+_{\varrho_0,\Lambda}(Z_0, t_0)), \quad m_2^- = u(A^-_{\varrho_0,\Lambda}(Z_0, t_0)), \end{aligned} \tag{4-12}$$

and assume $m_1^-, m_2^- > 0$*. Then there exist constants* $c_1 = c_1(N, M_1)$*,* $c_2 = c_2(N, M_1, m_1^+/m_1^-, m_2^+/m_2^-)$*,* $1 \le c_1, c_2 < \infty$*, and* $\sigma = \sigma(N, M_1, m_1^+/m_1^-, m_2^+/m_2^-)$*,* $\sigma \in (0, 1)$*, such that if we let* $\varrho_1 = \varrho_0/c_1$*, then*

$$\left| \frac{v(Z,t)}{u(Z,t)} - \frac{v(\widetilde{Z},\tilde{t})}{u(\widetilde{Z},\tilde{t})} \right| \le c_2 \left(\frac{d((Z,t),(\widetilde{Z},\tilde{t}))}{\varrho} \right)^{\sigma} \frac{v(A_{\varrho,\Lambda}(\widetilde{Z}_0,\tilde{t}_0))}{u(A_{\varrho,\Lambda}(\widetilde{Z}_0,\tilde{t}_0))},$$

whenever $(Z,t), (\widetilde{Z},\tilde{t}) \in \Omega \cap \mathcal{B}_{\varrho/c_1}(\widetilde{Z}_0,\tilde{t}_0)$*, for some* $0 < \varrho < \varrho_1$ *and* $(\widetilde{Z}_0,\tilde{t}_0) \in \partial\Omega \cap \mathcal{B}_{\varrho_1}(Z_0, t_0)$*.*

Proof. Again using (4-3) and (4-4) we see that the theorem is a special case of Theorem 1.2 in [Nyström and Polidoro 2016]. □

4C. *Doubling of parabolic measure and estimates of the kernel function.*

Lemma 4.10. *Let* $\Omega \subset \mathbb{R}^{N+1}$ *be a* Lip_K*-domain with constant* M_1*. Let* $\Lambda = \Lambda(N, M_1)$ *be in accordance with Remark 4.5. Let* $(Z_0, t_0) \in \partial\Omega$ *and* $r > 0$*. Then*

$$\omega\big(A^+_{r/c,\Lambda}(Z_0, t_0), \partial\Omega \cap \mathcal{B}_r(Z_0, t_0)\big) \ge c^{-1}.$$

Proof. This is an immediate consequence of Lemma 4.6. □

Lemma 4.11. *Let* $\Omega \subset \mathbb{R}^{N+1}$ *be a* Lip_K*-domain with constant* M_1*. Let* $\Lambda = \Lambda(N, M_1)$ *be in accordance with Remark 4.5. Let* $(Z_0, t_0) \in \partial\Omega$ *and* $r > 0$*. Let* $\omega(Z, t, \cdot)$ *be the Kolmogorov measure relative to* $(Z, t) \in \Omega$ *and* Ω *and let* $G(Z, t, \cdot, \cdot)$ *be the adjoint Green function for* Ω *with pole at* (Z, t)*. Then there exists* $c = c(N, M_1)$*,* $1 \le c < \infty$*, such that*

(1) $c^{-1} r^{\boldsymbol{q}-2} G(Z, t, A^+_{r,\Lambda}(Z_0, t_0)) \le \omega(Z, t, \partial\Omega \cap \mathcal{B}_r(Z_0, t_0))$,

(2) $\omega(Z, t, \partial\Omega \cap \mathcal{B}_{r/c}(Z_0, t_0)) \le c r^{\boldsymbol{q}-2} G(Z, t, A^-_{r,\Lambda}(Z_0, t_0))$

whenever $(Z, t) \in \Omega$*,* $t - t_0 \ge c r^2$*.*

Proof. This is a consequence of Lemma 4.1 in [Nyström and Polidoro 2016]. However, we emphasize that in that paper the definition of $\boldsymbol{q}$ is different compared to the definition used in this paper; see Remark 3.7. Based on the $\boldsymbol{q}$ used in this paper, $\boldsymbol{q} = 4m + 2$, (i) and (ii) are the correct formulation of the corresponding inequalities in Lemma 4.1 in [Nyström and Polidoro 2016]. □

Lemma 4.12. *Let* $\Omega \subset \mathbb{R}^{N+1}$ *be a* Lip_K*-domain with constant* M_1*. Let* $\Lambda = \Lambda(N, M_1)$ *be in accordance with Remark 4.5. Let* $(Z_0, t_0) \in \partial\Omega$ *and* $r > 0$*. Then there exists* $c = c(N, M_1)$*,* $1 \le c < \infty$*, such that*

$$\omega\big(A^+_{r,\Lambda}(Z_0, t_0), \partial\Omega \cap \mathcal{B}_{2\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0)\big) \le c\,\omega\big(A^+_{r,\Lambda}(Z_0, t_0), \partial\Omega \cap \mathcal{B}_{\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0)\big)$$

whenever $(\widetilde{Z}_0, \tilde{t}_0) \in \partial\Omega$*,* $\mathcal{B}_{\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0) \subset \mathcal{B}_{r/c}(Z_0, t_0)$*.*

Proof. This is a consequence of one of the main results, Theorem 1.3, proved in [Nyström and Polidoro 2016]. However, for the convenience of the reader we include here a proof using the results stated above. Consider $(\widetilde{Z}_0, \tilde{t}_0) \in \partial\Omega$, $\mathcal{B}_{\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0) \subset \mathcal{B}_{r/C}(Z_0, t_0)$, where $C = C(N, M_1) \gg 1$ is a degree of freedom. Choosing $\mathcal{C}$ large enough and using Lemma 4.11(ii) we see that

$$\omega\big(A^+_{r,\Lambda}(Z_0, t_0), \partial\Omega \cap \mathcal{B}_{2\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0)\big) \le \tilde{c}\tilde{r}^{\mathbf{q}-2} G\big(A^+_{r,\Lambda}(Z_0, t_0), A^-_{2c\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0)\big)$$

for some $\tilde{c} = \tilde{c}(N, M_1)$, $1 \le \tilde{c} < \infty$. Let

$$m^+ = G(A^+_{r,\Lambda}(Z_0, t_0), A^+_{r/1000,\Lambda}(Z_0, t_0)), \quad m^- = G(A^+_{r,\Lambda}(Z_0, t_0), A^-_{r/1000,\Lambda}(Z_0, t_0)). \tag{4-13}$$

By elementary estimates and the Harnack inequality, see Lemma 4.4, we see that

$$\bar{c}^{-1} \le r^{\mathbf{q}-2} m^+ \le \bar{c}, \qquad r^{\mathbf{q}-2} m^- \le \bar{c} \tag{4-14}$$

for some $\bar{c} = \bar{c}(N, M_1)$, $1 \le \bar{c} < \infty$. To prove the lemma we intend to use the adjoint version of Theorem 4.8 and hence we need to establish a lower bound on $r^{\mathbf{q}-2} m^-$. To establish this lower bound we first use the adjoint version of Lemma 4.7 to conclude that there exists $c = c(N, M_1)$, $1 \le c < \infty$, such that

$$\sup_{(Z,t)\in\Omega\cap\mathcal{B}_{r/C}(Z_0,t_0)} G\big(A^+_{r,\Lambda}(Z_0, t_0), (Z, t)\big) \le c m^-. \tag{4-15}$$

However,

$$\sup_{(Z,t)\in\Omega\cap\mathcal{B}_{r/C}(Z_0,t_0)} G\big(A^+_{r,\Lambda}(Z_0, t_0), (Z, t)\big) \ge G\big(A^+_{r,\Lambda}(Z_0, t_0), A^+_{r/(100C),\Lambda}(Z_0, t_0)\big) \ge c r^{2-\mathbf{q}} \tag{4-16}$$

by elementary estimates. In particular, (4-14)–(4-16) imply that

$$c^{-1} \le m^+/m^- \le c \quad \text{for some } c = c(N, M_1),\ 1 \le c < \infty.$$

Using this, the adjoint version of Theorem 4.8, and the scale invariance of Theorem 4.8, we deduce that there exist $\tilde{c} = \tilde{c}(N, M_1)$, $1 \le \tilde{c} < \infty$, such that

$$G\big(A^+_{r,\Lambda}(Z_0, t_0), A^-_{2c\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0)\big) \le \tilde{c} G\big(A^+_{r,\Lambda}(Z_0, t_0), A^+_{2c\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0)\big), \tag{4-17}$$

provided $(\widetilde{Z}_0, \tilde{t}_0) \in \partial\Omega$, $\mathcal{B}_{\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0) \subset \mathcal{B}_{r/C}(Z_0, t_0)$. Finally, using the (adjoint) Harnack inequality, we here use the adjoint version of Lemma 4.4, and Lemma 4.11, we see that

$$\begin{aligned}\tilde{r}^{\mathbf{q}-2} G\big(A^+_{r,\Lambda}(Z_0, t_0), A^+_{2c\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0)\big) &\le c\tilde{r}^{\mathbf{q}-2} G\big(A^+_{r,\Lambda}(Z_0, t_0), A^+_{\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0)\big)\\ &\le c^2 \omega\big(A^+_{r,\Lambda}(Z_0, t_0), \partial\Omega \cap \mathcal{B}_{\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0)\big)\end{aligned} \tag{4-18}$$

for some $c = c(N, M_1)$, $1 \le c < \infty$. Combining these, we can conclude the proof of Lemma 4.12. □

Lemma 4.13. *Let $\Omega \subset \mathbb{R}^{N+1}$ be a Lip_K-domain with constant M_1. Let $\Lambda = \Lambda(N, M_1)$ be in accordance with Remark 4.5. Let $(Z_0, t_0) \in \partial\Omega$ and $r > 0$. Let $(\widetilde{Z}_0, \tilde{t}_0) \in \partial\Omega$ and $\tilde{r} > 0$ be such that $\mathcal{B}_{\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0) \subset \mathcal{B}_r(Z_0, t_0)$. Then there exists $c = c(N, M_1)$, $1 \le c < \infty$, such that*

$$K(A^+_{c\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0), \bar{Z}, \bar{t}) := \lim_{\bar{r}\to 0} \frac{\omega(A^+_{c\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0), \partial\Omega \cap \mathcal{B}_{\bar{r}}(\bar{Z}, \bar{t}))}{\omega(A^+_{cr,\Lambda}(Z_0, t_0), \partial\Omega \cap \mathcal{B}_{\bar{r}}(\bar{Z}, \bar{t}))} \tag{4-19}$$

exists for $\omega(A^+_{cr,\Lambda}(Z_0,t_0),\cdot)$ *a.e.* $(\bar Z,\bar t)\in\partial\Omega\cap\mathcal{B}_{\tilde r}(\widetilde Z_0,\tilde t_0)$, *and*

$$c^{-1}\le\omega\bigl(A^+_{cr,\Lambda}(Z_0,t_0),\partial\Omega\cap\mathcal{B}_{\tilde r}(\widetilde Z_0,\tilde t_0)\bigr)K\left(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),\bar Z,\bar t\right)\le c \tag{4-20}$$

whenever $(\bar Z,\bar t)\in\partial\Omega\cap\mathcal{B}_{\tilde r}(\widetilde Z_0,\tilde t_0)$.

Proof. Using the Harnack inequality, see Lemma 4.4, we see that the only thing we have to prove is (4-20). To prove (4-20), consider $(\bar Z,\bar t)\in\partial\Omega\cap\mathcal{B}_{\tilde r}(\widetilde Z_0,\tilde t_0)$ and $\bar r\ll\tilde r$. Using Lemma 4.11 we see that there exists $c=c(N,M_1)$, $1\le c<\infty$, such that

$$\begin{aligned}c^{-1}\frac{G(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),A^+_{\bar r,\Lambda}(\bar Z,\bar t))}{G(A^+_{cr,\Lambda}(Z_0,t_0),A^-_{\bar r,\Lambda}(\bar Z,\bar t))}&\le\frac{\omega(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),\partial\Omega\cap\mathcal{B}_{\bar r}(\bar Z,\bar t))}{\omega(A^+_{cr,\Lambda}(Z_0,t_0),\partial\Omega\cap\mathcal{B}_{\bar r}(\bar Z,\bar t))}\\&\le c\frac{G(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),A^-_{\bar r,\Lambda}(\bar Z,\bar t))}{G(A^+_{cr,\Lambda}(Z_0,t_0),A^+_{\bar r,\Lambda}(\bar Z,\bar t))}.\end{aligned}\tag{4-21}$$

Furthermore, using the adjoint version of Theorem 4.8 and by arguing as in the proof of Lemma 4.12, we see that

$$\begin{aligned}c^{-1}\frac{G(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),A_{\bar r,\Lambda}(\bar Z,\bar t))}{G(A^+_{cr,\Lambda}(Z_0,t_0),A_{\bar r,\Lambda}(\bar Z,\bar t))}&\le\frac{G(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),A^+_{\bar r,\Lambda}(\bar Z,\bar t))}{G(A^+_{cr,\Lambda}(Z_0,t_0),A^-_{\bar r,\Lambda}(\bar Z,\bar t))},\\\frac{G(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),A^-_{\bar r,\Lambda}(\bar Z,\bar t))}{G(A^+_{cr,\Lambda}(Z_0,t_0),A^+_{\bar r,\Lambda}(\bar Z,\bar t))}&\le c\frac{G(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),A_{\bar r,\Lambda}(\bar Z,\bar t))}{G(A^+_{cr,\Lambda}(Z_0,t_0),A_{\bar r,\Lambda}(\bar Z,\bar t))}.\end{aligned}\tag{4-22}$$

Hence we can conclude that

$$\frac{\omega(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),\partial\Omega\cap\mathcal{B}_{\bar r}(\bar Z,\bar t))}{\omega(A^+_{cr,\Lambda}(Z_0,t_0),\partial\Omega\cap\mathcal{B}_{\bar r}(\bar Z,\bar t))}\sim\frac{G(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),A_{\bar r,\Lambda}(\bar Z,\bar t))}{G(A^+_{cr,\Lambda}(Z_0,t_0),A_{\bar r,\Lambda}(\bar Z,\bar t))},\tag{4-23}$$

where $\sim$ means that the quotient between the expression on the left-hand side and the expression on the right-hand side is bounded from above and below by constants depending only on N, M_1. Next, using the boundary Harnack inequality for solutions to the adjoint equation, which is a consequence of the adjoint version of Theorem 4.9, we deduce that

$$\frac{G(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),A_{\bar r,\Lambda}(\bar Z,\bar t))}{G(A^+_{cr,\Lambda}(Z_0,t_0),A_{\bar r,\Lambda}(\bar Z,\bar t))}\sim\frac{G(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),A_{\tilde r/\tilde c,\Lambda}(\widetilde Z_0,\tilde t_0))}{G(A^+_{cr,\Lambda}(Z_0,t_0),A_{\tilde r/\tilde c,\Lambda}(\widetilde Z_0,\tilde t_0))}\tag{4-24}$$

for some $\tilde c=\tilde c(N,M_1)\gg 1$. Combining the inequalities in the last two displays we see that

$$\frac{\omega(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),\partial\Omega\cap\mathcal{B}_{\bar r}(\bar Z,\bar t))}{\omega(A^+_{cr,\Lambda}(Z_0,t_0),\partial\Omega\cap\mathcal{B}_{\bar r}(\bar Z,\bar t))}\sim\frac{G(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),A_{\tilde r/\tilde c,\Lambda}(\widetilde Z_0,\tilde t_0))}{G(A^+_{cr,\Lambda}(Z_0,t_0),A_{\tilde r/\tilde c,\Lambda}(\widetilde Z_0,\tilde t_0))}\tag{4-25}$$

for some $\tilde c=\tilde c(N,M_1)\gg 1$. Finally, using this and arguing by the same principles, using the doubling properties of ω and related estimates, we can conclude that

$$\frac{\omega(A^+_{c\tilde r,\Lambda}(\widetilde Z_0,\tilde t_0),\partial\Omega\cap\mathcal{B}_{\bar r}(\bar Z,\bar t))}{\omega(A^+_{cr,\Lambda}(Z_0,t_0),\partial\Omega\cap\mathcal{B}_{\bar r}(\bar Z,\bar t))}\sim\frac{1}{\omega(A^+_{cr,\Lambda}(Z_0,t_0),\partial\Omega\cap\mathcal{B}_{\tilde r}(\widetilde Z_0,\tilde t_0)))}\tag{4-26}$$

whenever $(\bar Z,\bar t)\in\mathcal{B}_{\tilde r}(\widetilde Z_0,\tilde t_0)$. Using this and (4-19), we deduce (4-20) by letting $\bar r\to 0$. □

Lemma 4.14. *Let $\Omega \subset \mathbb{R}^{N+1}$ be a Lip_K-domain with constant M_1. Let $\Lambda = \Lambda(N, M_1)$ be in accordance with Remark 4.5. Let $(Z_0, t_0) \in \partial\Omega$ and $r > 0$. Let $(\widetilde{Z}_0, \tilde{t}_0) \in \partial\Omega$ and $\tilde{r} > 0$ be such that $\mathcal{B}_{\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0) \subset \mathcal{B}_r(Z_0, t_0)$. Then there exist $c = c(N, M_1)$, $1 \le c < \infty$, and $\tilde{c} = \tilde{c}(N, M_1)$, $1 \le \tilde{c} < \infty$, such that*

$$\tilde{c}^{-1}\omega(A^+_{c\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0), E) \le \frac{\omega(A^+_{cr,\Lambda}(Z_0, t_0), E)}{\omega(A^+_{cr,\Lambda}(Z_0, t_0), \partial\Omega \cap \mathcal{B}_{\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0))} \le \tilde{c}\,\omega(A^+_{c\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0), E)$$

whenever $E \subset \mathcal{B}_{\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0)$.

Proof. Consider $E \subset \mathcal{B}_{\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0)$. Then, by definition

$$\omega(A^+_{c\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0), E) = \int_E K(A^+_{c\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0), \bar{Z}, \bar{t})\, d\omega(A^+_{cr,\Lambda}(Z_0, t_0), \bar{Z}, \bar{t}).$$

Hence, using Lemma 4.13 we see that

$$\omega\big(A^+_{cr,\Lambda}(Z_0, t_0), \partial\Omega \cap \mathcal{B}_{\tilde{r}}(\widetilde{Z}_0, \tilde{t}_0)\big)\, \omega(A^+_{c\tilde{r},\Lambda}(\widetilde{Z}_0, \tilde{t}_0), E) \sim \omega(A^+_{cr,\Lambda}(Z_0, t_0), E),$$

which is the statement to be proved. □

5. Proof of Theorem 1.6

In order to introduce some efficient notation, we will use the terminology of spaces of homogeneous type in the sense of [Coifman and Weiss 1971]. Indeed, assuming that $\Omega = \Omega_\psi \subset \mathbb{R}^{N+1}$ is a Lip_K-domain, with constant M_1, in the sense of Definition 1.1, we let

$$\Sigma := \partial\Omega = \big\{(x, x_m, y, y_m, t) \in \mathbb{R}^{N+1} \,\big|\, x_m = \psi(x, y, t)\big\}.$$

Then $(\Sigma, d, d\sigma)$ is a space of homogeneous type, with homogeneous dimension $\boldsymbol{q} - 1$. Furthermore, $(\mathbb{R}^{N+1}, d, dZ\,dt)$ is also a space of homogeneous type, but with homogeneous dimension $\boldsymbol{q}$.

5A. *Dyadic grids, Whitney cubes and Carleson boxes.* By the results in [Christ 1990] there exists what we here will refer to as a dyadic grid on Σ having a number of important properties in relation to d. To formulate this we introduce, for any $(Z, t) = (X, Y, t) \in \Sigma$ and $E \subset \Sigma$,

$$\mathrm{dist}((Z, t), E) := \inf\big\{d((Z, t), (\widetilde{Z}, \tilde{t})) \,\big|\, (\widetilde{Z}, \tilde{t}) \in E\big\}, \tag{5-1}$$

and we let

$$\mathrm{diam}(E) := \sup\big\{d((Z, t), (\widetilde{Z}, \tilde{t})) \,\big|\, (Z, t), (\widetilde{Z}, \tilde{t}) \in E\big\}. \tag{5-2}$$

Using [Christ 1990] we can conclude that there exist constants $\alpha > 0$, $\beta > 0$ and $c_* < \infty$ such that for each $k \in \mathbb{Z}$ there exists a collection of Borel sets, $\mathbb{D}_k$, which we will call cubes, such that

$$\mathbb{D}_k := \{Q^k_j \subset \Sigma \mid j \in \mathfrak{I}_k\}, \tag{5-3}$$

where $\mathfrak{I}_k$ denotes some index set depending on k, satisfying:

(i) $\Sigma = \bigcup_j Q^k_j$ for each $k \in \mathbb{Z}$.

(ii) If $m \ge k$ then either $Q^m_i \subset Q^k_j$ or $Q^m_i \cap Q^k_j = \varnothing$.

(iii) For each (j,k) and each $m<k$, there is a unique i such that $Q_j^k \subset Q_i^m$.

(iv) $\operatorname{diam}(Q_j^k) \le c_* 2^{-k}$.

(v) Each Q_j^k contains $\Sigma \cap \mathcal{B}_{\alpha 2^{-k}}(Z_j^k, t_j^k)$ for some $(Z_j^k, t_j^k) \in \Sigma$.

(vi) $\sigma\big(\big\{(Z,t)\in Q_j^k \;\big|\; \operatorname{dist}((Z,t), \Sigma \setminus Q_j^k) \le \varrho\, 2^{-k}\big\}\big) \le c_*\, \varrho^\beta\, \sigma(Q_j^k)$ for all k, j and for all $\varrho \in (0,\alpha)$.

Let us make a few remarks concerning this result and discuss some related notation and terminology. First, in the setting of a general space of homogeneous type, this result is due to Christ [1990], with the dyadic parameter $\frac12$ replaced by some constant $\delta \in (0,1)$. In fact, one may always take $\delta = \frac12$; see [Hofmann et al. 2017, proof of Proposition 2.12]. We shall denote by $\mathbb{D} = \mathbb{D}(\Sigma)$ the collection of all Q_j^k; i.e.,

$$\mathbb{D} := \bigcup_k \mathbb{D}_k.$$

Note that (iv) and (v) imply that for each cube $Q \in \mathbb{D}_k$, there is a point $(Z_Q, t_Q) = (X_Q, Y_Q, t_Q) \in \Sigma$ and a ball $\mathcal{B}_r(Z_Q, t_Q)$ such that $r \approx 2^{-k} \approx \operatorname{diam}(Q)$ and

$$\Sigma \cap \mathcal{B}_r(Z_Q, t_Q) \subset Q \subset \Sigma \cap \mathcal{B}_{cr}(Z_Q, t_Q) \tag{5-4}$$

for some uniform constant c. We will denote the associated surface ball by

$$\Delta_Q := \Sigma \cap \mathcal{B}_r(Z_Q, t_Q), \tag{5-5}$$

and we shall refer to the point (Z_Q, t_Q) as the center of Q. Given a dyadic cube $Q \subset \Sigma$, we define its γ dilate by

$$\gamma Q := \Sigma \cap \mathcal{B}_{\gamma \operatorname{diam}(Q)}(Z_Q, t_Q). \tag{5-6}$$

For a dyadic cube $Q \in \mathbb{D}_k$, we let $\ell(Q) = 2^{-k}$, and we shall refer to this quantity as the length of Q. Clearly, $\ell(Q) \approx \operatorname{diam}(Q)$. For a dyadic cube $Q \in \mathbb{D}$, we let $k(Q)$ denote the dyadic generation to which Q belongs; i.e., we set $k = k(Q)$ if $Q \in \mathbb{D}_k$, thus, $\ell(Q) = 2^{-k(Q)}$. For any $Q \in \mathbb{D}(\Sigma)$, we set $\mathbb{D}_Q := \{Q' \in \mathbb{D} \mid Q' \subset Q\}$.

Using that also $(\mathbb{R}^{N+1}, d, dZ\,dt)$ is a space of homogeneous type, we see that we can partition Ω into a collection of (closed) dyadic Whitney cubes $\{I\}$, in the following denoted $\mathcal{W} = \mathcal{W}(\Omega)$, such that the cubes in $\mathcal{W}$ form a covering of Ω with nonoverlapping interiors, and which satisfy

$$4 \operatorname{diam}(I) \le \operatorname{dist}(4I, \Sigma) \le \operatorname{dist}(I, \Sigma) \le 40 \operatorname{diam}(I) \tag{5-7}$$

and

$$\operatorname{diam}(I_1) \approx \operatorname{diam}(I_2) \quad \text{whenever } I_1 \text{ and } I_2 \text{ touch.} \tag{5-8}$$

Given $I \in \mathcal{W}$, we let $\ell(I)$ denote its size. Given $Q \in \mathbb{D}(\Sigma)$, we set

$$\mathcal{W}_Q := \big\{I \in \mathcal{W} \;\big|\; 100^{-1}\ell(Q) \le \ell(I) \le 100\,\ell(Q) \text{ and } \operatorname{dist}(I, Q) \le 100\,\ell(Q)\big\}. \tag{5-9}$$

We fix a small, positive parameter τ, and given $I \in \mathcal{W}$, we let

$$I^* = I^*(\tau) := (1+\tau) I \tag{5-10}$$

denote the corresponding "fattened" Whitney cube. Choosing τ small, we see that the cubes I^* will retain the usual properties of Whitney cubes, in particular that

$$\operatorname{diam}(I) \approx \operatorname{diam}(I^*) \approx \operatorname{dist}(I^*, \Sigma) \approx \operatorname{dist}(I, \Sigma).$$

We then define a Whitney region with respect to Q by setting

$$U_Q := \bigcup_{I \in \mathcal{W}_Q} I^*. \tag{5-11}$$

Finally, given $Q \in \mathbb{D}(\Sigma)$, $\gamma \geq 0$, we let

$$T_Q := \operatorname{int}\biggl(\bigcup_{Q' \in \mathbb{D}_Q} U_{Q'}\biggr) \tag{5-12}$$

denote the Carleson box associated to Q. Furthermore, given $\gamma \geq 1$ we let

$$T_{\gamma Q} := \operatorname{int}\biggl(\bigcup_{Q' : Q' \cap (\gamma Q) \neq \varnothing} U_{Q'}\biggr) \tag{5-13}$$

denote the Carleson set associated to the γ dilate of Q.

5B. *Reduction of Theorem 1.6 to two key lemmas.* Using Lemma 4.12, we see that to prove Theorem 1.6 it suffices to prove the following version of Theorem 1.6.

Theorem 5.1. *Assume that $\Omega \subset \mathbb{R}^{N+1}$ is an (unbounded) admissible Lip_K-domain with constants (M_1, M_2) in the sense of Definition 1.1. Then there exist $\Lambda = \Lambda(N, M_1)$, $1 \leq \Lambda < \infty$, and $c = c(N, M_1)$, $1 \leq c < \infty$, and $\tilde{c} = \tilde{c}(N, M_1, M_2)$, $1 \leq \tilde{c} < \infty$, and $\eta = \eta(N, M_1, M_2)$, $0 < \eta < 1$, such that the following is true. Let $Q_0 \in \mathbb{D}$, $\varrho_0 := l(Q_0)$ and let $\omega(\,\cdot\,) := \omega(A^+_{c\varrho_0,\Lambda}(Z_{Q_0}, t_{Q_0}), \cdot\,)$. Then*

$$\tilde{c}^{-1}\biggl(\frac{\sigma(E)}{\sigma(Q)}\biggr)^{1/\eta} \leq \frac{\omega(E)}{\omega(Q)} \leq \tilde{c}\biggl(\frac{\sigma(E)}{\sigma(Q)}\biggr)^{\eta}$$

whenever $E \subset Q$, $Q \in \mathbb{D}$, $Q \subseteq Q_0$.

The proof of Theorem 5.1 is based on the following lemmas.

Lemma 5.2. *Let $Q_0 \in \mathbb{D}$ and let $\omega(\,\cdot\,)$ be as in the statement of Theorem 1.6. Let $\kappa \gg 1$ be given and consider $\delta_0 \in (0, 1)$. Assume that $E \subset Q_0$ with $\omega(E) \leq \delta_0$. If $\delta_0 = \delta_0(N, M_1, \kappa)$ is chosen sufficiently small, then there exist a Borel set $S \subset \partial\Omega$, and a constant $c = c(N, M_1)$, $1 \leq c < \infty$, such that if we let $u(Z, t) := \omega(Z, t, S)$, then*

$$\kappa^2 \sigma(E) \leq c \iint_{T_{cQ_0}} \bigl(|\nabla_X u|^2 \delta + |\nabla_Y u|^2 \delta^5 + |X_0(u)|^2 \delta^3\bigr)\, dZ\, dt.$$

Here $\delta = \delta(Z, t)$ is the distance from $(Z, t) \in \Omega$ to Σ and T_{cQ_0} is the Carleson set associated to cQ_0 as defined in (5-13).

Lemma 5.3. *Let $Q_0 \in \mathbb{D}$ and let $\omega(\,\cdot\,)$ be as in the statement of Theorem 1.6. Let $u(Z,t) := \omega(Z,t,S)$ and c be as stated in Lemma 5.2. Then there exists $\tilde c = \tilde c(N, M_1, M_2)$, $1 \le \tilde c < \infty$, such that*

$$\iint_{T_{cQ_0}} \big(|\nabla_X u|^2\delta + |\nabla_Y u|^2\delta^5 + |X_0(u)|^2\delta^3\big)\, dZ\, dt \le \tilde c\sigma(Q_0).$$

The proof of Lemma 5.2 is given below. The proof of Lemma 5.3 is given in the next section. We here prove Theorem 5.1, hence completing the proof of Theorem 1.6, assuming Lemmas 5.2 and 5.3. Indeed, first using Lemmas 4.14 and 4.12 we see that it suffices to prove Theorem 5.1 with $Q = Q_0$. Then, using Lemmas 5.2 and 5.3 we see that we can, for $\Gamma \gg 1$ given, choose $\delta_0 = \delta_0(N, M_1, \Gamma)$ so that if $E \subset Q_0$ with $\omega(E) \le \delta_0$, then

$$\Gamma^2\sigma(E) \le \hat c\bar\sigma(Q_0) \tag{5-14}$$

for some $\hat c = \hat c(N, M_1, M_2)$, $1 \le \hat c < \infty$. In particular, we can conclude that there exists, for every $\epsilon > 0$, a positive $\delta_0 = \delta_0(N, M_1, M_2, \epsilon)$ such that

$$\omega(E) \le \delta_0 \le c\delta_0\omega(Q_0) \implies \sigma(E) \le \epsilon\sigma(Q_0), \tag{5-15}$$

where we have also applied Lemma 4.10. Theorem 1.6 now follows from the doubling property of ω, see Lemma 4.12, and the classical result in [Coifman and Fefferman 1974].

5C. *Good ϵ_0 covers.* Recall that in the following $\omega(\,\cdot\,) := \omega(A^+_{c\varrho_0,\Lambda}(Z_{Q_0}, t_{Q_0}), \cdot)$, $Q_0 \in \mathbb{D}$, $\varrho_0 := l(Q_0)$.

Definition 5.4. Let $E \subset Q_0$ be given, let $\epsilon_0 \in (0,1)$ and let k be an integer. A good ϵ_0 cover of E, of length k, is a collection $\{\mathcal{O}_l\}_{l=1}^k$ of nested (relatively) open subsets of Q_0, together with collections $\mathcal{F}_l = \{\Delta_i^l\}_i \subset Q_0$, $\Delta_i^l \in \mathbb{D}$, such that

$$E \subset \mathcal{O}_k \subset \mathcal{O}_{k-1} \subset \cdots \subset \mathcal{O}_1 \subset Q_0, \tag{5-16}$$

$$\mathcal{O}_l = \bigcup_{\mathcal{F}_l} \Delta_i^l, \tag{5-17}$$

$$\omega(\mathcal{O}_l \cap \Delta_i^{l-1}) \le \epsilon_0\omega(\Delta_i^{l-1}) \quad \text{for all } \Delta_i^{l-1} \in \mathcal{F}_{l-1}. \tag{5-18}$$

Lemma 5.5. *Let $E \subset Q_0$ be given and consider $\epsilon_0 \in (0,1)$. There exists $\gamma = \gamma(N, M_1)$, $0 < \gamma \ll 1$, and $\Gamma = \Gamma(N, M_1)$, $1 \ll \Gamma$, such that if we let $\delta_0 = \gamma(\epsilon_0/\Gamma)^k$, and if $\omega(E) \le \delta_0$, then E has a good ϵ_0 cover of length k.*

Proof. Let $k \in \mathbb{Z}_+$ be given. Let γ, $0 < \gamma \ll 1$, and Γ, $1 \ll \Gamma$, be degrees of freedom to be chosen depending only on N and M_1. Let $\delta_0 = \gamma(\epsilon_0/\Gamma)^k$. Suppose that $\omega(E) \le \delta_0$. Using that ω is a regular Borel measure, we see that there exists a (relatively) open subset of Q_0, containing E, which we denote by $\mathcal{O}_{k+1}$, satisfying $\omega(\mathcal{O}_{k+1}) \le 2\omega(E)$. Using Lemma 4.10 and the Harnack inequality, see Lemma 4.4, we see that there exists $c = c(N, M_1)$, $1 \le c < \infty$, such that

$$\omega(\mathcal{O}_{k+1}) \le 2\delta_0 \le c\delta_0\omega(Q_0) \le \frac{1}{2}\Big(\frac{\epsilon_0}{\Gamma}\Big)^k\omega(Q_0) \tag{5-19}$$

if we let $\gamma := 2/c$. Let $f \in L^1_{\mathrm{loc}}(\Sigma, d\omega)$, and let

$$M_\omega(f)(Z,t) := \sup_{\mathfrak{B}} \frac{1}{\omega(\mathcal{B}_r(\widetilde{Z},\tilde{t}))} \int_{\mathcal{B}_r(\widetilde{Z},\tilde{t})} f\, d\omega,$$

where $\mathfrak{B} = \{\mathcal{B}_r(\widetilde{Z},\tilde{t}) \mid (\widetilde{Z},\tilde{t}) \in \partial\Omega,\ (Z,t) \in \mathcal{B}_r(\widetilde{Z},\tilde{t})\}$, denote the Hardy–Littlewood maximal function of f, with respect to ω, and where the supremum is taken over all balls $\mathcal{B}_r(\widetilde{Z},\tilde{t})$, $(\widetilde{Z},\tilde{t}) \in \partial\Omega$, containing (Z,t). Set

$$\mathcal{O}_k := \{(Z,t) \in Q_0 \mid M_\omega(1_{\mathcal{O}_{k+1}}) \geq \epsilon_0/\bar{c}\},$$

where we let $\bar{c} = \bar{c}(N, M_1)$, $1 \leq \bar{c} < \infty$, denote the constant appearing in Lemma 4.12. Then, by construction, $\mathcal{O}_{k+1} \subset \mathcal{O}_k$, $\mathcal{O}_k$ is relatively open in Q_0 and $\mathcal{O}_k$ is properly contained in Q_0. As ω is doubling, see Lemma 4.12, $(2Q_0, d, \omega)$ is a space of homogeneous type, weak L^1 estimates for the Hardy–Littlewood maximal function apply and hence

$$\omega(\mathcal{O}_k) \leq \tilde{c}\,\frac{\bar{c}}{\epsilon_0}\,\omega(\mathcal{O}_{k+1}) \leq \frac{1}{2}\left(\frac{\epsilon_0}{\bar{c}}\right)^{k-1} \omega(Q_0), \tag{5-20}$$

if we let $\Gamma = \tilde{c}\bar{c}$ and where $\tilde{c} = \tilde{c}(N, M_1)$, $1 \leq \tilde{c} < \infty$. By definition and by the construction, see (i)–(iii) on page 1733, Q_0 can be dyadically subdivided, and we can select a collection $\mathcal{F}_k = \{\Delta_i^k\}_i \subset Q_0$, comprised of the cubes that are maximal with respect to containment in $\mathcal{O}_k$, and thus $\mathcal{O}_k := \bigcup_i \Delta_i^k$. Then, by the maximality of the cubes in $\mathcal{F}_k$, and by the doubling property of ω, we find that

$$\omega(\mathcal{O}_{k+1} \cap \Delta_i^k) \leq \epsilon_0 \omega(\Delta_i^k) \quad \text{for all } \Delta_i^k \in \mathcal{F}_k. \tag{5-21}$$

We now iterate this argument, to construct $\mathcal{O}_{j-1}$ from $\mathcal{O}_j$ for $2 \leq j \leq k$, just as we constructed $\mathcal{O}_k$ from $\mathcal{O}_{k+1}$. It is then a routine matter to verify that the sets $\mathcal{O}_1, \dots, \mathcal{O}_k$, form a good ϵ_0 cover of E. We omit further details. □

Remark 5.6. From now on we fix a small dyadic number $\eta = 2^{-k_0}$, where k_0 is to be chosen. Given $Q \in \mathbb{D}$, we consider the k_0-grandchildren of Q, i.e., the subcubes $Q' \subset Q$, $Q' \in \mathbb{D}$, with length $l(Q') = \eta l(Q)$. We let $\widetilde{Q}$ denote the particular such grandchild which contains the center of Q, (Z_Q, t_Q).

Remark 5.7. Given $Q \in \mathbb{D}$ we let $A_Q^+ = A^+_{cl(Q),\Lambda}(Z_Q, t_Q)$ and $A^+_{\widetilde{Q}} = A^+_{cl(\widetilde{Q}),\Lambda}(Z_{\widetilde{Q}}, t_{\widetilde{Q}})$, where $\widetilde{Q}$ was defined in Remark 5.6.

Remark 5.8. Consider the special case $\Delta := \Delta_i^l \in \mathcal{F}_l$; i.e., Δ is a cube arising in some good ϵ_0 cover. We then set $\tilde{\Delta}_i^l := \tilde{\Delta}$, where $\tilde{\Delta}$ is defined as in Remark 5.6, and we define

$$\widetilde{\mathcal{O}}_l := \bigcup_{\Delta_i^l \in \mathcal{F}_l} \tilde{\Delta}_i^l. \tag{5-22}$$

Remark 5.9. Let $E \subset Q$ and consider the setup of Lemma 5.5. We note that for every $(Z_0, t_0) \in E$ we have $(Z_0, t_0) \in \mathcal{O}_l$ for all $l = 1, 2, \dots, k$, and therefore there exists, for each l, a cube $\Delta_i^l = \Delta_i^l(Z_0, t_0) \in \mathcal{F}_l$ containing (Z_0, t_0). With (Z_0, t_0) fixed, we let $\hat{\Delta}_i^l = \hat{\Delta}_i^l(Z_0, t_0)$ denote the particular k_0-grandchild, as defined in Remark 5.6, of Δ_i^l that contains (Z_0, t_0).

5D. *Proof of Lemma 5.2.* To prove Lemma 5.2, let $\epsilon_0 > 0$ be a degree of freedom to be specified below and depending only on N, M_1, let $\delta_0 = \gamma(\epsilon_0/\Gamma)^k$ be as specified in Lemma 5.5, where k is to be chosen depending only on N, M_1 and κ. Consider $E \subset Q_0$ with $\omega(E) \leq \delta_0$. Using Lemma 5.5, we see that E has a good ϵ_0 cover of length k, $\{\mathcal{O}_l\}_{l=1}^k$ with corresponding collections $\mathcal{F}_l = \{\Delta_i^l\}_i \subset Q_0$. Let $\{\widetilde{\mathcal{O}}_l\}_{l=1}^k$ be defined as in (5-22). Using this good ϵ_0 cover of E we let

$$F(Z,t) := \sum_{j=2}^{k} \chi_{\widetilde{\mathcal{O}}_{j-1}\setminus\mathcal{O}_j}(Z,t),$$

where $\chi_{\widetilde{\mathcal{O}}_{j-1}\setminus\mathcal{O}_j}$ is the indicator function for the set $\chi_{\widetilde{\mathcal{O}}_{j-1}\setminus\mathcal{O}_j}$. Then F equals the indicator function of some Borel set $S \subset \Sigma$ and we let $u(Z,t) := \omega(Z,t,S)$. Consider

$$(Z_0,t_0) \in E \quad \text{and} \quad \text{an index } l \in \{1,\dots,k\}.$$

In the following let

$$\Delta_i^l \in \mathcal{F}_l \text{ be a cube in the collection } \mathcal{F}_l \text{ which contains } (Z_0,t_0).$$

Given $k_0 \in \mathbb{Z}_+$ we let

$$\tilde{\Delta}_i^l \text{ be the } k_0\text{-grandchild of } \Delta_i^l \text{ which contains } (Z_{\Delta_i^l}, t_{\Delta_i^l}).$$

With (Z_0,t_0) and Δ_i^l fixed, we let $\hat{\Delta}_i^l$ be defined as in Remark 5.9; i.e., we let

$$\hat{\Delta}_i^l \text{ be the } k_0\text{-grandchild of } \Delta_i^l \text{ which contains } (Z_0,t_0).$$

Finally, we let

$$\bar{\Delta}_i^l \text{ be the } k_0\text{-grandchild of } \hat{\Delta}_i^l \text{ which contains } (Z_{\hat{\Delta}_i^l}, t_{\hat{\Delta}_i^l}).$$

Hence, based on $(Z_0,t_0) \in E$ and an index $l \in \{1,\dots,k\}$, we have specified Δ_i^l, $\tilde{\Delta}_i^l$, $\hat{\Delta}_i^l$ and $\bar{\Delta}_i^l$ satisfying

$$\tilde{\Delta}_i^l \subset \Delta_i^l, \quad \bar{\Delta}_i^l \subset \hat{\Delta}_i^l \subset \Delta_i^l.$$

We let

$$A^+_{\tilde{\Delta}_i^l} = A^+_{cl(\tilde{\Delta}_i^l),\Lambda}(Z_{\tilde{\Delta}_i^l}, t_{\tilde{\Delta}_i^l}), \quad A^+_{\bar{\Delta}_i^l} = A^+_{cl(\bar{\Delta}_i^l),\Lambda}(Z_{\bar{\Delta}_i^l}, t_{\bar{\Delta}_i^l}). \tag{5-23}$$

We first intend to prove that there exists $\beta > 0$, depending only on N, M_1, κ, such that if ϵ_0 and $\eta = 2^{-k_0}$ are chosen sufficiently small, then

$$|u(A^+_{\tilde{\Delta}_i^l}) - u(A^+_{\bar{\Delta}_i^l})| \geq \beta. \tag{5-24}$$

To estimate $u(A^+_{\tilde{\Delta}_i^l})$ we write

$$u(A^+_{\tilde{\Delta}_i^l}) = \int_{Q_0\setminus\Delta_i^l} F(\bar{Z},\bar{t})\, d\omega(A^+_{\tilde{\Delta}_i^l}, \bar{Z},\bar{t}) + \int_{\Delta_i^l} F(\bar{Z},\bar{t})\, d\omega(A^+_{\tilde{\Delta}_i^l}, \bar{Z},\bar{t}) =: I + II. \tag{5-25}$$

Using Lemma 4.6 and the definition of $A^+_{\tilde{\Delta}_i^l}$ we see that

$$|I| \leq \omega(A^+_{\tilde{\Delta}_i^l}, Q_0 \setminus \Delta_i^l) \leq c\eta^\sigma \tag{5-26}$$

for some $c = c(N, M_1)$, $\sigma = \sigma(N, M_1) \in (0, 1)$. Furthermore, by the definition of F we see that

$$II = II_1 + II_2 + II_3, \tag{5-27}$$

where

$$\begin{aligned} II_1 &:= \sum_{j=2}^{l} \int_{\Delta_i^l} 1_{\widetilde{\mathcal{O}}_{j-1}\setminus\mathcal{O}_j}\, d\omega(A^+_{\tilde{\Delta}_i^l}, \bar{Z}, \bar{t}),\\ II_2 &:= \sum_{j=l+2}^{k} \int_{\Delta_i^l} 1_{\widetilde{\mathcal{O}}_{j-1}\setminus\mathcal{O}_j}\, d\omega(A^+_{\tilde{\Delta}_i^l}, \bar{Z}, \bar{t}),\\ III_3 &:= \int_{\Delta_i^l} 1_{\widetilde{\mathcal{O}}_{l}\setminus\mathcal{O}_{l+1}}\, d\omega(A^+_{\tilde{\Delta}_i^l}, \bar{Z}, \bar{t}). \end{aligned} \tag{5-28}$$

Note that if $j \le l$, then $\Delta_i^l \subset \mathcal{O}_l \subset \mathcal{O}_j$ and $(\widetilde{\mathcal{O}}_{j-1} \setminus \mathcal{O}_j) \cap \Delta_i^l = \varnothing$. Hence $II_1 = 0$. Obviously,

$$|II_2| \le \sum_{j=l+2}^{k} \omega\big(A^+_{\tilde{\Delta}_i^l}, (\widetilde{\mathcal{O}}_{j-1} \setminus \mathcal{O}_j) \cap \Delta_i^l\big) \le c_\eta \sum_{j=l+2}^{k} \omega\big(A^+_{\Delta_i^l}, (\widetilde{\mathcal{O}}_{j-1} \setminus \mathcal{O}_j) \cap \Delta_i^l\big), \tag{5-29}$$

where we in the second estimate have used the Harnack inequality; see Lemma 4.4. Consider $(\bar{Z}, \bar{t}) \in (\widetilde{\mathcal{O}}_{j-1} \setminus \mathcal{O}_j) \cap \Delta_i^l$. Then, using Lemma 4.13 we have

$$K(A^+_{\Delta_i^l}, \bar{Z}, \bar{t}) := \lim_{\varrho \to 0} \frac{\omega(A^+_{\Delta_i^l}, \partial\Omega \cap \mathcal{B}_\varrho(\bar{Z}, \bar{t}))}{\omega(\partial\Omega \cap \mathcal{B}_\varrho(\bar{Z}, \bar{t}))} \tag{5-30}$$

exists for ω a.e. $(\bar{Z}, \bar{t}) \in \Delta_i^l$, and

$$K(A^+_{\Delta_i^l}, \bar{Z}, \bar{t}) \le \frac{c}{\omega(\Delta_i^l)} \quad \text{whenever } (\bar{Z}, \bar{t}) \in \Delta_i^l. \tag{5-31}$$

In the last conclusion we have also used Lemma 4.12. Using these facts, and using the definition of the good ϵ_0 cover, we see that

$$\begin{aligned} |II_2| &\le \frac{c_\eta}{\omega(\Delta_i^l)} \sum_{j=l+2}^{k} \omega((\widetilde{\mathcal{O}}_{j-1} \setminus \mathcal{O}_j) \cap \Delta_i^l)\\ &\le \frac{c_\eta}{\omega(\Delta_i^l)} \sum_{j=l+2}^{k} \omega(\mathcal{O}_{j-1} \cap \Delta_i^l) \le \frac{c_\eta}{\omega(\Delta_i^l)} \sum_{j=l+2}^{k} \epsilon_0^{j-1-l} \omega(\Delta_i^l) \le c_\eta \epsilon_0. \end{aligned} \tag{5-32}$$

To estimate the term II_3 we first observe that $\Delta_i^l \cap \widetilde{\mathcal{O}}_l = \tilde{\Delta}_i^l$ by the definition of $\widetilde{\mathcal{O}}_l$. Hence,

$$\begin{aligned} II_3 &= \omega(A^+_{\tilde{\Delta}_i^l}, \Delta_i^l \cap \widetilde{\mathcal{O}}_l) - \omega(A^+_{\tilde{\Delta}_i^l}, \Delta_i^l \cap \mathcal{O}_{l+1})\\ &= \omega(A^+_{\tilde{\Delta}_i^l}, \tilde{\Delta}_i^l) - \omega(A^+_{\tilde{\Delta}_i^l}, \Delta_i^l \cap \mathcal{O}_{l+1}) =: II_{31} + II_{32}. \end{aligned} \tag{5-33}$$

Arguing as in (5-32), we see that

$$|II_{32}| \le \frac{c_\eta}{\omega(\Delta_i^l)} \omega(\Delta_i^l \cap \mathcal{O}_{l+1}) \le C_\eta \epsilon_0, \tag{5-34}$$

by the construction. Putting these together we can conclude that so far we have proved that

$$|u(A^+_{\tilde{\Delta}^l_i}) - II_{31}| \leq c\eta^\sigma + c_\eta\epsilon_0, \tag{5-35}$$

and it remains to analyze II_{31}. However, using Lemma 4.10, and elementary estimates, we see that there exists $\tilde{c} = \tilde{c}(N, M_1)$, $1 \leq \tilde{c} < \infty$, such that

$$\tilde{c}^{-1} \leq II_{31} \leq 1 - \tilde{c}^{-1}.$$

Combining the last result and (5-35) we can conclude, by first choosing $\eta = \eta(N, M_1)$ small and then $\epsilon_0 = \epsilon_0(N, M_1, \eta)$ small, that

$$\tfrac{3}{4}\tilde{c}^{-1} \leq u(A^+_{\tilde{\Delta}^l_i}) \leq 1 - \tfrac{3}{4}\tilde{c}^{-1}. \tag{5-36}$$

To estimate $u(A^+_{\hat{\Delta}^l_i})$ we write

$$u(A^+_{\hat{\Delta}^l_i}) = \int_{Q_0\setminus\hat{\Delta}^l_i} F(\bar{Z},\bar{t})\,d\omega(A^+_{\hat{\Delta}^l_i},\bar{Z},\bar{t}) + \int_{\hat{\Delta}^l_i} F(\bar{Z},\bar{t})\,d\omega(A^+_{\hat{\Delta}^l_i},\bar{Z},\bar{t}) =: \hat{I} + \widehat{II}. \tag{5-37}$$

We split $\widehat{II}$ as

$$\widehat{II} = \widehat{II}_1 + \widehat{II}_2 + \widehat{II}_3, \tag{5-38}$$

where

$$\begin{aligned} \widehat{II}_1 &:= \sum_{j=2}^{l} \int_{\hat{\Delta}^l_i} 1_{\widetilde{\mathcal{O}}_{j-1}\setminus\mathcal{O}_j}\,d\omega(A^+_{\hat{\Delta}^l_i},\bar{Z},\bar{t},),\\ \widehat{II}_2 &:= \sum_{j=l+2}^{k} \int_{\hat{\Delta}^l_i} 1_{\widetilde{\mathcal{O}}_{j-1}\setminus\mathcal{O}_j}\,d\omega(A^+_{\hat{\Delta}^l_i},\bar{Z},\bar{t}),\\ \widehat{II}_3 &:= \int_{\hat{\Delta}^l_i} 1_{\widetilde{\mathcal{O}}_l\setminus\mathcal{O}_{l+1}}\,d\omega(A^+_{\hat{\Delta}^l_i},\bar{Z},\bar{t}). \end{aligned} \tag{5-39}$$

We can now conclude, by essentially repeating the estimates in the corresponding estimate for $u(A^+_{\tilde{\Delta}^l_i})$ above, that

$$|\hat{I}| + |\widehat{II}_1| + |\widehat{II}_2| \leq c\eta^\sigma + c_\eta\epsilon_0. \tag{5-40}$$

Note that the key estimate is now the kernel estimate

$$K(A^+_{\hat{\Delta}^l_i},\bar{Z},\bar{t}) \leq \frac{c}{\omega(\Delta^l_i)} \quad \text{whenever } (\bar{Z},\bar{t}) \in \Delta^l_i, \tag{5-41}$$

which again follows from Lemma 4.13. We now focus on $\widehat{II}_3$ and we observe that

$$\widehat{II}_3 = \omega(A^+_{\hat{\Delta}^l_i}, \hat{\Delta}^l_i \cap (\widetilde{\mathcal{O}}_l \setminus \mathcal{O}_{l+1})) = \omega(A^+_{\hat{\Delta}^l_i}, (\hat{\Delta}^l_i \cap \tilde{\Delta}^l_i) \setminus \mathcal{O}_{l+1}), \tag{5-42}$$

and that either $(\hat{\Delta}^l_i \cap \tilde{\Delta}^l_i) = \varnothing$ or $\hat{\Delta}^l_i = \tilde{\Delta}^l_i$. If $(\hat{\Delta}^l_i \cap \tilde{\Delta}^l_i) = \varnothing$, then $\widehat{II}_3 = 0$. If $\hat{\Delta}^l_i = \tilde{\Delta}^l_i$, then

$$\widehat{II}_3 = \omega(A^+_{\hat{\Delta}^l_i}, \hat{\Delta}^l_i) - \omega(A^+_{\hat{\Delta}^l_i}, \hat{\Delta}^l_i \setminus \mathcal{O}_{l+1}) =: \widehat{II}_{31} + \widehat{II}_{32}. \tag{5-43}$$

Also in this case

$$\widehat{II}_{32} \leq c_\eta\epsilon_0,$$

and we are left with $\widehat{II}_{31}$. However, by Lemma 4.6 we deduce that

$$\widehat{II}_{31} = \omega(A^+_{\tilde{\Delta}^l_i}, \hat{\Delta}^l_i) \geq 1 - c\eta^\sigma,$$

where $c = c(N, M_1)$, $\sigma = \sigma(N, M_1) \in (0, 1)$. Combining our estimates we find that either

$$0 \leq u(A^+_{\tilde{\Delta}^l_i}) \leq c\eta^\sigma + c_\eta\epsilon_0 \quad \text{or} \quad u(A^+_{\tilde{\Delta}^l_i}) \geq 1 - (c\eta^\sigma + c_\eta\epsilon_0). \tag{5-44}$$

Combining (5-36) and (5-44) we can conclude, in either case, that

$$|u(A^+_{\tilde{\Delta}^l_i}) - u(A^+_{\tilde{\Delta}^l_i})| \geq \tfrac{3}{4}\tilde{c}^{-1} - c\eta^\sigma + c_\eta\epsilon_0 \geq \tfrac{1}{2}\tilde{c}^{-1} \tag{5-45}$$

by first choosing $\eta = \eta(N, M_1)$ small and then choosing $\epsilon_0 = \epsilon_0(N, M_1, \eta)$ small. Hence the proof of (5-24) is complete.

Next, using (5-24), Lemma 2.5(i), an elementary connectivity/covering argument and the Poincaré inequality, see Lemma 2.4, we see that

$$c^{-1}\beta^2 \leq \iint_{\widetilde{W}_{\Delta^l_i}} \big(|\nabla_X u|^2\delta^{2-\mathbf{q}} + |\nabla_Y u|^2\delta^{6-\mathbf{q}} + |X_0(u)|^2\delta^{4-\mathbf{q}}\big)\, dZ\, dt,$$

where $\widetilde{W}_{\Delta^l_i}$ is a natural Whitney-type region associated to Δ^l_i, $\delta = \delta(Z, t)$ is the distance from (Z, t) to Σ, and $c = c(N, M_1, \eta)$, $1 \leq c < \infty$. Consequently, for $(Z_0, t_0) \in E$ fixed we find, by summing over all indices i, l, such that $(Z_0, t_0) \in \Delta^l_i$, that

$$c^{-1}\beta^2 k \leq \sum_{i,l:(Z_0,t_0)\in\Delta^l_i} \left(\iint_{\widetilde{W}_{\Delta^l_i}} \big(|\nabla_X u|^2\delta^{2-\mathbf{q}} + |\nabla_Y u|^2\delta^{6-\mathbf{q}} + |X_0(u)|^2\delta^{4-\mathbf{q}}\big)\, dZ\, dt\right). \tag{5-46}$$

The construction can be made so that the Whitney-type regions $\{\widetilde{W}_{\Delta^l_i}\}$ have bounded overlaps measured by a constant depending only on N, M_1, and such that $W_{Q^l_i} \subset T_{cQ_0}$ for some $c = c(N, M_1)$, $1 \leq c < \infty$, where T_{cQ_0} is defined in (5-13). Hence, integrating with respect to $d\sigma$, we deduce that

$$c^{-1}\beta^2 k\sigma(E) \leq \left(\iint_{T_{cQ_0}} \big(|\nabla_X u|^2\delta + |\nabla_Y u|^2\delta^5 + |X_0(u)|^2\delta^3\big)\, dZ\, dt\right), \tag{5-47}$$

where, resolving the dependencies, $c = c(N, M_1)$, $1 \leq c < \infty$. Furthermore,

$$k \approx \frac{\log(\delta_0)}{\log(\epsilon_0)},$$

where η and ϵ_0 now have been fixed, and δ_0 is at our disposal. Given κ we obtain the conclusion of the lemma by specifying $\delta_0 = \delta_0(N, M_1, \kappa)$ sufficiently small. This completes the proof of Lemma 5.2.

6. The square function estimate: proof of Lemma 5.3

The purpose of the section is to prove Lemma 5.3. Hence we consider $Q_0 \in \mathbb{D}(\Sigma)$, we let $\varrho_0 = l(Q_0)$, $u(Z, t) := \omega(Z, t, S)$ and we let c be as stated in Lemma 5.2. We want to prove that there exists

$\tilde{c} = \tilde{c}(N, M_1, M_2),\ 1 \le \tilde{c} < \infty$, such that

$$\iint_{T_c Q_0} \big(|\nabla_X u|^2 \delta + |\nabla_Y u|^2 \delta^5 + |X_0(u)|^2 \delta^3\big)\, dZ\, dt \le \tilde{c} \varrho_0^{\boldsymbol{q}-1}. \tag{6-1}$$

However, using Lemma 2.8, and a simple covering argument, we first note that to prove (6-1) it suffices to prove that

$$\iint_{T_c Q_0} |\nabla_X u|^2 \delta\, dZ\, dt \le \tilde{c} \varrho_0^{\boldsymbol{q}-1} \tag{6-2}$$

for potentially new constants c, $\tilde{c}$ having the same dependence as the original constants c, $\tilde{c}$. Inequality (6-2) will be proved using partial integration. To enable partial integration, we perform the change of variables

$$(w, w_m, y, y_m, t) \in U \to (w, w_m + P_{\gamma w_m} \psi(w, y, t), y, y_m, t) \tag{6-3}$$

where

$$U = \big\{(W, Y, t) = (w, w_m, y, y_m, t) \in \mathbb{R}^{m-1} \times \mathbb{R} \times \mathbb{R}^{m-1} \times \mathbb{R} \times \mathbb{R} \;\big|\; w_m > 0\big\}. \tag{6-4}$$

Then, by a straightforward calculation we see that u satisfies $\mathcal{K}u = 0$ in Ω if and only if $v(W, Y, t) := u(w, w_m + P_{\gamma w_m} \psi(w, y, t), y, y_m, t)$ satisfies

$$\nabla_W \cdot (A \nabla_W v) + B \cdot \nabla_W v + \big((w, w_m + P_{\gamma w_m} \psi(w, y, t)) \cdot \nabla_Y - \partial_t\big) v = 0 \quad \text{in } U. \tag{6-5}$$

Here A is an $m \times m$-matrix-valued function, $B : U \to \mathbb{R}^m$ and

$$\begin{aligned} a_{m,m} &= \frac{1 + |\nabla_w P_{\gamma w_m} \psi|^2}{(1 + \partial_{w_m} P_{\gamma w_m} \psi)^2}, \\ a_{j,m} = a_{m,j} &= -\frac{\partial_{w_j} P_{\gamma w_m} \psi}{(1 + \partial_{w_m} P_{\gamma w_m} \psi)}, \qquad j = 1, \dots, m-1, \\ a_{i,j} &= \delta_{i,j}, \quad i, j \in \{1, \dots, m-1\}, \end{aligned} \tag{6-6}$$

and

$$\begin{aligned} b_m &= \frac{\partial_{w_m w_m} P_{\gamma w_m} \psi}{(1 + \partial_{w_m} P_{\gamma w_m} \psi)} + \frac{\big((w, w_m + P_{\gamma w_m} \psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)(P_{\gamma w_m} \psi)}{(1 + \partial_{w_m} P_{\gamma w_m} \psi)}, \\ b_j &= \frac{\partial_{w_m w_j} P_{\gamma w_m} \psi}{(1 + \partial_{w_m} P_{\gamma w_m} \psi)}, \qquad j = 1, \dots, m-1. \end{aligned} \tag{6-7}$$

Choosing $\gamma = \gamma(N, M_1),\ \gamma > 0$, small enough, see Remark 2.3, we have

$$(c(N, M_1))^{-1} |\xi|^2 \le \sum_{i,j=1}^{m} a_{i,j}(W, Y, t) \xi_i \xi_j \le c(N, M_1) |\xi|^2, \quad c(N, M_1) \ge 1, \tag{6-8}$$

for all $(W, Y, t) \in \mathbb{R}^{N+1}$ and for all $\xi \in \mathbb{R}^m$. Let $\theta, \tilde{\theta} \ge 0$ be integers, let $(\phi_1, \dots, \phi_{m-1})$ and $(\tilde{\phi}_1, \dots, \tilde{\phi}_{m-1})$ denote multi-indices and let $\ell = (\theta + |\phi| + 3|\tilde{\phi}| + 2\tilde{\theta})$. Then, using Lemma 2.1 and the fact that ψ is independent of y_m, we deduce that

$$\begin{aligned}\left|\frac{\partial^{\theta+|\phi|+|\tilde{\phi}|}}{\partial w_m^\theta \partial w^\phi \partial y^{\tilde{\phi}}}\Big(\big((w, w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)^{\tilde{\theta}}(A(W,Y,t))\Big)\right| &\leq c(m,l,\gamma)w_m^{-l}M_1,\\ \left|\frac{\partial^{\theta+|\phi|+|\tilde{\phi}|}}{\partial w_m^\theta \partial w^\phi \partial y^{\tilde{\phi}}}\Big(\big((w, w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)^{\tilde{\theta}}(B(W,Y,t))\Big)\right| &\leq c(m,l,\gamma)w_m^{-1-l}M_1\end{aligned}\tag{6-9}$$

whenever $(W,Y,t)\in U$. Similarly, using Lemma 2.2 we see that if we let $d\mu_i=d\mu_i(W,Y,t),\ i\in\{1,2\}$,

$$\begin{aligned}d\mu_1 &:= \left|\frac{\partial^{\theta+|\phi|+|\tilde{\phi}|}}{\partial w_m^\theta \partial w^\phi \partial y^{\tilde{\phi}}}\Big(\big((w, w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)^{\tilde{\theta}}(A(W,Y,t))\Big)\right|^2 w_m^{2l-1}\,dW\,dy\,dt,\\ d\mu_2 &:= \left|\frac{\partial^{\theta+|\phi|+|\tilde{\phi}|}}{\partial w_m^\theta \partial w^\phi \partial y^{\tilde{\phi}}}\Big(\big((w, w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)^{\tilde{\theta}}(B(W,Y,t))\Big)\right|^2 w_m^{2l+1}\,dW\,dy\,dt,\end{aligned}$$

be defined on U, then

$$\mu_1(U\cap\mathcal{B}_\varrho(w_0,0,Y_0,t_0))+\mu_2(U\cap\mathcal{B}_\varrho(w_0,0,Y_0,t_0))\leq c(m,l,\gamma,M_1,M_2)\varrho^{q-1}\tag{6-10}$$

whenever $(w_0,0,Y_0,t_0)\in\partial U,\ \varrho>0$, and $\mathcal{B}_\varrho(w_0,0,Y_0,t_0)\subset\mathbb{R}^{N+1}$. We emphasize that

$$A \text{ and } B \text{ are independent of } y_m.\tag{6-11}$$

As the equation $\mathcal{K}u=0$ and the statements in (6-9)–(6-11) are invariant under left translation defined by $\circ$, we can in the following without loss of generality assume that

$$(Z_{Q_0},t_{Q_0})=\big(x_{Q_0},\psi(x_{Q_0},y_{Q_0},t_{Q_0}),y_{Q_0},y_{m,Q_0},t_{Q_0}\big)=(0,0,0,0,0).\tag{6-12}$$

Furthermore, given $c\geq 1$ we claim that there exist $\alpha=\alpha(N,c)\geq 1$ and $\beta=\beta(N,M_1,c)\geq 1$ such that if we define

$$\square_0=\big\{(W,Y,t)=(w,w_m,y,y_m,t)\ \big|\ |w|<\alpha\varrho_0,\ 0<w_m<\beta\varrho_0,\ |Y|<\alpha^3\varrho_0^3,\ |t|<\alpha^2\varrho_0^2\big\},\tag{6-13}$$

then

$$T_{cQ_0}\subset\big\{\big(w,w_m+P_{\gamma w_m}\psi(w,y,t),y,y_m,t\big)\ \big|\ (w,w_m,y,y_m,t)\in\square_0\big\}.\tag{6-14}$$

In the following we let $2\square_0$ be defined as in (6-13) but with ϱ_0 replaced by $2\varrho_0$. Letting $(\square_0)^*$ and $(2\square_0)^*$ denote the sets we get if we reflect $\square_0$ and $2\square_0$, respectively, in the boundary ∂U, in the following we let $\zeta\in C_0^\infty(2\square_0\cup(2\square_0)^*)$ be such that $0\leq\zeta\leq 1,\ \zeta\equiv 1$ on $\square_0\cup\square_0^*$, and such that

$$\varrho_0|\nabla_W\zeta|+\varrho_0^3|\nabla_Y\zeta|+\varrho_0^2|(W\cdot\nabla_Y-\partial_t)(\zeta)|\leq c(N,M_1).\tag{6-15}$$

Letting $v(W,Y,t)=u\big(w,w_m+P_{\gamma w_m}\psi(w,y,t),y,y_m,t\big)$, where $u(Z,t)=\omega(Z,t,S)$, we see that

$$|v(W,Y,t)|\leq 1\quad\text{whenever } (W,Y,t)\in U.\tag{6-16}$$

Using (6-10) we see that

$$\iint_{2\square_0} |v(W,Y,t)|^2\, d\mu_i(W,Y,t) \le c(m,l,\gamma,M_1,M_2,i)\varrho_0^{\boldsymbol{q}-1}, \quad i\in\{1,2\}. \tag{6-17}$$

To prove (6-2), and hence to complete the proof of Lemma 5.3, it suffices to prove that

$$\iint_{2\square_0} |\nabla_W v(W,Y,t)|^2 \zeta^2 w_m\, dW\, dy\, dt \le c(N,M_1,M_2)\varrho_0^{\boldsymbol{q}-1}. \tag{6-18}$$

The rest of the proof is devoted to the proof of (6-18) and in the proof of (6-18) we will use the notation

$$\begin{aligned} T_1 &:= \iint_{2\square_0} |\nabla_W v|^2 \zeta^2 w_m\, dW\, dy\, dt,\\ T_2 &:= \iint_{2\square_0} \big|\big((w, w_m + P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y - \partial_t\big)(v)\big|^2 \zeta^4 w_m^3\, dW\, dy\, dt. \end{aligned} \tag{6-19}$$

Inequality (6-18) is a consequence of the following two lemmas.

Lemma 6.1. *Let $\square_0$, ζ and v be as above. Then there exists, for $\epsilon>0$ given, $c=c(N,M_1,M_2,\epsilon)$, $1\le c<\infty$, such that*

$$T_1 \le c\varrho_0^{\boldsymbol{q}-1} + \epsilon T_2.$$

Lemma 6.2. *Let $\square_0$, ζ and v be as above. Then there exists $c=c(N,M_1,M_2)$, $1\le c<\infty$, such that*

$$T_2 \le c(T_1 + \varrho_0^{\boldsymbol{q}-1}).$$

6A. ***Proof of Lemma 6.1.*** Using (6-8) we see that

$$T_1 \le cI, \quad I := \sum_{i,j=1}^{m} I_{i,j},$$

where

$$I_{i,j} := 2\iint_{2\square_0} \Big(\frac{1}{a_{m,m}}\Big) a_{i,j}(\partial_{w_i} v)(\partial_{w_j} v) w_m \zeta^2\, dW\, dY\, dt.$$

Assume first that $i\neq m$. Then, integrating by parts in $I_{i,j}$ with respect to w_i we see that

$$\begin{aligned} I_{i,j} = &-2\iint_{2\square_0} \Big(\frac{1}{a_{m,m}}\Big) v\,\partial_{w_i}(a_{i,j}\partial_{w_j} v) w_m \zeta^2\, dW\, dY\, dt\\ &-2\iint_{2\square_0} \partial_{w_i}\Big(\frac{1}{a_{m,m}}\Big) a_{i,j} v(\partial_{w_j} v) w_m \zeta^2\, dW\, dY\, dt\\ &-4\iint_{2\square_0} \Big(\frac{a_{i,j}}{a_{m,m}}\Big) v(\partial_{w_j} v) w_m \zeta\,\partial_{w_i}\zeta\, dW\, dY\, dt. \end{aligned}$$

Similarly we see that

$$\begin{aligned}
I_{m,j} = &- \lim_{\delta\to 0} 2\int_{2\square_0\cap\{w_m=\delta\}} \left(\frac{a_{m,j}}{a_{m,m}}\right)(w,\delta,Y,t)v(w,\delta,Y,t)(\partial_{w_j}v(w,\delta,Y,t))\delta\zeta^2\,dw\,dY\,dt\\
&-2\iint_{2\square_0}\left(\frac{1}{a_{m,m}}\right)v\partial_{w_m}(a_{m,j}\partial_{w_j}v)w_m\zeta^2\,dW\,dY\,dt\\
&-2\iint_{2\square_0}\partial_{w_m}\left(\frac{1}{a_{m,m}}\right)a_{m,j}v(\partial_{w_j}v)w_m\zeta^2\,dW\,dY\,dt\\
&-2\iint_{2\square_0}\left(\frac{a_{m,j}}{a_{m,m}}\right)v(\partial_{w_j}v)\zeta^2\,dW\,dY\,dt\\
&-4\iint_{2\square_0}\left(\frac{a_{m,j}}{a_{m,m}}\right)v(\partial_{w_j}v)w_m\zeta\partial_{w_m}\zeta\,dW\,dY\,dt.
\end{aligned}$$

Combining the above,

$$I = \lim_{\delta\to 0} I_1^\delta + I_2 + I_3 + I_4 + I_5,$$

where

$$\begin{aligned}
I_1^\delta &:= -2\sum_j\int_{2\square_0\cap\{w_m=\delta\}}\left(\frac{a_{m,j}}{a_{m,m}}\right)(w,\delta,Y,t)v(w,\delta,Y,t)(\partial_{w_j}v(w,\delta,Y,t))\delta\zeta^2\,dw\,dY\,dt,\\
I_2 &:= -2\sum_{i,j}\iint_{2\square_0}\left(\frac{1}{a_{m,m}}\right)v\partial_{w_i}(a_{i,j}\partial_{w_j}v)w_m\zeta^2\,dW\,dY\,dt,\\
I_3 &:= -2\sum_{i,j}\iint_{2\square_0}\partial_{w_i}\left(\frac{1}{a_{m,m}}\right)a_{i,j}v(\partial_{w_j}v)w_m\zeta^2\,dW\,dY\,dt,\\
I_4 &:= -4\sum_{i,j}\iint_{2\square_0}\left(\frac{a_{i,j}}{a_{m,m}}\right)v(\partial_{w_j}v)w_m\zeta\partial_{w_i}\zeta\,dW\,dy\,dt,\\
I_5 &:= -2\sum_{j}\iint_{2\square_0}\left(\frac{a_{m,j}}{a_{m,m}}\right)v(\partial_{w_j}v)\zeta^2\,dW\,dy\,dt.
\end{aligned}$$

Using (6-9), Lemma 2.5(iii) and (6-15) we see that

$$|I_1^\delta| \le c\varrho_0^{q-1}. \tag{6-20}$$

We next analyze I_2. Using the equation

$$\begin{aligned}
I_2 = 2\iint_{2\square_0}\left(\frac{1}{a_{m,m}}\right)v\big(((w,w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t)v\big)w_m\zeta^2\,dW\,dy\,dt&\\
+2\sum_i\iint_{2\square_0}\left(\frac{1}{a_{m,m}}\right)vb_i\partial_{w_i}vw_m\zeta^2\,dW\,dy\,dt&\\
=: I_{21}+I_{22},&
\end{aligned}$$

we have

$$I_{22}\le c\left(\iint_{2\square_0}v^2|B|^2w_m\zeta^2\,dW\,dy\,dt\right)^{1/2}\left(\iint_{2\square_0}|\nabla_Wv|^2w_m\zeta^2\,dW\,dy\,dt\right)^{1/2}\le c(\epsilon)\varrho_0^{q-1}+\epsilon T_1$$

by (6-10), (6-17) and where ϵ is a degree of freedom. Furthermore, integrating by parts with respect to w_m we see that

$$I_{21} = \lim_{\delta \to 0} I_{211}^{\delta} + I_{212} + I_{213} + I_{214} + I_{215},$$

where

$$\begin{aligned}
I_{211}^{\delta} &= \int_{2\square_0 \cap \{w_m = \delta\}} \left(\frac{1}{a_{m,m}}\right) v\big(((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t)v\big)\delta^2\zeta^2 \, dw \, dY \, dt, \\
I_{212} &= -\iint_{2\square_0} \partial_{w_m}\left(\frac{1}{a_{m,m}}\right) v\big(((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t)v\big) w_m^2 \zeta^2 \, dW \, dY \, dt, \\
I_{213} &= -\iint_{2\square_0} \left(\frac{1}{a_{m,m}}\right) \partial_{w_m} v\big(((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t)v\big) w_m^2 \zeta^2 \, dW \, dY \, dt, \\
I_{214} &= -\iint_{2\square_0} \left(\frac{1}{a_{m,m}}\right) v\partial_{w_m}\big(((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t)v\big) w_m^2 \zeta^2 \, dW \, dY \, dt, \\
I_{215} &= -2\iint_{2\square_0} \left(\frac{1}{a_{m,m}}\right) v\big(((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t)v\big) w_m^2 \zeta \partial_{w_m}\zeta \, dW \, dY \, dt.
\end{aligned}$$

To estimate I_{211}^{δ} we again have to use Lemma 2.5. Indeed, using that

$$v(W, Y, t) = u\big(w, w_m + P_{\gamma w_m}\psi(w, y, t), y, y_m, t\big) = u(x, x_m, y, y_m, t)$$

we see that

$$\begin{aligned}
&\big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)v \\
&\qquad = (X \cdot \nabla_Y - \partial_t)u(X, Y, t) + \partial_{x_m}u(X, Y, t)\big((w \cdot \nabla_y - \partial_t)(\mathcal{P}_{\gamma w_m}\psi(x, y, t))\big).
\end{aligned}$$

Hence, using (6-9), Lemma 2.5(iii) and Lemma 2.1 we first see that

$$\big|\big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)v\big| \le c\delta^{-2}$$

whenever $(W, Y, t) \in 2\square_0 \cap \{w_m = \delta\}$ and then that $|I_{211}^{\delta}| \le c\varrho_0^{\mathbf{q}-1}$. Focusing on I_{212} we see that

$$\begin{aligned}
I_{212} &= -\iint_{2\square_0} \left(\frac{\partial_{w_m} a_{m,m}}{a_{m,m}^2}\right) v\big(((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t)v\big) w_m^2 \zeta^2 \, dW \, dy \, dt \\
&\le c\left(\iint_{2\square_0} |\partial_{w_m} a_{m,m}|^2 v^2 w_m \, dW \, dy \, dt\right) T_2^{1/2} \\
&\le c(\epsilon)\varrho_0^{\mathbf{q}-1} + \epsilon T_2,
\end{aligned}$$

by (6-10), (6-17), and where ϵ is a degree of freedom. To continue we see that

$$\begin{aligned}
I_{214} = &-\iint_{2\square_0} \left(\frac{1}{a_{m,m}}\right) v\big(((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t)\partial_{w_m} v\big) w_m^2 \zeta^2 \, dW \, dy \, dt \\
&\qquad - \iint_{2\square_0} \left(\frac{1}{a_{m,m}}\right) v(1 + \partial_{w_m} P_{\gamma w_m}\psi(w, y, t))(\partial_{y_m} v) w_m^2 \zeta^2 \, dW \, dy \, dt \\
&=: I_{2141} + I_{2142}.
\end{aligned}$$

To estimate I_{2142} we write

$$
\begin{aligned}
I_{2142} &= -\frac{1}{2}\iint_{2\square_0}\Big(\frac{1}{a_{m,m}}\Big)(1+\partial_{w_m}P_{\gamma w_m}\psi(w,y,t))(\partial_{y_m}v^2)w_m^2\zeta^2\,dW\,dy\,dt\\
&= \iint_{2\square_0}\Big(\frac{1}{a_{m,m}}\Big)(1+\partial_{w_m}P_{\gamma w_m}\psi(w,y,t))v^2w_m^2\zeta\,\partial_{y_m}\zeta\,dW\,dy\,dt,
\end{aligned}
$$

where we have used that ψ is independent of y_m. In particular, $|I_{2142}|\le c\varrho_0^{\boldsymbol{q}-1}$. Focusing on I_{2141},

$$
\begin{aligned}
I_{2141} = {}&\iint_{2\square_0}\big((w,w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)\Big(\frac{1}{a_{m,m}}\Big)v(\partial_{w_m}v)w_m^2\zeta^2\,dW\,dy\,dt\\
&+\iint_{2\square_0}\Big(\frac{1}{a_{m,m}}\Big)\big(\big((w,w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)v\big)(\partial_{w_m}v)w_m^2\zeta^2\,dW\,dy\,dt,\\
&+2\iint_{2\square_0}\Big(\frac{1}{a_{m,m}}\Big)v(\partial_{w_m}v)w_m^2\zeta\big((w,w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)\zeta\,dW\,dy\,dt\\
=:{}& I_{21411}+I_{21412}+I_{21413}.
\end{aligned}
$$

Again using (6-10), (6-17), (6-9), and elementary estimates we see that

$$|I_{21411}|+|I_{21413}|\le c(\epsilon)\varrho_0^{\boldsymbol{q}-1}+\epsilon T_1,$$

where ϵ is a degree of freedom. Furthermore,

$$I_{21412}=-I_{213}.$$

Finally,

$$
\begin{aligned}
I_{215} &= -\iint_{2\square_0}\Big(\frac{1}{a_{m,m}}\Big)\big(\big((w,w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)v^2\big)w_m^2\zeta\,\partial_{w_m}\zeta\,dW\,dy\,dt\\
&= \iint_{2\square_0}\big((w,w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)\Big(\frac{1}{a_{m,m}}\Big)v^2w_m^2\zeta\,\partial_{w_m}\zeta\,dW\,dy\,dt\\
&\quad+\iint_{2\square_0}\Big(\frac{1}{a_{m,m}}\Big)v^2w_m^2\big(\big((w,w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)\zeta\big)\partial_{w_m}\zeta\,dW\,dy\,dt\\
&\quad+\iint_{2\square_0}\Big(\frac{1}{a_{m,m}}\Big)v^2w_m^2\zeta\big((w,w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)\partial_{w_m}\zeta\,dW\,dy\,dt\\
&=: I_{2151}+I_{2152}+I_{2153}.
\end{aligned}
$$

Using (6-9), (6-10), (6-17), (6-12), (6-15), and by now familiar arguments, we see that $|I_{215}|\le c\varrho_0^{\boldsymbol{q}-1}$. Combining the above, we can conclude that

$$
\begin{aligned}
|I_2| &\le |I_{211}^{\delta}|+|I_{212}|+|I_{213}|+|I_{214}|+|I_{215}|+|I_{22}|\\
&\le c(\epsilon,\tilde\epsilon)\varrho_0^{\boldsymbol{q}-1}+\epsilon T_1+\tilde\epsilon T_2,
\end{aligned}
$$

where ϵ and $\tilde\epsilon$ are degrees of freedom. Similarly,

$$|I_3|+|I_4|\le c(\epsilon)\varrho_0^{\boldsymbol{q}-1}+\epsilon T_1,$$

and we can conclude that

$$|I_1^\delta|+|I_2|+|I_3|+|I_4|\le c(\epsilon)\varrho_0^{\boldsymbol{q}-1}+\epsilon T_1.$$

Finally we consider I_5,

$$I_5=-2\sum_j\iint_{2\square_0}\left(\frac{a_{m,j}}{a_{m,m}}\right)v(\partial_{w_j}v)\zeta^2\,dW\,dy\,dt.$$

First we consider the term in the definition of I_5 which corresponds to $j=m$. Then

$$\begin{aligned}-\iint_{2\square_0}&\partial_{w_m}(v^2)\zeta^2\,dW\,dy\,dt\\&=-\lim_{\delta\to0}\int_{2\square_0\cap\{w_m=\delta\}}(v^2)(w,\delta,Y,t)\zeta^2\,dw\,dY\,dt+2\iint_{2\square_0}v^2\zeta\,\partial_{w_m}\zeta\,dW\,dY\,dt,\end{aligned}$$

and obviously the absolute value of the terms on the right-hand side is bounded by $c\varrho_0^{\boldsymbol{q}-1}$. Next we consider the terms in the definition of I_5 which correspond to $j\neq m$. By integration by parts we see that

$$\begin{aligned}-2\iint_{2\square_0}&\left(\frac{a_{m,j}}{a_{m,m}}\right)v(\partial_{w_j}v)\partial_{w_m}(w_m)\zeta^2\,dW\,dY\,dt\\
&=\lim_{\delta\to0}2\int_{2\square_0\cap\{w_m=\delta\}}\left(\frac{a_{m,j}}{a_{m,m}}\right)(w,\delta,Y,t)v(w,\delta,Y,t)\partial_{w_j}v(w,\delta,Y,t)\delta\zeta^2\,dw\,dY\,dt\\
&\quad+2\iint_{2\square_0}\partial_{w_m}\left(\frac{a_{m,j}}{a_{m,m}}\right)v\partial_{w_j}v\,w_m\zeta^2\,dW\,dY\,dt\\
&\quad+2\iint_{2\square_0}\left(\frac{a_{m,j}}{a_{m,m}}\right)\partial_{w_m}v\partial_{w_j}v\,w_m\zeta^2\,dW\,dY\,dt\\
&\quad+2\iint_{2\square_0}\left(\frac{a_{m,j}}{a_{m,m}}\right)v\partial_{w_mw_j}v\,w_m\zeta^2\,dW\,dY\,dt\\
&\quad+4\iint_{2\square_0}\left(\frac{a_{m,j}}{a_{m,m}}\right)v\partial_{w_j}v\,w_m\zeta\,\partial_{w_m}\zeta\,dW\,dY\,dt.\end{aligned}$$

Let

$$I_{51}:=2\sum_{j\neq m}\iint_{2\square_0}\left(\frac{a_{m,j}}{a_{m,m}}\right)\partial_{w_m}v\partial_{w_j}v\,w_m\zeta^2\,dW\,dY\,dt,$$

$$I_{52}:=2\sum_{j\neq m}\iint_{2\square_0}\left(\frac{a_{m,j}}{a_{m,m}}\right)v\partial_{w_mw_j}v\,w_m\zeta^2\,dW\,dY\,dt.$$

By the above deductions, and using by now familiar arguments, we can conclude that

$$|I_5-I_{51}-I_{52}|\le c(\epsilon)\varrho_0^{\boldsymbol{q}-1}+\epsilon T_1.$$

To estimate I_{52} we use that $j \neq m$. Integrating by parts

$$\begin{aligned} I_{52} = &-2 \sum_{j\neq m} \iint_{2\square_0} \partial_{w_j}\left(\frac{a_{m,j}}{a_{m,m}}\right) v\partial_{w_m} v w_m \zeta^2 \, dW\, dY\, dt \\ &-2 \sum_{j\neq m} \iint_{2\square_0} \left(\frac{a_{m,j}}{a_{m,m}}\right) \partial_{w_j} v\partial_{w_m} v w_m \zeta^2 \, dW\, dY\, dt \\ &-4 \sum_{j\neq m} \iint_{2\square_0} \left(\frac{a_{m,j}}{a_{m,m}}\right) v\partial_{w_m} v w_m \zeta \partial_{w_j}\zeta \, dW\, dY\, dt \\ := &\, I_{521} + I_{522} + I_{523}. \end{aligned}$$

Note that

$$I_{522} = -I_{51},$$

and that

$$|I_{521}| + |I_{523}| \leq c(\epsilon)\varrho_0^{q-1} + \epsilon T_1$$

by familiar arguments. Summarizing we can conclude that

$$c^{-1}T_1 \leq I \leq |I_1^\delta| + |I_2| + |I_3| + |I_4| + |I_5| \leq c(\epsilon, \tilde{\epsilon})\varrho_0^{q-1} + \epsilon T_1 + \tilde{\epsilon} T_2,$$

where ϵ, $\tilde{\epsilon}$ are degrees of freedom. This completes the proof of the lemma.

6B. ***Additional technical estimates.*** In this subsection we prove some additional technical estimates that will be used in the proof of Lemma 6.2. Let

$$\begin{aligned} T_3 &= \sum_{i=1}^m \iint_{2\square_0} |\nabla_W(\partial_{w_i} v)|^2 w_m^3 \zeta^4 \, dW\, dY\, dt, \\ T_4 &= \sum_{i=1}^m \iint_{2\square_0} |\nabla_W(\partial_{y_i} v)|^2 w_m^7 \zeta^8 \, dW\, dY\, dt, \\ T_5 &= \iint_{2\square_0} |\nabla_Y v|^2 w_m^5 \zeta^6 \, dW\, dY\, dt. \end{aligned} \tag{6-21}$$

Lemma 6.3. *Let $\square_0$, ζ and v be as in Lemma 6.1. Then there exists, for positive $\epsilon_1 - \epsilon_4$ given, $c = c(N, M_1, M_2, \epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)$, $1 \leq c < \infty$, such that*

(i) $T_3 \leq c\varrho_0^{q-1} + cT_1 + \epsilon_1 T_5$,

(ii) $T_4 \leq c\varrho_0^{q-1} + cT_5 + \epsilon_2 T_1 + \epsilon_3 T_3$,

(iii) $T_5 \leq c\varrho_0^{q-1} + cT_1 + cT_2 + \epsilon_4 T_4$.

Proof. To prove (i) we introduce $\tilde{v} = \partial_{w_i} v$. Using (6-5) we see that $\tilde{v}$ solves

$$\begin{aligned} &\nabla_W \cdot (A\nabla_W \tilde{v}) + B\cdot\nabla_W \tilde{v} + \big((w, w_m + P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y - \partial_t\big)\tilde{v} \\ &\qquad = -\nabla_W\cdot(\partial_{w_i} A\nabla_W v) - \partial_{w_i} B\cdot\nabla_W v - \partial_{y_i} v - \partial_{w_i} P_{\gamma w_m}\psi(w,y,t)\partial_{y_m} v \end{aligned} \tag{6-22}$$

in U. Multiplying the equation in (6-22) with $\tilde v w_m^3\zeta^4$ and integrating we see that

$$J := -\iint_{2\square_0}(\nabla_W\cdot(A\nabla_W\tilde v))\tilde v w_m^3\zeta^4\,dW\,dY\,dt = J_1+J_2+J_3+J_4, \tag{6-23}$$

where

$$\begin{aligned}
J_1 &:= \iint_{2\square_0}\big(\big((w, w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)\tilde v\big)\tilde v w_m^3\zeta^4\,dW\,dy\,dt,\\
J_2 &:= \iint_{2\square_0}\big(\nabla_W\cdot((\partial_{w_i}A)\nabla_W v)\big)\tilde v w_m^3\zeta^4\,dW\,dy\,dt,\\
J_3 &:= \iint_{2\square_0}(\partial_{w_i}(B\cdot\nabla_W v))\tilde v w_m^3\zeta^4\,dW\,dy\,dt,\\
J_4 &:= \iint_{2\square_0}\big(\partial_{y_i}v+\partial_{w_i}P_{\gamma w_m}\psi(w,y,t)\partial_{y_m}v\big)\tilde v w_m^3\zeta^4\,dW\,dy\,dt.
\end{aligned} \tag{6-24}$$

Using (6-9) we immediately see that

$$|J_2|+|J_3|+|J_4|\le cT_1+cT_1^{1/2}T_3^{1/2}+cT_1^{1/2}T_5^{1/2}. \tag{6-25}$$

Furthermore, using (6-12), (6-15) we see that

$$2|J_1|\le 4\left|\iint_{2\square_0}\tilde v^2\big((w, w_m+P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y-\partial_t\big)(\zeta)w_m^3\zeta^3\,dW\,dy\,dt\right|\le cT_1, \tag{6-26}$$

and we can conclude that

$$|J_1|+|J_2|+|J_3|+|J_4|\le c(\epsilon,\tilde\epsilon)T_1+\epsilon T_3+\tilde\epsilon T_5, \tag{6-27}$$

where ϵ and $\tilde\epsilon$ are positive degrees of freedom. Next, integrating by parts in J we see that

$$\begin{aligned}
J &= -\sum_j\int_{2\square_0\cap\{w_m=\delta\}} a_{m,j}\partial_{w_j}\tilde v\tilde v\delta^3\zeta^4\,dw\,dY\,dt\\
&\quad+\sum_j\iint_{2\square_0} a_{m,j}(\partial_{w_j}\tilde v)\partial_{w_m}(\tilde v w_m^3\zeta^4)\,dW\,dy\,dt\\
&\quad+\sum_{i\ne m}\sum_j\iint_{2\square_0} a_{i,j}(\partial_{w_j}\tilde v)\partial_{w_i}(\tilde v w_m^3\zeta^4)\,dW\,dy\,dt\\
&= \sum_i\sum_j\iint_{2\square_0} a_{i,j}(\partial_{w_j}\tilde v)(\partial_{w_i}\tilde v)(w_m^3\zeta^4)\,dW\,dy\,dt+\tilde J,
\end{aligned}$$

where

$$\begin{aligned}
\tilde J &:= -\sum_j\int_{2\square_0\cap\{w_m=\delta\}} a_{m,j}(\partial_{w_j}\tilde v)\tilde v\delta^3\zeta^4\,dw\,dY\,dt\\
&\quad+4\sum_i\sum_j\iint_{2\square_0} a_{i,j}(\partial_{w_j}\tilde v)\tilde v\zeta^3\partial_{w_i}\zeta\, w_m^3\zeta^3\,dW\,dy\,dt\\
&\quad+3\sum_j\iint_{2\square_0} a_{m,j}(\partial_{w_j}\tilde v)\tilde v w_m^2\zeta^4\,dW\,dy\,dt.
\end{aligned}$$

Using this notation we see that

$$c^{-1}T_3 \le |\tilde{J}| + |J_1| + |J_2| + |J_3| + |J_4|. \tag{6-28}$$

Furthermore, using (6-9) and Lemma 2.5 it is easy to see that

$$|\tilde{J}| \le c\varrho_0^{\boldsymbol{q}-1} + cT_1 + \epsilon T_3, \tag{6-29}$$

where ϵ is a degree of freedom. Combining (6-27) and (6-29), (i) follows. To prove (ii) we introduce $\tilde{v} = \partial_{y_i} v$. Again using (6-5) we see that $\tilde{v}$ solves

$$\begin{aligned}\nabla_W \cdot (A\nabla_W \tilde{v}) + B \cdot \nabla_W \tilde{v} + \big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)\tilde{v}& \\ = -\nabla_W \cdot (\partial_{y_i} A\nabla_W v) - \partial_{y_i} B \cdot \nabla_W v + \partial_{y_i} P_{\gamma w_m}\psi(w, y, t)\partial_{y_m} v&\end{aligned} \tag{6-30}$$

in U. Arguing similarly to the proof of (i) we derive that

$$T_4 \le c\varrho_0^{\boldsymbol{q}-1} + cT_5 + cT_4^{1/2}T_5^{1/2} + cT_1^{1/2}T_5^{1/2} + cT_4^{1/2}T_5^{1/2} + cT_3^{1/2}T_5^{1/2}. \tag{6-31}$$

Hence,

$$T_4 \le c\varrho_0^{\boldsymbol{q}-1} + c(\epsilon, \tilde{\epsilon})T_5 + \epsilon T_1 + \tilde{\epsilon}T_3, \tag{6-32}$$

where ϵ and $\tilde{\epsilon}$ are positive degrees of freedom. To prove (iii) we have to estimate

$$T_5 = \sum_{i=1}^{m} T_{5,i}, \;\; T_{5,i} := \iint_{2\square_0} (\partial_{y_i} v)(\partial_{y_i} v)w_m^5\zeta^6\, dW\, dy\, dt. \tag{6-33}$$

Note that

$$\begin{aligned}\partial_{y_i} v = -\big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)(\partial_{w_i} v)& \\ +\partial_{w_i}\big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)(v) - (\partial_{w_i} P_{\gamma w_m}\psi(w, y, t))\partial_{y_m} v.&\end{aligned} \tag{6-34}$$

Hence,

$$\begin{aligned}T_{5,i} = &-\iint_{2\square_0} (\partial_{y_i} v)\big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)(\partial_{w_i} v)w_m^5\zeta^6\, dW\, dy\, dt \\ &+\iint_{2\square_0} (\partial_{y_i} v)\partial_{w_i}\big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)(v)w_m^5\zeta^6\, dW\, dy\, dt \\ &-\iint_{2\square_0} (\partial_{y_i} v)\big((\partial_{w_i} P_{\gamma w_m}\psi(w, y, t))\partial_{y_m} v\big)w_m^5\zeta^6\, dW\, dy\, dt \\ =:&\; T_{5,i,1} + T_{5,i,2} + T_{5,i,3}.\end{aligned} \tag{6-35}$$

Using partial integration we immediately see that

$$|T_{5,i,2}| \le c\varrho_0^{\boldsymbol{q}-1} + cT_4^{1/2}T_2^{1/2} + cT_{5,i}^{1/2}T_2^{1/2}. \tag{6-36}$$

Furthermore,

$$
\begin{aligned}
T_{5,i,1} &= \iint_{2\square_0} \big((w, w_m + P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y - \partial_t\big)(\partial_{y_i} v)(\partial_{w_i} v) w_m^5 \zeta^6 \, dW\, dy\, dt \\
&\quad + 6 \iint_{2\square_0} (\partial_{y_i} v)(\partial_{w_i} v) w_m^5 \zeta^5 \big((w, w_m + P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y - \partial_t\big)\zeta \, dW\, dy\, dt \\
&= \iint_{2\square_0} \partial_{y_i}\big((w, w_m + P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y - \partial_t\big)(v)(\partial_{w_i} v) w_m^5 \zeta^6 \, dW\, dy\, dt \\
&\quad - \iint_{2\square_0} (\partial_{y_i} P_{\gamma w_m}\psi(w,y,t))(\partial_{y_m} v)(\partial_{w_i} v) w_m^5 \zeta^6 \, dW\, dy\, dt \\
&\quad + 6 \iint_{2\square_0} (\partial_{y_i} v)(\partial_{w_i} v) w_m^5 \zeta^5 \big((w, w_m + P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y - \partial_t\big)\zeta \, dW\, dy\, dt.
\end{aligned}
$$

Integrating by parts we have

$$
\begin{aligned}
T_{5,i,1} &= -\iint_{2\square_0} \big((w, w_m + P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y - \partial_t\big)(v)(\partial_{w_i y_i} v) w_m^5 \zeta^6 \, dW\, dy\, dt \\
&\quad - 6 \iint_{2\square_0} \big((w, w_m + P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y - \partial_t\big)(v)(\partial_{w_i} v) w_m^5 \zeta^5 \partial_{y_i}\zeta \, dW\, dy\, dt \\
&\quad - \iint_{2\square_0} (\partial_{y_i} P_{\gamma w_m}\psi(w,y,t))(\partial_{y_m} v)(\partial_{w_i} v) w_m^5 \zeta^6 \, dW\, dy\, dt \\
&\quad + 6 \iint_{2\square_0} (\partial_{y_i} v)(\partial_{w_i} v) w_m^5 \zeta^5 \big((w, w_m + P_{\gamma w_m}\psi(w,y,t))\cdot\nabla_Y - \partial_t\big)\zeta \, dW\, dy\, dt.
\end{aligned}
$$

Hence,

$$
|T_{5,i,1}| \le cT_2^{1/2}T_4^{1/2} + cT_1^{1/2}T_2^{1/2} + cT_1^{1/2}(T_{5,i}^{1/2} + T_{5,m}^{1/2}) \tag{6-37}
$$

and

$$
|T_{5,i}| \le c\varrho_0^{\boldsymbol{q}-1} + cT_2^{1/2}T_4^{1/2} + cT_1^{1/2}T_2^{1/2} + cT_1^{1/2}(T_{5,i}^{1/2} + T_{5,m}^{1/2}) + cT_2^{1/2}T_{5,i}^{1/2} + |T_{5,i,3}|. \tag{6-38}
$$

We now first consider the case $i = m$. Using Remark 2.3 and (6-38) we immediately see that

$$
T_{5,m} \le c\varrho_0^{\boldsymbol{q}-1} + cT_1 + cT_2 + \epsilon T_4, \tag{6-39}
$$

where ϵ is a positive degree of freedom. Consider now $i \ne m$. Then, using (6-38) we have

$$
|T_{5,i}| \le c\varrho_0^{\boldsymbol{q}-1} + cT_2^{1/2}T_4^{1/2} + cT_1^{1/2}T_2^{1/2} + cT_1^{1/2}(T_{5,i}^{1/2} + T_{5,m}^{1/2}) + cT_2^{1/2}T_{5,i}^{1/2} + cT_{5,i}^{1/2}T_{5,m}^{1/2}, \tag{6-40}
$$

and hence

$$
T_{5,i} \le c\varrho_0^{\boldsymbol{q}-1} + cT_1 + cT_2 + \epsilon T_4 + cT_{5,m}. \tag{6-41}
$$

Using (6-39) and (6-41) we can now complete the proof of (iii) and the lemma. $\square$

6C. *Proof of Lemma 6.2*. To start the proof of Lemma 6.2 we first use the equation in (6-5) and write

$$T_2 = -\iint_{2\square_0} \big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)(v)\big(\nabla_W \cdot (A\nabla_W v) + B \cdot \nabla_W v\big)\zeta^4 w_m^3 \, dW\, dy\, dt$$

and

$$T_2 = T_{21} + T_{22} + T_{23} + T_{24}, \tag{6-42}$$

where

$$\begin{aligned}
T_{21} &:= -\sum_j \iint_{2\square_0} \big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big) v \partial_{w_m}(a_{m,j}\partial_{w_j} v) w_m^3 \zeta^4 \, dW\, dy\, dt,\\
T_{22} &:= -\sum_{i\neq m} \iint_{2\square_0} \big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big) v \partial_{w_i}(a_{i,m}\partial_{w_m} v) w_m^3 \zeta^4 \, dW\, dy\, dt,\\
T_{23} &:= -\sum_{i\neq m}\sum_{j\neq m} \iint_{2\square_0} \big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big) v \partial_{w_i}(a_{i,j}\partial_{w_j} v) w_m^3 \zeta^4 \, dW\, dy\, dt,\\
T_{24} &:= -\sum_i \iint_{2\square_0} \big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big) v b_i \partial_{w_i} v w_m^3 \zeta^4 \, dW\, dy\, dt.
\end{aligned}$$

Using (6-9) we immediately see that

$$|T_{21}| + |T_{22}| + |T_{24}| \leq cT_1^{1/2}T_2^{1/2} + cT_3^{1/2}T_2^{1/2}. \tag{6-43}$$

Next, focusing on T_{23}, and integrating by parts with respect to w_i, we see that

$$\begin{aligned}
T_{23} &= \sum_{i\neq m}\sum_{j\neq m} \iint_{2\square_0} \partial_{w_i}\big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big) v (a_{i,j}\partial_{w_j} v) w_m^3 \zeta^4 \, dW\, dy\, dt\\
&\quad + \sum_{i\neq m}\sum_{j\neq m} \iint_{2\square_0} \big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big) v (a_{i,j}\partial_{w_j} v)\partial_{w_i}(\zeta) w_m^3 \zeta^3 \, dW\, dy\, dt\\
&=: T_{231} + T_{232},
\end{aligned}$$

and that $|T_{232}| \leq T_1^{1/2}T_2^{1/2}$. Furthermore

$$\begin{aligned}
T_{231} &= \sum_{i\neq m}\sum_{j\neq m} \iint_{2\square_0} (\partial_{y_i} v)(a_{i,j}\partial_{w_j} v) w_m^3 \zeta^4 \, dW\, dy\, dt\\
&\quad + \sum_{i\neq m}\sum_{j\neq m} \iint_{2\square_0} \big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)(\partial_{w_i} v)(a_{i,j}\partial_{w_j} v) w_m^3 \zeta^4 \, dW\, dy\, dt\\
&\quad + \sum_{i\neq m}\sum_{j\neq m} \iint_{2\square_0} (\partial_{w_i} P_{\gamma w_m}\psi(w, y, t))(\partial_{y_m} v)(a_{i,j}\partial_{w_j} v) w_m^3 \zeta^4 \, dW\, dy\, dt\\
&=: T_{2311} + T_{2312} + T_{2313}.
\end{aligned}$$

Then

$$|T_{2311}| + |T_{2313}| \leq cT_5^{1/2}T_1^{1/2}. \tag{6-44}$$

To estimate T_{2312} we lift the vector field $\big((w, w_m + P_{\gamma w_m}\psi(w, y, t)) \cdot \nabla_Y - \partial_t\big)$ through partial integration and use the symmetry of the matrix $\{a_{i,j}\}$ to see that

$$\begin{aligned}2T_{2312} = &-\sum_{i\neq m}\sum_{j\neq m}\iint_{2\square_0} (\partial_{w_i} v)\Big(\big((w, w_m+P_{\gamma w_m}\psi(w, y, t))\cdot\nabla_Y-\partial_t\big)(a_{i,j})\partial_{w_j} v\Big) w_m^3\zeta^4\, dW\, dy\, dt\\ &-4\sum_{i\neq m}\sum_{j\neq m}\iint_{2\square_0} (\partial_{w_i} v)(a_{i,j}\partial_{w_j} v)\big((w, w_m+P_{\gamma w_m}\psi(w, y, t))\cdot\nabla_Y-\partial_t\big)(\zeta) w_m^3\zeta^3\, dW\, dy\, dt\\ =:&\ T_{23121}+T_{23122}.\end{aligned}$$

Then, by familiar arguments,

$$|T_{23121}| + |T_{23122}| \leq cT_1. \tag{6-45}$$

Putting all estimates together we can conclude that

$$T_2 \leq |T_{21}| + |T_{22}| + |T_{23}| + |T_{24}| \leq cT_1 + cT_1^{1/2}T_2^{1/2} + cT_2^{1/2}T_3^{1/2} + cT_1^{1/2}T_5^{1/2}. \tag{6-46}$$

Hence

$$T_2 \leq c(T_1 + T_3) + \epsilon_1 T_5, \tag{6-47}$$

where ϵ_1 is a positive degree of freedom. Now, using Lemma 6.3 we see, given positive degrees of freedom $\epsilon_2 - \epsilon_5$, that there exists $c = c(N, M_1, M_2, \epsilon_2 - \epsilon_5)$, $1 \leq c < \infty$, such that

$$\begin{aligned} T_3 &\leq c\varrho_0^{q-1} + cT_1 + \epsilon_2 T_5,\\ T_4 &\leq c\varrho_0^{q-1} + cT_5 + \epsilon_3 T_1 + \epsilon_4 T_3,\\ T_5 &\leq c\varrho_0^{q-1} + cT_1 + cT_2 + \epsilon_5 T_4.\end{aligned} \tag{6-48}$$

Using the estimates on the last two lines in (6-48) we see that

$$T_5 \leq c\varrho_0^{q-1} + cT_1 + cT_2 + \epsilon_6 T_3, \tag{6-49}$$

where again ϵ_6 is a degree of freedom. Using (6-49) in the first estimate in (6-48) we deduce that

$$T_3 \leq c\varrho_0^{q-1} + cT_1 + c\epsilon_2 T_2 + \epsilon_2\epsilon_6 T_3, \tag{6-50}$$

and hence, consuming ϵ_6,

$$\begin{aligned} T_3 &\leq c\varrho_0^{q-1} + cT_1 + \epsilon_7 T_2,\\ T_5 &\leq c\varrho_0^{q-1} + cT_1 + cT_2,\end{aligned} \tag{6-51}$$

for yet another degree of freedom ϵ_7. Putting the estimates from (6-51) into (6-47), we deduce that

$$T_2 \leq c_1\varrho_0^{q-1} + c_1 T_1 + c_2 T_2(\epsilon_1 + \epsilon_7), \tag{6-52}$$

where $c_1 = c_1(N, M_1, M_2, \epsilon_1, \epsilon_7)$ and $c_2 = c_2(N, M_1, M_2)$. Elementary manipulations now complete the proof of the lemma. □

Acknowledgements

The author would like to express his gratitude to Sergio Polidoro for introducing him to operators of Kolmogorov type a few years ago and for fruitful collaboration. The author also thanks Steve Hofmann for communications.

References

[Barucci et al. 2001] E. Barucci, S. Polidoro, and V. Vespri, "Some results on partial differential equations and Asian options", *Math. Models Methods Appl. Sci.* **11**:3 (2001), 475–497. MR Zbl

[Bernardin et al. 2009] F. Bernardin, M. Bossy, C. Chauvin, P. Drobinski, A. Rousseau, and T. Salameh, "Stochastic downscaling method: application to wind refinement", *Stoch. Environ. Res. Risk Assess.* **23**:6 (2009), 851–859. MR

[Bernardin et al. 2010] F. Bernardin, M. Bossy, C. Chauvin, J.-F. Jabir, and A. Rousseau, "Stochastic Lagrangian method for downscaling problems in computational fluid dynamics", *M2AN Math. Model. Numer. Anal.* **44**:5 (2010), 885–920. MR Zbl

[Bossy et al. 2011] M. Bossy, J.-F. Jabir, and D. Talay, "On conditional McKean Lagrangian stochastic models", *Probab. Theory Related Fields* **151**:1-2 (2011), 319–351. MR Zbl

[Chauvin et al. 2010] C. Chauvin, F. Bernardin, M. Bossy, and A. Rousseau, "Wind simulation refinement: some new challenges for particle methods", pp. 765–770 in *Progress in Industrial Mathematics at ECMI 2008*, edited by A. D. Fitt et al., Mathematics in Industry **15**, 2010. Zbl

[Christ 1990] M. Christ, "A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral", *Colloq. Math.* **60/61**:2 (1990), 601–628. MR Zbl

[Cinti et al. 2010] C. Cinti, K. Nyström, and S. Polidoro, "A note on Harnack inequalities and propagation sets for a class of hypoelliptic operators", *Potential Anal.* **33**:4 (2010), 341–354. MR Zbl

[Cinti et al. 2012] C. Cinti, K. Nyström, and S. Polidoro, "A boundary estimate for non-negative solutions to Kolmogorov operators in non-divergence form", *Ann. Mat. Pura Appl.* (4) **191**:1 (2012), 1–23. MR Zbl

[Cinti et al. 2013] C. Cinti, K. Nyström, and S. Polidoro, "A Carleson-type estimate in Lipschitz type domains for non-negative solutions to Kolmogorov operators", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **12**:2 (2013), 439–465. MR Zbl

[Citti and Sarti 2006] G. Citti and A. Sarti, "A cortical based model of perceptual completion in the roto-translation space", *J. Math. Imaging Vision* **24**:3 (2006), 307–326. MR Zbl

[Citti and Sarti 2014] G. Citti and A. Sarti (editors), *Neuromathematics of vision*, Springer, 2014. MR Zbl

[Coifman and Fefferman 1974] R. R. Coifman and C. Fefferman, "Weighted norm inequalities for maximal functions and singular integrals", *Studia Math.* **51**:3 (1974), 241–250. MR Zbl

[Coifman and Weiss 1971] R. R. Coifman and G. Weiss, *Analyse harmonique non-commutative sur certains espaces homogènes: étude de certaines intégrales singulières*, Lecture Notes in Mathematics **242**, Springer, 1971. MR Zbl

[Fabes and Safonov 1997] E. B. Fabes and M. V. Safonov, "Behavior near the boundary of positive solutions of second order parabolic equations", *J. Fourier Anal. Appl.* **3**:Special Issue (1997), 871–882. MR Zbl

[Fabes et al. 1986] E. B. Fabes, N. Garofalo, and S. Salsa, "A backward Harnack inequality and Fatou theorem for nonnegative solutions of parabolic equations", *Illinois J. Math.* **30**:4 (1986), 536–565. MR Zbl

[Fabes et al. 1999] E. B. Fabes, M. V. Safonov, and Y. Yuan, "Behavior near the boundary of positive solutions of second order parabolic equations, II", *Trans. Amer. Math. Soc.* **351**:12 (1999), 4947–4961. MR Zbl

[Garofalo and Lanconelli 1990] N. Garofalo and E. Lanconelli, "Level sets of the fundamental solution and Harnack inequality for degenerate equations of Kolmogorov type", *Trans. Amer. Math. Soc.* **321**:2 (1990), 775–792. MR Zbl

[Hofmann 1997] S. Hofmann, "Parabolic singular integrals of Calderón-type, rough operators, and caloric layer potentials", *Duke Math. J.* **90**:2 (1997), 209–259. MR Zbl

[Hofmann and Lewis 1996] S. Hofmann and J. L. Lewis, "L^2 solvability and representation by caloric layer potentials in time-varying domains", *Ann. of Math.* (2) **144**:2 (1996), 349–420. MR Zbl

[Hofmann and Lewis 1999] S. Hofmann and J. L. Lewis, "The L^p regularity problem for the heat equation in non-cylindrical domains", *Illinois J. Math.* **43**:4 (1999), 752–769. MR Zbl

[Hofmann and Lewis 2001a] S. Hofmann and J. L. Lewis, *The Dirichlet problem for parabolic operators with singular drift terms*, Mem. Amer. Math. Soc. **719**, Amer. Math. Soc., Providence, RI, 2001. MR Zbl

[Hofmann and Lewis 2001b] S. Hofmann and J. L. Lewis, "Square functions of Calderón type and applications", *Rev. Mat. Iberoamericana* **17**:1 (2001), 1–20. MR Zbl

[Hofmann and Lewis 2005] S. Hofmann and J. L. Lewis, "The L^p Neumann problem for the heat equation in non-cylindrical domains", *J. Funct. Anal.* **220**:1 (2005), 1–54. MR Zbl

[Hofmann et al. 2003] S. Hofmann, J. L. Lewis, and K. Nyström, "Existence of big pieces of graphs for parabolic problems", *Ann. Acad. Sci. Fenn. Math.* **28**:2 (2003), 355–384. MR Zbl

[Hofmann et al. 2004] S. Hofmann, J. L. Lewis, and K. Nyström, "Caloric measure in parabolic flat domains", *Duke Math. J.* **122**:2 (2004), 281–346. MR Zbl

[Hofmann et al. 2017] S. Hofmann, D. Mitrea, M. Mitrea, and A. J. Morris, *L^p-square function estimates on spaces of homogeneous type and on uniformly rectifiable sets*, Mem. Amer. Math. Soc. **1159**, Amer. Math. Soc., Providence, RI, 2017. MR

[Hörmander 1967] L. Hörmander, "Hypoelliptic second order differential equations", *Acta Math.* **119** (1967), 147–171. MR Zbl

[Kenig et al. 2016] C. Kenig, B. Kirchheim, J. Pipher, and T. Toro, "Square functions and the A_∞ property of elliptic measures", *J. Geom. Anal.* **26**:3 (2016), 2383–2410. MR Zbl

[Kolmogorov 1934] A. Kolmogorov, "Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)", *Ann. of Math.* (2) **35**:1 (1934), 116–117. MR Zbl

[Lanconelli and Polidoro 1994] E. Lanconelli and S. Polidoro, "On a class of hypoelliptic evolution operators", *Rend. Sem. Mat. Univ. Politec. Torino* **52**:1 (1994), 29–63. MR Zbl

[Lewis and Murray 1995] J. L. Lewis and M. A. M. Murray, *The method of layer potentials for the heat equation in time-varying domains*, Mem. Amer. Math. Soc. **545**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Lewis and Silver 1988] J. L. Lewis and J. Silver, "Parabolic measure and the Dirichlet problem for the heat equation in two dimensions", *Indiana Univ. Math. J.* **37**:4 (1988), 801–839. MR Zbl

[Lieberman 1996] G. M. Lieberman, *Second order parabolic differential equations*, World Scientific Publishing Co., River Edge, NJ, 1996. MR Zbl

[Manfredini 1997] M. Manfredini, "The Dirichlet problem for a class of ultraparabolic equations", *Adv. Differential Equations* **2**:5 (1997), 831–866. MR Zbl

[Nyström 1997] K. Nyström, "The Dirichlet problem for second order parabolic operators", *Indiana Univ. Math. J.* **46**:1 (1997), 183–245. MR Zbl

[Nyström and Polidoro 2016] K. Nyström and S. Polidoro, "Kolmogorov–Fokker–Planck equations: comparison principles near Lipschitz type boundaries", *J. Math. Pures Appl.* (9) **106**:1 (2016), 155–202. MR Zbl

[Pascucci 2011] A. Pascucci, *PDE and martingale methods in option pricing*, Bocconi & Springer Series **2**, Springer, 2011. MR Zbl

[Pope 2000] S. B. Pope, *Turbulent flows*, Cambridge University Press, 2000. MR Zbl

[Safonov and Yuan 1999] M. V. Safonov and Y. Yuan, "Doubling properties for second order parabolic equations", *Ann. of Math.* (2) **150**:1 (1999), 313–327. MR Zbl

[Salsa 1981] S. Salsa, "Some properties of nonnegative solutions of parabolic differential operators", *Ann. Mat. Pura Appl.* (4) **128** (1981), 193–206. MR Zbl

Received 1 Feb 2017. Accepted 17 Jun 2017.

KAJ NYSTRÖM: kaj.nystrom@math.uu.se
Department of Mathematics, Uppsala University, Uppsala, Sweden

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 10, No. 7, 2017
dx.doi.org/10.2140/apde.2017.10.1757

L^2-BETTI NUMBERS OF RIGID C^*-TENSOR CATEGORIES AND DISCRETE QUANTUM GROUPS

DAVID KYED, SVEN RAUM, STEFAAN VAES AND MATTHIAS VALVEKENS

We compute the L^2-Betti numbers of the free C^*-tensor categories, which are the representation categories of the universal unitary quantum groups $A_u(F)$. We show that the L^2-Betti numbers of the dual of a compact quantum group $\mathbb{G}$ are equal to the L^2-Betti numbers of the representation category $\mathrm{Rep}(\mathbb{G})$ and thus, in particular, invariant under monoidal equivalence. As an application, we obtain several new computations of L^2-Betti numbers for discrete quantum groups, including the quantum permutation groups and the free wreath product groups. Finally, we obtain upper bounds for the first L^2-Betti number in terms of a generating set of a C^*-tensor category.

1. Introduction

The framework of rigid C^*-tensor categories unifies a number of structures encoding various kinds of quantum symmetry, including standard invariants of Jones' subfactors, representation categories of compact quantum groups, in particular of q-deformations of compact simple Lie groups, and ordinary discrete groups. In several respects, rigid C^*-tensor categories are quantum analogues of discrete groups.

Using this point of view, the unitary representation theory for rigid C^*-tensor categories was introduced in [Popa and Vaes 2015]. This allowed for the definition of typical geometric group theory properties like the Haagerup property and property (T) intrinsically for standard invariants of subfactors and for rigid C^*-tensor categories. It was then proved in [Popa and Vaes 2015], using [Arano 2016; De Commer et al. 2014], that the Temperley–Lieb–Jones category $\mathrm{Rep}(\mathrm{SU}_q(2))$ has the Haagerup property, while $\mathrm{Rep}(\mathrm{SU}_q(3))$ has Kazhdan's property (T). Equivalent formulations of the unitary representation theory of a rigid C^*-tensor category were found in [Neshveyev and Yamashita 2016; Ghosh and Jones 2016] and are introduced below.

In [Popa et al. 2017], a comprehensive (co)homology theory for standard invariants of subfactors and rigid C^*-tensor categories was introduced. Taking the appropriate Murray–von Neumann dimension for (co)homology with L^2-coefficients, this provides a definition of L^2-Betti numbers.

The first goal of this article is to compute the L^2-Betti numbers for the representation category $\mathcal{C}$ of a free unitary quantum group $A_u(F)$. Here, $A_u(F)$ is the universal compact quantum group (in the sense of Woronowicz) generated by a single irreducible unitary representation. As a C^*-tensor category, $\mathcal{C}$ is the free rigid C^*-tensor category generated by a single irreducible object u. The irreducible objects of $\mathcal{C}$

Kyed is supported by the Villum foundation grant 7423. Vaes and Valvekens are supported by European Research Council Consolidator Grant 614195, and by long term structural funding — a Methusalem grant of the Flemish Government.
MSC2010: primary 46L37; secondary 16E40, 18D10, 20G42.
Keywords: L^2-Betti numbers, rigid C^*-tensor categories, discrete quantum groups, subfactors, compact quantum groups.

are then labeled by all words in u and $\bar{u}$ and can thus be identified with the free monoid $\mathbb{N} * \mathbb{N}$. We prove that $\beta_1^{(2)}(\mathcal{C}) = 1$ and that the other L^2-Betti numbers vanish.

For compact quantum groups $\mathbb{G}$ of Kac type (a unimodularity assumption that is equivalent with the traciality of the Haar state), the L^2-Betti numbers $\beta_n^{(2)}(\widehat{\mathbb{G}})$ of the dual discrete quantum group $\widehat{\mathbb{G}}$ were defined in [Kyed 2008]. The second main result of our paper is that these L^2-Betti numbers only depend on the representation category of $\mathbb{G}$ and are given by $\beta_n^{(2)}(\mathrm{Rep}(\mathbb{G}))$. This is surprising for two reasons. The L^2-Betti numbers $\beta_n^{(2)}(\mathrm{Rep}(\mathbb{G}))$ are well defined for all compact quantum groups, without a unimodularity assumption. And secondly, taking arbitrary coefficients instead of L^2-cohomology, there is no possible identification between the (co)homology of $\widehat{\mathbb{G}}$ and $\mathrm{Rep}(\mathbb{G})$. Indeed, by [Collins et al. 2009, Theorem 3.2], homology with trivial coefficients distinguishes between the quantum groups $A_o(k)$, but does not distinguish between their representation categories $\mathrm{Rep}(A_o(k))$ by Corollary 6.2 below. As an application, we compute the L^2-Betti numbers for several families of Kac-type discrete quantum groups, including the duals of the quantum permutation groups S_m^+, the hyperoctahedral series H_m^{s+} of [Banica and Vergnioux 2009] and the free wreath product groups $\mathbb{H} \wr_* \mathbb{F}$ of [Bichon 2004].

One of the equivalent definitions in [Popa et al. 2017] for the (co)homology of a rigid C^*-tensor category $\mathcal{C}$ is given by the Hochschild (co)homology of the associated tube algebra $\mathcal{A}$ together with its counit $\varrho : \mathcal{A} \to \mathbb{C}$ as the augmentation. In [Neshveyev and Yamashita 2015], it is proved that when $\mathcal{C} = \mathrm{Rep}(\mathbb{G})$ is the representation category of a compact quantum group $\mathbb{G}$, then the tube algebra $\mathcal{A}$ is strongly Morita equivalent with the Drinfeld double algebra of $\mathbb{G}$. This is one of the main tools in our paper. As a side result, applying this to $\mathbb{G} = \mathrm{SU}_q(2)$, so that $\mathcal{C}$ becomes the Temperley–Lieb–Jones category TLJ, we can transfer the resolution of [Bichon 2013] to a length-3 resolution for the tube algebra of TLJ; see Theorem 6.1. This allows us in particular to compute the (co)homology of TLJ with trivial coefficients, giving $\mathbb{C}$ in degree 0 and degree 3, and giving 0 in all other degrees. This completes the computation in [Popa et al. 2017, Proposition 9.13], which went up to degree 2, and this was also obtained in an unpublished note of Y. Arano.

In the second part of this paper, we focus on the first L^2-Betti number of a rigid C^*-tensor category. For an infinite group Γ generated by n elements $g_1, \dots, g_n$, it is well known that $\beta_1^{(2)}(\Gamma) \le n-1$. The reason for this is that a 1-cocycle on Γ is completely determined by the values it takes on the generators $g_1, \dots, g_n$. In Section 7, we explain how to realize the first cohomology of a rigid C^*-tensor category $\mathcal{C}$ by a space of maps D, similar to derivations, and prove that D is indeed determined by its values on a generating set of irreducible objects. We then deduce an upper bound for $\beta_1^{(2)}(\mathcal{C})$ and show in Section 8 that this upper bound is precisely reached for the universal (or free) category $\mathcal{C} = \mathrm{Rep}(A_u(F))$.

2. Preliminaries

2A. ***The tube algebra of a rigid C*-tensor category.*** Let $\mathcal{C}$ be a *rigid C^*-tensor category*, i.e., a C^*-tensor category with irreducible unit object $\varepsilon \in \mathcal{C}$ such that every object $\alpha \in \mathcal{C}$ has a conjugate $\bar{\alpha} \in \mathcal{C}$. In particular, this implies that every object in $\mathcal{C}$ decomposes into finitely many irreducibles. The essential results on rigid C^*-tensor categories, which we will use without further reference, are covered in [Neshveyev and Tuset 2013, Chapter 2]. For $\alpha, \beta \in \mathcal{C}$, we denote the (necessarily finite-dimensional) Banach space of morphisms $\alpha \to \beta$ by (β, α).

The set of isomorphism classes of irreducible objects of $\mathcal{C}$ will be denoted by $\mathrm{Irr}(\mathcal{C})$. In what follows, we do not distinguish between irreducible objects and their respective isomorphism classes and we fix representatives for all isomorphism classes once and for all. Additionally, we always identify (α, α) with $\mathbb{C}$ when $\alpha \in \mathrm{Irr}(\mathcal{C})$. The *multiplicity* of γ in α when $\alpha \in \mathcal{C}$ and $\gamma \in \mathrm{Irr}(\mathcal{C})$ is defined by

$$\mathrm{mult}(\gamma, \alpha) = \dim_{\mathbb{C}}(\alpha, \gamma).$$

For $\alpha, \beta \in \mathcal{C}$, we write $\beta \prec \alpha$ whenever β is isomorphic with a subobject of α. When there is no danger of confusion, we denote the tensor product of α and β by $\alpha\beta$.

The rigidity assumption says that every object $\alpha \in \mathcal{C}$ admits a *solution to the conjugate equations* [Neshveyev and Tuset 2013, Section 2.2], i.e., an object $\bar{\alpha} \in \mathcal{C}$ and a pair of morphisms $s_\alpha \in (\alpha\bar{\alpha}, \varepsilon)$ and $t_\alpha \in (\bar{\alpha}\alpha, \varepsilon)$ satisfying the relations

$$(t_\alpha^* \otimes 1)(1 \otimes s_\alpha) = 1 \quad \text{and} \quad (s_\alpha^* \otimes 1)(1 \otimes t_\alpha) = 1.$$

A *standard* solution for the conjugate equations for $\alpha \in \mathcal{C}$ additionally satisfies

$$s_\alpha^*(T \otimes 1)s_\alpha = t_\alpha^*(1 \otimes T)t_\alpha$$

for all $T \in (\alpha, \alpha)$. The adjoint object $\bar{\alpha}$ and the standard solutions for the conjugate equations are unique up to unitary equivalence. Throughout this article, we always fix standard solutions for all $\alpha \in \mathrm{Irr}(\mathcal{C})$, and extend by naturality to arbitrary objects $\alpha \in \mathcal{C}$; see [Neshveyev and Tuset 2013, Definition 2.2.14]. The positive real number defined by $d(\alpha) = t_\alpha^* t_\alpha = s_\alpha^* s_\alpha$ is referred to as the *quantum dimension* of α.

These standard solutions also give rise to canonical tracial functionals Tr_α on (α, α) via

$$\mathrm{Tr}_\alpha(T) = s_\alpha^*(T \otimes 1)s_\alpha = t_\alpha^*(1 \otimes T)t_\alpha.$$

Note that these traces are typically not normalized, since $\mathrm{Tr}_\alpha(1) = d(\alpha)$. It is sometimes convenient to work with the *partial traces* defined by

$$\begin{aligned} \mathrm{Tr}_\alpha \otimes \mathrm{id} : (\alpha\beta, \alpha\gamma) \to (\beta, \gamma), &\quad T \mapsto (t_\alpha^* \otimes 1)(1 \otimes T)(t_\alpha \otimes 1), \\ \mathrm{id} \otimes \mathrm{Tr}_\alpha : (\beta\alpha, \gamma\alpha) \to (\beta, \gamma), &\quad T \mapsto (1 \otimes s_\alpha^*)(T \otimes 1)(1 \otimes s_\alpha), \end{aligned}$$

for $\alpha, \beta, \gamma \in \mathcal{C}$. These satisfy $\mathrm{Tr}_\beta \circ (\mathrm{Tr}_\alpha \otimes \mathrm{id}) = \mathrm{Tr}_{\alpha\beta} = \mathrm{Tr}_\alpha \circ (\mathrm{id} \otimes \mathrm{Tr}_\beta)$. For all $\alpha, \beta \in \mathcal{C}$, the categorical traces induce an inner product on (α, β), given by

$$\langle T, S \rangle = \mathrm{Tr}_\alpha(TS^*) = \mathrm{Tr}_\beta(S^*T). \tag{2-1}$$

Throughout, the notation $\mathrm{onb}(\alpha, \beta)$ will refer to some choice of orthonormal basis of (α, β) with respect to this inner product. Finally, the standard solutions of the conjugate equations induce the *Frobenius reciprocity maps*, which are the unitary isomorphisms given by

$$\begin{aligned} (\alpha\beta, \gamma) \to (\alpha, \gamma\bar{\beta}), &\quad T \mapsto (1 \otimes s_\beta^*)(T \otimes 1), \\ (\alpha\beta, \gamma) \to (\beta, \bar{\alpha}\gamma), &\quad T \mapsto (t_\alpha^* \otimes 1)(1 \otimes T), \end{aligned} \tag{2-2}$$

where $\alpha, \beta, \gamma \in \mathrm{Irr}(\mathcal{C})$.

The *tube algebra* $\mathcal{A}$ of a rigid C^*-tensor category was first defined by Ocneanu [1994] for categories with finitely many irreducibles. For convenience, we recall some of the exposition from [Popa et al. 2017]. The tube algebra is defined by the vector space direct sum

$$\mathcal{A} = \bigoplus_{i,j,\alpha\in\mathrm{Irr}(\mathcal{C})} (i\alpha, \alpha j).$$

For general $\alpha \in \mathcal{C}$ and $i, j \in \mathrm{Irr}(\mathcal{C})$, a morphism $V \in (i\alpha, \alpha j)$ also defines an element of $\mathcal{A}$ via

$$V \mapsto \sum_{\gamma\in\mathrm{Irr}(\mathcal{C})} d(\gamma) \sum_{W\in\mathrm{onb}(\alpha,\gamma)} (1 \otimes W^*) V (W \otimes 1). \tag{2-3}$$

It should be noted that this map is generally not an embedding of $(i\alpha, \alpha j)$ into $\mathcal{A}$. One easily checks that $\mathcal{A}$ is a $*$-algebra for the operations

$$\begin{aligned} V \cdot W &= \delta_{j,j'} (V \otimes 1)(1 \otimes W) \in (i\alpha\beta, \alpha\beta k), \\ V^{\#} &= (t_\alpha^* \otimes 1 \otimes 1)(1 \otimes V^* \otimes 1)(1 \otimes 1 \otimes s_\alpha) \in (j\bar{\alpha}, \bar{\alpha} i), \end{aligned}$$

where $V \in (i\alpha, \alpha j)$, $W \in (j'\beta, \beta k)$ and where the map in (2-3) is used to view $V \cdot W$ as an element of $\mathcal{A}$. We follow the notational convention from [Popa et al. 2017] and explicitly denote the tube algebra operations by $\cdot$ and $^{\#}$, to avoid confusion with composition and adjunction of morphisms. It should be noted that $\mathcal{A}$ is not unital, unless $\mathrm{Irr}(\mathcal{C})$ is finite.

For $i \in \mathrm{Irr}(\mathcal{C})$, the identity map on i is an element of $(i\varepsilon, \varepsilon i)$. So it can be considered as an element $p_i \in \mathcal{A}$. As the notation suggests, p_i is a self-adjoint idempotent in $\mathcal{A}$, and it is easy to see that $p_i \cdot V \cdot p_j = \delta_{ik}\delta_{jk'} V$ when $V \in (k\alpha, \alpha k')$. The corner $p_i \cdot \mathcal{A} \cdot p_i$ is a unital $*$-algebra and the projections p_i, $i \in \mathrm{Irr}(\mathcal{C})$, serve as local units for $\mathcal{A}$. In particular, for all purposes of homological algebra, we can work with $\mathcal{A}$ as if it were a unital algebra.

The corner $p_\varepsilon \cdot \mathcal{A} \cdot p_\varepsilon$ is canonically isomorphic to the *fusion $*$-algebra* $\mathbb{C}[\mathcal{C}]$. This algebra is formed by taking the free vector space over $\mathrm{Irr}(\mathcal{C})$, and defining multiplication by the fusion rules, i.e.,

$$\alpha \cdot \beta = \sum_{\gamma\in\mathrm{Irr}(\mathcal{C})} \mathrm{mult}(\gamma, \alpha \otimes \beta)\gamma.$$

The involution on $\mathbb{C}[\mathcal{C}]$ is given by conjugation in $\mathcal{C}$.

The tube algebra comes with a faithful trace τ; see [Popa et al. 2017, Proposition 3.10]. For $V \in (i\alpha, \alpha j)$ with $i, j, \alpha \in \mathrm{Irr}(\mathcal{C})$, this trace is given by

$$\tau(V) = \begin{cases} \mathrm{Tr}_i(V), & i = j,\ \alpha = \varepsilon, \\ 0, & \text{otherwise.} \end{cases}$$

In [Popa et al. 2017], it is also shown that every involutive action of $\mathcal{A}$ on a pre-Hilbert space is automatically by bounded operators. In particular, this allows us to define a von Neumann algebra $\mathcal{A}''$ by considering the faithful action of $\mathcal{A}$ on $L^2(\mathcal{A}, \tau)$ by left multiplication, and then taking the bicommutant. Additionally, the trace τ uniquely extends to a faithful normal semifinite trace on $\mathcal{A}''$.

For $i, j, \alpha \in \mathrm{Irr}(\mathcal{C})$, we now have two inner products on $(i\alpha, \alpha j)$, related by

$$\mathrm{Tr}_{\alpha j}(W^* V) = d(\alpha)\, \tau(W^\# \cdot V).$$

We will however always work with the inner product given by $\mathrm{Tr}_{\alpha j}(W^* V)$, because it is compatible with the inner product in (2-1), which is defined on all spaces of intertwiners and which makes the Frobenius reciprocity maps (2-2) unitary.

2B. ***Representation theory for rigid C*-tensor categories.*** The unitary representation theory for rigid C^*-tensor categories was introduced in [Popa and Vaes 2015] and several equivalent formulations were found in [Neshveyev and Yamashita 2016; Ghosh and Jones 2016; Popa et al. 2017]. Following [Ghosh and Jones 2016], a unitary representation of $\mathcal{C}$ is given by a *nondegenerate $*$-representation of the tube algebra* of $\mathcal{C}$. Following [Neshveyev and Yamashita 2016], a unitary representation of $\mathcal{C}$ is given by a *unitary half-braiding on an ind-object of* $\mathcal{C}$, i.e., an object in the *unitary Drinfeld center* $\mathcal{Z}(\text{ind-}\mathcal{C})$. Here, the category ind-$\mathcal{C}$ may be thought of as a completion of $\mathcal{C}$ with infinite direct sums, giving rise to a (nonrigid) C^*-tensor category. A *unitary half-braiding* on an ind-object $X \in \text{ind-}\mathcal{C}$ is a natural unitary isomorphism $\sigma_- : - \otimes X \to X \otimes -$ that satisfies the half-braiding condition

$$\sigma_{Y\otimes Z} = (\sigma_Y \otimes 1)(1 \otimes \sigma_Z)$$

for all $Y, Z \in \text{ind-}\mathcal{C}$. The collection of unitary half-braidings on ind-$\mathcal{C}$ is denoted by $\mathcal{Z}(\text{ind-}\mathcal{C})$. We refer to [Neshveyev and Yamashita 2016] for rigorous definitions and basic properties of these objects.

By [Popa et al. 2017, Proposition 3.14], there is the following bijective correspondence between nondegenerate right Hilbert $\mathcal{A}$-modules $\mathcal{K}$ and unitary half-braidings (X, σ). Given $(X, \sigma) \in \mathcal{Z}(\text{ind-}\mathcal{C})$, one defines $\mathcal{K}$ as the Hilbert space direct sum of the Hilbert spaces (X, i), $i \in \mathrm{Irr}(\mathcal{C})$. To turn $\mathcal{K}$ into a right $\mathcal{A}$-module, we let $V \in (i\alpha, \alpha j)$ act on a vector $\xi \in (X, i')$ by

$$\xi \cdot V = \delta_{ii'} (\mathrm{Tr}_\alpha \otimes \mathrm{id})(\sigma_\alpha^*(\xi \otimes 1)V) \in (X, j). \tag{2-4}$$

In particular, we see that $\mathcal{K} \cdot p_i = (X, i)$.

2C. ***(Co)homology and L²-Betti numbers for rigid C*-tensor categories.*** (Co)homology for rigid C^*-tensor categories was introduced in [Popa et al. 2017]. One of the equivalent ways to describe this (co)homology theory is as Hochschild (co)homology for the tube algebra $\mathcal{A}$; see [Popa et al. 2017, Section 7.2]. Concretely, we equip $\mathcal{A}$ with the augmentation (or counit)

$$\varrho : \mathcal{A} \to \mathbb{C}, \qquad V \in (i\alpha, \alpha j) \mapsto \delta_{ij\varepsilon}\, \mathrm{Tr}_\alpha(V).$$

Since ϱ is a $*$-homomorphism, we can view $\mathbb{C}$ as an $\mathcal{A}$-module, which should be considered as the trivial representation of $\mathcal{C}$. Let $\mathcal{K}$ be a nondegenerate right Hilbert $\mathcal{A}$-module. We denote the (algebraic) linear span of $\mathcal{K} \cdot p_i$ for $i \in \mathrm{Irr}(\mathcal{C})$ by $\mathcal{K}^0$. Following [Popa et al. 2017], the homology of $\mathcal{C}$ with coefficients in $\mathcal{K}^0$ is then defined by

$$H_\bullet(\mathcal{C}, \mathcal{K}^0) = \mathrm{Tor}_\bullet^{\mathcal{A}}(\mathcal{K}^0, \mathbb{C}).$$

Similarly, the cohomology of $\mathcal{C}$ with coefficients in $\mathcal{K}^0$ is given by

$$H^\bullet(\mathcal{C}, \mathcal{K}^0) = \mathrm{Ext}^\bullet_{\mathcal{A}}(\mathbb{C}, \mathcal{K}^0).$$

Note that, in the special case where $\mathcal{K} = L^2(\mathcal{A})$, the left $\mathcal{A}''$-module structure on $L^2(\mathcal{A})$ induces a natural left $\mathcal{A}''$-module structure on the (co)homology spaces. As in [Popa et al. 2017], one then defines the n-th *L^2-Betti number* of $\mathcal{C}$ as

$$\beta_n^{(2)}(\mathcal{C}) = \dim_{\mathcal{A}''} H^n(\mathcal{C}, L^2(\mathcal{A})^0) = \dim_{\mathcal{A}''} H_n(\mathcal{C}, L^2(\mathcal{A})^0), \tag{2-5}$$

where $\dim_{\mathcal{A}''}$ is the Lück dimension with respect to the normal semifinite trace τ on $\mathcal{A}''$.

We refer to [Lück 2002, Section 6.1], [Kyed et al. 2015, Section A.4] and Remark 3.8 for the relevant definitions and properties of the dimension function $\dim_N$ on arbitrary N-modules, associated with a von Neumann algebra N equipped with a faithful normal semifinite trace Tr. Note that the second equality in (2-5) is nontrivial and was proved in [Popa et al. 2017, Proposition 6.4]. When $\mathcal{C}$ is a discrete group, all these notions reduce to the familiar ones for groups.

3. A scaling formula for L^2-Betti numbers

3A. *Index of a subcategory.*

Definition 3.1. Let $\mathcal{C}$ be a rigid C^*-tensor category and $\mathcal{C}_1 \subset \mathcal{C}$ a full C^*-tensor subcategory of $\mathcal{C}$. For an object $\alpha \in \mathcal{C}$, we define $[\alpha]_{\mathcal{C}_1}$ as the largest subobject of α that belongs to $\mathcal{C}_1$. We denote the orthogonal projection of α onto $[\alpha]_{\mathcal{C}_1}$ by $P^\alpha_{\mathcal{C}_1} \in (\alpha, \alpha)$. Fixing $\alpha \in \mathrm{Irr}(\mathcal{C})$, we define the $\mathcal{C}_1$-*orbit* of α as

$$\alpha \cdot \mathcal{C}_1 = \{\beta \in \mathrm{Irr}(\mathcal{C}) \mid \exists \gamma \in \mathcal{C}_1 \text{ such that } \beta \prec \alpha\gamma\}.$$

Note that in this definition, we can replace $\mathcal{C}_1$ by $\mathrm{Irr}(\mathcal{C}_1)$ without changing the orbit. By Frobenius reciprocity, the orbits form a partition of $\mathrm{Irr}(\mathcal{C})$. If $\alpha_1, \dots, \alpha_k$ are representatives of $\mathcal{C}_1$-orbits, the *index* of $\mathcal{C}_1 \subset \mathcal{C}$ is defined as

$$[\mathcal{C} : \mathcal{C}_1] = \sum_{i=1}^{k} \frac{d(\alpha_i)^2}{d([\bar{\alpha}_i \alpha_i]_{\mathcal{C}_1})}. \tag{3-1}$$

If the set of orbits is infinite, we put $[\mathcal{C} : \mathcal{C}_1] = \infty$.

In Lemma 3.2, we show that the index is well defined. In Proposition 3.12, we prove that $[\mathcal{C} : \mathcal{C}_1]$ equals the Jones index for an associated inclusion of von Neumann algebra completions of tube algebras. In Proposition 3.3, we prove the formula $[\mathcal{C} : \mathcal{C}_2] = [\mathcal{C} : \mathcal{C}_1]\,[\mathcal{C}_1 : \mathcal{C}_2]$ when $\mathcal{C}_2 \subset \mathcal{C}_1 \subset \mathcal{C}$. So, the above definition of $[\mathcal{C} : \mathcal{C}_1]$ is indeed natural.

When $\mathcal{C}_1 = \{\varepsilon\}$, the index defined above coincides with the global index $d(\mathcal{C})$ of $\mathcal{C}$. When $\mathcal{C}$ has only finitely many irreducible objects, we have $[\mathcal{C} : \mathcal{C}_1] = d(\mathcal{C})/d(\mathcal{C}_1)$; see Proposition 3.3.

Another extreme situation arises when

$$N(\mathcal{C}) = \{\gamma \in \mathrm{Irr}(\mathcal{C}) \mid \exists \alpha_1, \dots, \alpha_k \in \mathrm{Irr}(\mathcal{C}) \text{ such that } \gamma \prec \alpha_1 \cdots \alpha_k \bar{\alpha}_k \cdots \bar{\alpha}_1\} \tag{3-2}$$

is a subset of $\mathrm{Irr}(\mathcal{C}_1)$. In this case, the index simply counts the number of orbits. In particular, we recover the index for subgroups when $\mathcal{C}_1 \subset \mathcal{C}$ are both groups considered as C^*-tensor categories.

Lemma 3.2. *Let $\mathcal{C}$ be a rigid C^*-tensor category with full C^*-tensor subcategory $\mathcal{C}_1$. Then, for $\alpha, \beta \in \operatorname{Irr}(\mathcal{C})$ with $\beta \in \alpha \cdot \mathcal{C}_1$, we have*

$$\frac{d([\bar{\alpha}\alpha]_{\mathcal{C}_1})}{d(\alpha)^2} = \frac{d([\bar{\alpha}\beta]_{\mathcal{C}_1})}{d(\alpha)d(\beta)} = \frac{d([\bar{\beta}\beta]_{\mathcal{C}_1})}{d(\beta)^2}. \tag{3-3}$$

Proof. For arbitrary $\alpha, \beta \in \operatorname{Irr}(\mathcal{C})$, we have

$$(\operatorname{Tr}_{\bar{\alpha}} \otimes \operatorname{id})(P^{\bar{\alpha}\beta}_{\mathcal{C}_1}) = d(\beta)^{-1} \operatorname{Tr}_{\bar{\alpha}\beta}(P^{\bar{\alpha}\beta}_{\mathcal{C}_1})\, 1 = \frac{d([\bar{\alpha}\beta]_{\mathcal{C}_1})}{d(\beta)}\, 1,$$

by irreducibility of β. Now suppose that α, β satisfy the conditions of the lemma. Choose $\gamma \in \operatorname{Irr}(\mathcal{C}_1)$ such that $\beta \prec \alpha\gamma$. For any isometry $W : \beta \to \alpha\gamma$, we compute

$$\begin{aligned} W(\operatorname{Tr}_{\bar{\alpha}} \otimes \operatorname{id})(P^{\bar{\alpha}\beta}_{\mathcal{C}_1}) &= (\operatorname{Tr}_{\bar{\alpha}} \otimes \operatorname{id})((1 \otimes W) P^{\bar{\alpha}\beta}_{\mathcal{C}_1}) \\ &= (\operatorname{Tr}_{\bar{\alpha}} \otimes \operatorname{id})(P^{\bar{\alpha}\alpha\gamma}_{\mathcal{C}_1}(1 \otimes W)) = ((\operatorname{Tr}_{\bar{\alpha}} \otimes \operatorname{id})(P^{\bar{\alpha}\alpha}_{\mathcal{C}_1}) \otimes 1) W = \frac{d([\bar{\alpha}\alpha]_{\mathcal{C}_1})}{d(\alpha)} W, \end{aligned}$$

where we used that $P^{\bar{\alpha}\alpha\gamma}_{\mathcal{C}_1} = P^{\bar{\alpha}\alpha}_{\mathcal{C}_1} \otimes 1$, as is easy to see by splitting $\bar{\alpha}\alpha$ into irreducible components. Multiplying by W^* on the left, we find that

$$(\operatorname{Tr}_{\bar{\alpha}} \otimes \operatorname{id})(P^{\bar{\alpha}\beta}_{\mathcal{C}_1}) = \frac{d([\bar{\alpha}\alpha]_{\mathcal{C}_1})}{d(\alpha)} 1.$$

We already proved that the left-hand side equals $d([\bar{\alpha}\beta]_{\mathcal{C}_1})/d(\beta)\, 1$. So, the first equality in (3-3) follows. The second one is proven analogously. □

Proposition 3.3. *Let $\mathcal{C}$ be a rigid C^*-tensor category with full C^*-tensor subcategories $\mathcal{C}_2 \subset \mathcal{C}_1 \subset \mathcal{C}$. Then,*

$$[\mathcal{C} : \mathcal{C}_2] = [\mathcal{C} : \mathcal{C}_1]\, [\mathcal{C}_1 : \mathcal{C}_2].$$

In particular, if $\mathcal{C}$ is a rigid C^*-tensor category with finitely many irreducible objects and if $\mathcal{C}_1 \subset \mathcal{C}$ is a full C^*-tensor subcategory, then $[\mathcal{C} : \mathcal{C}_1] = d(\mathcal{C})/d(\mathcal{C}_1)$, where $d(\mathcal{C})$ and $d(\mathcal{C}_1)$ denote the global index of $\mathcal{C}$ and $\mathcal{C}_1$.

Since a short proof for Proposition 3.3 can be given using the language of Markov inclusions, we postpone the proof until the end of Section 3C.

In the concrete computations of L^2-Betti numbers in this paper, we only need the particularly easy tensor subcategories $\mathcal{C}_1 \subset \mathcal{C}$ that arise from a homomorphism to a finite group. More precisely, assume that we are given a group Λ and a map $\Xi : \operatorname{Irr}(\mathcal{C}) \to \Lambda$ satisfying the following two properties:

(i) For all $\alpha, \beta, \gamma \in \operatorname{Irr}(\mathcal{C})$ with $\gamma \prec \alpha\beta$, we have $\Xi(\gamma) = \Xi(\alpha)\Xi(\beta)$.

(ii) For all $\alpha \in \operatorname{Irr}(\mathcal{C})$, we have $\Xi(\bar{\alpha}) = \Xi(\alpha)^{-1}$.

Defining $\operatorname{Ker}(\Xi) \subset \mathcal{C}$ as those objects in $\mathcal{C}$ that can be written as a direct sum of irreducible objects $\gamma \in \operatorname{Irr}(\mathcal{C})$ with $\Xi(\gamma) = e$, we obtain a full C^*-tensor subcategory $\operatorname{Ker}(\Xi) \subset \mathcal{C}$ of index $|\Lambda|$.

Note that $N(\mathcal{C})$, as defined in (3-2), always is a subset of $\operatorname{Ker}(\Xi)$. Actually, denoting by Γ the set of orbits for the left (or right) action of $N(\mathcal{C})$ on $\operatorname{Irr}(\mathcal{C})$, we get that Γ has a natural group structure and we can view Γ as the largest group quotient of $\mathcal{C}$.

3B. ***Markov inclusions of tracial von Neumann algebras.*** In [Popa 1994, Section 1.1.4], the concept of a λ-Markov inclusion $N \subset (M, \tau)$ of tracial von Neumann algebras was introduced. More generally, Popa [1995, Section 1.2] defined the λ-Markov property for arbitrary inclusions of von Neumann algebras $N \subset M$ together with a faithful normal conditional expectation $E : M \to N$. Taking in the tracial setting the unique trace-preserving conditional expectation, both notions coincide.

In this paper, we need a slight variant of this concept for inclusions $N \subset M$ where both N and M are equipped with fixed faithful normal semifinite traces, denoted by Tr_N and Tr_M, but the inclusion need not be trace-preserving. In particular, there is no canonical conditional expectation of M onto N.

Recall that an element $v \in M$ is called right N-bounded if there exists a $\kappa > 0$ such that $\mathrm{Tr}_M(a^* v^* v a) \leq \kappa \,\mathrm{Tr}_N(a^* a)$ for all $a \in N$. We denote by $L_v : L^2(N, \mathrm{Tr}_N) \to L^2(M, \mathrm{Tr}_M)$ the associated bounded operator, which is right N-linear and given by $L_v(a) = va$ for all $a \in N \cap L^2(N, \mathrm{Tr}_N)$. A family $(v_i)_{i \in I}$ of right N-bounded vectors in M is called a Pimsner–Popa basis for $N \subset M$ if

$$\sum_{i \in I} L_{v_i} L_{v_i}^* = 1.$$

Definition 3.4. Let (N, Tr_N) and (M, Tr_M) be von Neumann algebras equipped with faithful normal semifinite traces. Assume that $N \subset M$, but without assuming that this inclusion is trace-preserving. We say that the inclusion is λ-Markov for a given number $\lambda > 0$ if a Pimsner–Popa basis $(v_i)_{i \in I}$ satisfies

$$\sum_{i \in I} v_i v_i^* = \lambda^{-1}\, 1.$$

One checks that this definition does not depend on the choice of the Pimsner–Popa basis.

Definition 3.5. Given a von Neumann algebra M equipped with a faithful normal semifinite trace Tr, we call an (algebraic) right M-module $\mathcal{E}$ *locally finite* if for every $\xi \in \mathcal{E}$, there exists a projection $p \in M$ with $\mathrm{Tr}(p) < \infty$ and $\xi = \xi p$.

Note that for every projection $p \in M$ with $\mathrm{Tr}(p) < \infty$, the right M-module $pL^2(M)$ is locally finite, because for every $\xi \in L^2(M)$, the right support projection of $p\xi$ has finite trace.

For our computations, the following scaling formula is essential.

Proposition 3.6. *Let (N, Tr_N) and (M, Tr_M) be von Neumann algebras equipped with faithful normal semifinite traces. Assume that $N \subset M$ and that $\lambda > 0$. The inclusion is λ-Markov if and only if $\dim_M(\mathcal{E}) = \lambda \dim_N(\mathcal{E})$ for every locally finite M-module $\mathcal{E}$.*

We have $\dim_M(\mathcal{E}) = \lambda \dim_N(\mathcal{E})$ for arbitrary M-modules $\mathcal{E}$ if and only if the inclusion is λ-Markov and the restriction of Tr_M to N is semifinite.

Proof. Fix a Pimsner–Popa basis $(v_i)_{i \in I}$ for $N \subset M$, with respect to the traces Tr_N, Tr_M. Define the projection $q \in B(\ell^2(I)) \mathbin{\overline{\otimes}} N$ given by $q_{ij} = L_{v_i}^* L_{v_j}$. Then,

$$U : L^2(M, \mathrm{Tr}_M) \to q(\ell^2(I) \otimes L^2(N, \mathrm{Tr}_N)), \quad U(x) = \sum_{i \in I} e_i \otimes L_{v_i}^*(x),$$

is a well-defined right N-linear unitary operator. Whenever $a \in M$, the operator UaU^* commutes with the right N-action and so, we get a well-defined unital $*$-homomorphism

$$\alpha : M \to q(B(\ell^2(I)) \mathbin{\overline{\otimes}} N)q : \alpha(a) = UaU^*.$$

A direct computation gives that

$$(\mathrm{Tr} \otimes \mathrm{Tr}_N)(\alpha(a)) = \sum_{i \in I} \mathrm{Tr}_M(v_i^* a v_i)$$

for all $a \in M^+$. So the inclusion $N \subset M$ is λ-Markov if and only if $\mathrm{Tr}_M(p) = \lambda\,(\mathrm{Tr} \otimes \mathrm{Tr}_N)(\alpha(p))$ for every projection $p \in M$. Note that the left-hand side equals $\dim_{-M}(pL^2(M))$, while the right-hand side equals $\lambda \dim_{-N}(pL^2(M))$. So if the formula $\dim_M(\mathcal{E}) = \lambda\, \dim_N(\mathcal{E})$ holds for all locally finite M-modules, it holds in particular for $\mathcal{E} = pL^2(M)$ for every projection $p \in M$ with $\mathrm{Tr}_M(p) < \infty$ and we conclude that $\mathrm{Tr}_M(p) = \lambda\,(\mathrm{Tr} \otimes \mathrm{Tr}_N)(\alpha(p))$ for every projection $p \in M$ with $\mathrm{Tr}_M(p) < \infty$. An arbitrary projection $p \in M$ can be written as the limit of an increasing net of finite trace projections, so that the same formula holds for all projections $p \in M$ and thus, $N \subset M$ is λ-Markov.

Conversely, assume that $N \subset M$ is λ-Markov. We prove that $\dim_M(\mathcal{E}) = \lambda\, \dim_N(\mathcal{E})$ for every locally finite M-module $\mathcal{E}$. Denote by $\mathcal{L}_1$ the class of M-modules that are isomorphic with $p(\mathbb{C}^n \otimes M)$ for some $n \in \mathbb{N}$ and some projection $p \in M_n(\mathbb{C}) \otimes M$ having finite trace. We start by proving that $\dim_N(\mathcal{E}) = \lambda^{-1} \dim_M(\mathcal{E})$ for all $\mathcal{E} \in \mathcal{L}_1$.

Take a finite trace projection $p \in M_n(\mathbb{C}) \otimes M$ such that $\mathcal{E} \cong p(\mathbb{C}^n \otimes M)$. We have

$$p(\mathbb{C}^n \otimes M) \subset p(\mathbb{C}^n \otimes L^2(M, \mathrm{Tr}_M)) \cong (\mathrm{id} \otimes \alpha)(p)\big(\mathbb{C}^n \otimes \ell^2(I) \otimes L^2(N, \mathrm{Tr}_N)\big).$$

Therefore,

$$\begin{aligned}\dim_{-N}(p(\mathbb{C}^n \otimes M)) &\leq \dim_{-N}\big((\mathrm{id} \otimes \alpha)(p)(\mathbb{C}^n \otimes \ell^2(I) \otimes L^2(N, \mathrm{Tr}_N))\big) \\ &= (\mathrm{Tr} \otimes \mathrm{Tr} \otimes \mathrm{Tr}_N)(\mathrm{id} \otimes \alpha)(p) \\ &= \lambda^{-1}\,(\mathrm{Tr} \otimes \mathrm{Tr}_M)(p) = \lambda^{-1}\, \dim_{-M}(p(\mathbb{C}^n \otimes M)).\end{aligned}$$

Conversely, since

$$(1 \otimes U^*)\big((\mathrm{id} \otimes \alpha)(p)(e_i \otimes e_j \otimes a)\big) = p(e_i \otimes v_j a)$$

for all $i \in \{1, \dots, n\}$, $j \in I$ and $a \in N$, we get for every finite subset $I_0 \subset I$ the injective N-module map

$$(\mathrm{id} \otimes \alpha)(p)(\mathbb{C}^n \otimes \ell^2(I_0) \otimes N) \hookrightarrow p(\mathbb{C}^n \otimes M).$$

Letting I_0 increase and taking $\dim_{-N}$, it follows that

$$(\mathrm{Tr} \otimes \mathrm{Tr} \otimes \mathrm{Tr}_N)(\mathrm{id} \otimes \alpha)(p) \leq \dim_{-N}(p(\mathbb{C}^n \otimes M)).$$

The left-hand side equals $\lambda^{-1}\,(\mathrm{Tr} \otimes \mathrm{Tr}_M)(p) = \lambda^{-1}\, \dim_{-M}(p(\mathbb{C}^n \otimes M))$. In combination with the converse inequality above, we have proved that $\dim_M(\mathcal{E}) = \lambda\, \dim_N(\mathcal{E})$ for every $\mathcal{E} \in \mathcal{L}_1$.

Next denote by $\mathcal{L}_2$ the class of all M-modules that arise as the quotient of an M-module in $\mathcal{L}_1$. Let $\mathcal{E} \in \mathcal{L}_2$ and let $0 \to \mathcal{E}_0 \to \mathcal{E}_1 \to \mathcal{E} \to 0$ be an exact sequence of M-modules, with $\mathcal{E}_1 \in \mathcal{L}_1$. Since every finitely generated M-submodule of an M-module in $\mathcal{L}_1$ again belongs to $\mathcal{L}_1$, we can write $\mathcal{E}_0$ as the union of an increasing family of M-submodules $\mathcal{E}_j \subset \mathcal{E}_0$ with $\mathcal{E}_j \in \mathcal{L}_1$ for all j. Since both $\dim_M$ and $\dim_N$ are continuous when taking increasing unions (see Remark 3.7), we get $\dim_M(\mathcal{E}_0) = \lambda \dim_N(\mathcal{E}_0)$. Since both $\dim_M$ and $\dim_N$ are additive with respect to short exact sequences (see Remark 3.7 as well), we conclude that

$$\dim_M(\mathcal{E}) = \dim_M(\mathcal{E}_1) - \dim_M(\mathcal{E}_0) = \lambda \dim_N(\mathcal{E}_1) - \lambda \dim_N(\mathcal{E}_0) = \lambda \dim_N(\mathcal{E}).$$

Finally, every locally finite M-module can be written as the union of an increasing family of M-submodules in $\mathcal{L}_2$. So again using the continuity of the dimension function, we find that $\dim_M(\mathcal{E}) = \lambda \dim_N(\mathcal{E})$ for all locally finite M-modules $\mathcal{E}$.

Next assume that $N \subset M$ is λ-Markov and that the restriction of Tr_M to N is semifinite. We can then choose an increasing net of projections $p_n \in N$, converging to 1 strongly, with $\mathrm{Tr}_N(p_n) < \infty$ and $\mathrm{Tr}_M(p_n) < \infty$ for all n. Let $\mathcal{E}$ be an arbitrary M-module. By [Kyed et al. 2015, Lemmas A.15 and A.16], we have $\dim_M(\mathcal{E}) = \lim_n \dim_M(\mathcal{E} p_n M)$. For each n, the M-module $\mathcal{E} p_n M$ is locally finite. Therefore, $\dim_M(\mathcal{E} p_n M) = \lambda \dim_N(\mathcal{E} p_n M)$. Since $\mathcal{E} p_n M \subset \mathcal{E}$, it follows that $\dim_M(\mathcal{E}) \leq \lambda \dim_N(\mathcal{E})$. Conversely, $\mathcal{E} p_n N \subset \mathcal{E} p_n M$, so that

$$\dim_N(\mathcal{E} p_n M) \geq \dim_N(\mathcal{E} p_n N).$$

Again using [Kyed et al. 2015, Lemmas A.15 and A.16], we have $\lim_n \dim_N(\mathcal{E} p_n N) = \dim_N(\mathcal{E})$, so that the inequality $\dim_M(\mathcal{E}) \geq \lambda \dim_N(\mathcal{E})$ follows.

Finally, assume that the restriction of Tr_M to N is not semifinite. We then find a nonzero projection $p \in N$ such that $\mathrm{Tr}_M(x) = +\infty$ for every nonzero element $x \in pN^+p$. Define the two-sided ideal $M_0 \subset M$ consisting of all elements $x \in M$ whose left (equivalently right) support projection has finite Tr_M. Define $\mathcal{E} = M/M_0$ and view $\mathcal{E}$ as a right M-module. Whenever $p \in M$ is a projection with $\mathrm{Tr}_M(p) < \infty$, we have $\mathcal{E} p = \{0\}$. By [Kyed et al. 2015, Definition A.14], we have $\dim_{-M}(\mathcal{E}) = 0$. On the other hand, the map $pN \to \mathcal{E}$, $x \mapsto x + M_0$, is N-linear and injective because $pN \cap M_0 = \{0\}$. Therefore, $\dim_N(\mathcal{E}) \geq \mathrm{Tr}_N(p) > 0$. So, the dimension scaling formula fails in general when the restriction of Tr_M to N is no longer semifinite. □

Remark 3.7. In the proof of Proposition 3.6, we made use of the following continuity and additivity properties of the dimension function $\dim_M$ associated with a von Neumann algebra M equipped with a faithful normal semifinite trace Tr:

(i) Assume that $\mathcal{E}$ is an M-module and $\mathcal{E}_j \subset \mathcal{E}$ is an increasing net of M-submodules with $\bigcup_j \mathcal{E}_j = \mathcal{E}$. Then, $\dim_M(\mathcal{E}) = \lim_j \dim_M(\mathcal{E}_j)$.

(ii) Assume that $0 \to \mathcal{E}_1 \to \mathcal{E} \to \mathcal{E}_2 \to 0$ is an exact sequence of M-modules. Then, $\dim_M(\mathcal{E}) = \dim_M(\mathcal{E}_1) + \dim_M(\mathcal{E}_2)$.

When Tr is a tracial state, meaning that $\mathrm{Tr}(1) = 1$, these properties are proved in [Lück 2002, Theorem 6.7(4)]. When Tr is semifinite, we can take an increasing net of projections $p_i \in M$ with $\mathrm{Tr}(p_i) < \infty$ for all i and $p_i \to 1$ strongly. Define the tracial state τ_i on $p_i M p_i$ given by $\tau_i(x) = \mathrm{Tr}(p_i)^{-1} \mathrm{Tr}(x)$. Then [Kyed et al. 2015, Lemma A.16] says that for every M-module $\mathcal{E}$, the net $\mathrm{Tr}(p_i) \dim_{(p_i M p_i, \tau_i)}(\mathcal{E} p_i)$ is increasing and converges to $\dim_M(\mathcal{E})$. Therefore, the continuity and additivity properties (i) and (ii) above are also valid for $\dim_M$.

Remark 3.8. Let (M, Tr) be a von Neumann algebra equipped with a faithful normal semifinite trace. Proposition 3.6 shows that the dimension function $\dim_M$ has a subtle behavior. We therefore also want to clarify why [Kyed et al. 2015, Definition A.14], given by

$$\dim_M(\mathcal{E}) = \sup\{\mathrm{Tr}(q)\dim_{qMq}(\mathcal{E}q) \mid q \in M \text{ is a projection with } \mathrm{Tr}(q) < \infty\} \tag{3-4}$$

and making use of the dimension function for $(qMq, \mathrm{Tr}(q)^{-1}\mathrm{Tr}(\,\cdot\,))$, coincides with [Petersen 2012, Definition B.17], given by

$$\dim_M(\mathcal{E}) = \sup\{(\mathrm{Tr}\otimes\mathrm{Tr})(p) \mid p \in M_n(\mathbb{C})\otimes M \text{ is a projection with finite trace and } p(\mathbb{C}^n\otimes M) \hookrightarrow \mathcal{E} \text{ as } M\text{-modules}\}. \tag{3-5}$$

Whenever $p(\mathbb{C}^n\otimes M) \hookrightarrow \mathcal{E}$, we have $p(\mathbb{C}^n\otimes Mq) \hookrightarrow \mathcal{E}q$. Denoting by $z_q \in \mathcal{Z}(M)$ the central support of q, it follows from [Kyed et al. 2015, Lemma A.15] that

$$\mathrm{Tr}(q)\dim_{qMq}(\mathcal{E}q) \geq \mathrm{Tr}(q)\dim_{qMq}(p(\mathbb{C}^n\otimes Mq)) = (\mathrm{Tr}\otimes\mathrm{Tr})(p(1\otimes z_q)).$$

Taking the supremum over all finite trace projections $q \in M$ and all embeddings $p(\mathbb{C}^n\otimes M) \hookrightarrow \mathcal{E}$, it follows that the dimension in (3-5) is bounded above by the dimension in (3-4).

Conversely, $\mathrm{Tr}(q)\dim_{qMq}(\mathcal{E}q)$ can be computed as the supremum of $(\mathrm{Tr}\otimes\mathrm{Tr})(p)$, where $p \in M_n(\mathbb{C})\otimes qMq$ is a projection and $\theta : p(\mathbb{C}^n\otimes Mq) \hookrightarrow \mathcal{E}q$. Defining $\xi \in \overline{\mathbb{C}}^n\otimes\mathcal{E}q$ by

$$\xi = \sum_{i=1}^{n} e_i^*\otimes\theta(p(e_i\otimes q)),$$

it follows that $\xi = \xi p$ and $\theta(x) = \xi x$ for all $x \in p(\mathbb{C}^n\otimes Mq)$. Then $\psi : p(\mathbb{C}^n\otimes M) \to \mathcal{E}$, $\psi(x) = \xi x$, is M-linear. We claim that ψ remains injective. Indeed, if $\psi(x) = 0$, then for all $i \in \{1,\dots,n\}$, also $\theta(xx^*(e_i\otimes 1)) = \psi(x)x^*(e_i\otimes 1) = 0$. So, $xx^*(e_i\otimes 1) = 0$ for all i and thus, $x = 0$. It follows that the dimension in (3-4) is bounded above by the dimension in (3-5).

3C. *The scaling formula.* The goal of this section is to prove the following scaling formula for L^2-Betti numbers under finite-index inclusions.

Theorem 3.9. *Let $\mathcal{C}_1 \subset \mathcal{C}$ be a finite-index inclusion of rigid C^*-tensor categories. Then*

$$\beta_n^{(2)}(\mathcal{C}_1) = [\mathcal{C} : \mathcal{C}_1]\,\beta_n^{(2)}(\mathcal{C})$$

for all $n \geq 0$.

For the rest of this section, fix a rigid C^*-tensor category $\mathcal{C}$ and a full C^*-tensor subcategory $\mathcal{C}_1 \subset \mathcal{C}$. The tube algebra $\mathcal{A}_1$ of $\mathcal{C}_1$ naturally is a unital $*$-subalgebra of a *corner* of the tube algebra $\mathcal{A}$ of $\mathcal{C}$. In dimension computations, this causes a number of issues that can be avoided by considering the $*$-subalgebra $\widetilde{\mathcal{A}}_1 \subset \mathcal{A}$ given by

$$\widetilde{\mathcal{A}}_1 = \bigoplus_{i,j\in\mathrm{Irr}(\mathcal{C})}\ \bigoplus_{\alpha\in\mathrm{Irr}(\mathcal{C}_1)} (i\alpha, \alpha j). \tag{3-6}$$

We still have a natural trace τ on $\widetilde{\mathcal{A}}_1$ and the inclusion $\widetilde{\mathcal{A}}_1 \subset \mathcal{A}$ is trace-preserving.

As a first lemma, we prove that the homology of $\mathcal{C}_1$ can be computed as the Hochschild homology of $\widetilde{\mathcal{A}}_1$ with the counit augmentation $\varrho : \widetilde{\mathcal{A}}_1 \to \mathbb{C}$.

Lemma 3.10. *Define the central projection p_1 in the multiplier algebra of $\widetilde{\mathcal{A}}_1$ given by $p_1 = \sum_{i \in \mathrm{Irr}(\mathcal{C}_1)} p_i$. Note that $p_1 \cdot \widetilde{\mathcal{A}}_1 \cdot p_1 \cong \mathcal{A}_1$ naturally.*

For every nondegenerate right Hilbert $\widetilde{\mathcal{A}}_1$-module $\mathcal{K}$, there are natural isomorphisms

$$H_\bullet(\mathcal{C}_1, \mathcal{K} \cdot p_1) \cong \mathrm{Tor}_\bullet^{\widetilde{\mathcal{A}}_1}(\mathcal{K}^0, \mathbb{C}) \quad \text{and} \quad H^\bullet(\mathcal{C}_1, \mathcal{K} \cdot p_1) \cong \mathrm{Ext}^\bullet_{\widetilde{\mathcal{A}}_1}(\mathbb{C}, \mathcal{K}^0).$$

We also have

$$\beta_n^{(2)}(\mathcal{C}_1) = \dim_{\widetilde{\mathcal{A}}_1''} \mathrm{Tor}_n^{\widetilde{\mathcal{A}}_1}(L^2(\widetilde{\mathcal{A}}_1)^0, \mathbb{C}).$$

Proof. If $i \in \mathrm{Irr}(\mathcal{C})$ and $\alpha, j \in \mathrm{Irr}(\mathcal{C}_1)$, then $(i\alpha, \alpha j)$ can only be nonzero if $i \in \mathrm{Irr}(\mathcal{C}_1)$, by Frobenius reciprocity. Interchanging the roles of i and j, we conclude that p_1 is central in the multiplier algebra $M(\widetilde{\mathcal{A}}_1)$. Because $p_\varepsilon \le p_1$, it follows that

$$\begin{aligned}\widetilde{\mathcal{A}}_1 \otimes_{\mathcal{B}} \cdots \otimes_{\mathcal{B}} \widetilde{\mathcal{A}}_1 \cdot p_\varepsilon &= p_1 \cdot \widetilde{\mathcal{A}}_1 \cdot p_1 \otimes_{\mathcal{B}} \cdots \otimes_{\mathcal{B}} p_1 \cdot \widetilde{\mathcal{A}}_1 \cdot p_\varepsilon \\ &\cong \mathcal{A}_1 \otimes_{\mathcal{B}_1} \cdots \otimes_{\mathcal{B}_1} \mathcal{A}_1 \cdot p_\varepsilon,\end{aligned}$$

and the right bar resolution is similar. Since the bar resolutions associated to $\mathcal{A}_1$ and $\widetilde{\mathcal{A}}_1$ are equal, the respective Tor and Ext functors must also be the same. □

The following formula, generalizing [Popa et al. 2017, Lemma 3.9], is crucial for us since we deduce from it that $\mathcal{A}$ is a projective $\widetilde{\mathcal{A}}_1$-module and also that in the finite-index case, the inclusion $\widetilde{\mathcal{A}}_1'' \subset \mathcal{A}''$ is λ-Markov in the sense of Definition 3.4.

Lemma 3.11. *For $\alpha \in \mathrm{Irr}(\mathcal{C})$, we denote by $e_{\alpha\cdot\mathcal{C}_1}$ the orthogonal projection of $L^2(\mathcal{A})$ onto the closed linear span of all $(i\beta, \beta j)$ with $i, j \in \mathrm{Irr}(\mathcal{C})$ and $\beta \in \alpha \cdot \mathcal{C}_1$.*

Then, for all $i \in \mathrm{Irr}(\mathcal{C})$ and $\alpha \in \mathrm{Irr}(\mathcal{C})$, we have

$$\sum_{j \in \mathrm{Irr}(\mathcal{C})} \; \sum_{W \in \mathrm{onb}(i\alpha, \alpha j)} d(j)\, W \cdot e_{\mathcal{C}_1} \cdot W^\# = \frac{d([\bar\alpha\alpha]_{\mathcal{C}_1})}{d(\alpha)}\, p_i \cdot e_{\alpha\cdot\mathcal{C}_1} \tag{3-7}$$

as operators on $L^2(\mathcal{A})$.

Proof. Both the left- and the right-hand sides of (3-7) vanish on $(i_1\beta, \beta k) \subset L^2(\mathcal{A})$ if $i_1 \ne i$. So we fix $k, \beta \in \mathrm{Irr}(\mathcal{C})$ and $V \in (i\beta, \beta k)$ and prove that both sides of (3-7) agree on V.

For every $j \in \mathrm{Irr}(\mathcal{C})$ and $W \in (i\alpha, \alpha j)$, we have

$$(e_{\mathcal{C}_1} \cdot W^\#)(V) = \sum_{\gamma \in \mathrm{Irr}(\mathcal{C}_1)} \; \sum_{U \in \mathrm{onb}(\bar\alpha\beta, \gamma)} d(\gamma)(1 \otimes U^*)(W^\# \otimes 1)(1 \otimes V)(U \otimes 1).$$

We claim that $(e_{\mathcal{C}_1} \cdot W^\#)(V)$ is the image in $\mathcal{A}$ under the map in (2-3) of the element

$$(W^\# \otimes 1)(1 \otimes V)(P_{\mathcal{C}_1}^{\bar\alpha\beta} \otimes 1) \in (j(\bar\alpha\beta), (\bar\alpha\beta)k). \tag{3-8}$$

The claim follows because that image is given by

$$\sum_{\gamma\in\mathrm{Irr}(\mathcal{C})}\ \sum_{U\in\mathrm{onb}(\bar\alpha\beta,\gamma)} d(\gamma)(1\otimes U^*)(W^\#\otimes 1)(1\otimes V)(P_{\mathcal{C}_1}^{\bar\alpha\beta}\otimes 1)(U\otimes 1)$$
$$= \sum_{\gamma\in\mathrm{Irr}(\mathcal{C}_1)}\ \sum_{U\in\mathrm{onb}(\bar\alpha\beta,\gamma)} d(\gamma)(1\otimes U^*)(W^\#\otimes 1)(1\otimes V)(U\otimes 1),$$

because $P_{\mathcal{C}_1}^{\bar\alpha\beta}U$ equals 0 when $\gamma\notin\mathrm{Irr}(\mathcal{C}_1)$ and equals U when $\gamma\in\mathrm{Irr}(\mathcal{C}_1)$.

It then follows that $(W\cdot e_{\mathcal{C}_1}\cdot W^\#)(V)$ is the image in $\mathcal{A}$ of the element

$$(W\otimes 1\otimes 1)(1\otimes W^\#\otimes 1)(1\otimes 1\otimes V)(1\otimes P_{\mathcal{C}_1}^{\bar\alpha\beta}\otimes 1)\in(i(\alpha\bar\alpha\beta),(\alpha\bar\alpha\beta)k). \tag{3-9}$$

By Frobenius reciprocity,

$$\{W_Z=(s_\alpha^*\otimes 1\otimes 1)(1\otimes Z)\mid Z\in\mathrm{onb}(\bar\alpha i\alpha, j)\}$$

is an orthonormal basis of $(i\alpha,\alpha j)$, and any orthonormal basis can be written in this form.

With this notation, we find that

$$\sum_{j\in\mathrm{Irr}(\mathcal{C})}\ \sum_{Z\in\mathrm{onb}(\bar\alpha i\alpha,j)} d(j)(W_Z\otimes 1)(1\otimes W_Z^\#)=\sum_{j\in\mathrm{Irr}(\mathcal{C})}\ \sum_{Z\in\mathrm{onb}(\bar\alpha i\alpha,j)} d(j)(s_\alpha^*\otimes 1^{\otimes 3})(1\otimes ZZ^*\otimes 1)(1^{\otimes 3}\otimes s_\alpha)$$
$$=s_\alpha^*\otimes 1\otimes s_\alpha=(1\otimes s_\alpha)(s_\alpha^*\otimes 1). \tag{3-10}$$

Combining (3-9) and (3-10), we thus obtain

$$\sum_{j\in\mathrm{Irr}(\mathcal{C})}\ \sum_{W\in\mathrm{onb}(i\alpha,\alpha j)} d(j)(W\cdot e_{\mathcal{C}_1}\cdot W^\#)(V)$$
$$=\sum_{\gamma\in\mathrm{Irr}(\mathcal{C})}\ \sum_{U\in\mathrm{onb}(\alpha\bar\alpha\beta,\gamma)} d(\gamma)(1\otimes U^*)(1\otimes s_\alpha\otimes 1)V(s_\alpha^*\otimes 1\otimes 1)(1\otimes P_{\mathcal{C}_1}^{\bar\alpha\beta}\otimes 1)(U\otimes 1).$$

Choosing the orthonormal basis of $(\alpha\bar\alpha\beta,\gamma)$ by first decomposing $\alpha\bar\alpha$, we see that only one of the $U^*(s_\alpha\otimes 1)$ is nonzero and conclude that

$$\sum_{j\in\mathrm{Irr}(\mathcal{C})}\ \sum_{W\in\mathrm{onb}(i\alpha,\alpha j)} d(j)(W\cdot e_{\mathcal{C}_1}\cdot W^\#)(V)=V(s_\alpha^*\otimes 1\otimes 1)(1\otimes P_{\mathcal{C}_1}^{\bar\alpha\beta}\otimes 1)(s_\alpha\otimes 1\otimes 1)$$
$$=V\big((\mathrm{Tr}_{\bar\alpha}\otimes\mathrm{id})(P_{\mathcal{C}_1}^{\bar\alpha\beta})\otimes 1\big).$$

Using Lemma 3.2, we get

$$V((\mathrm{Tr}_{\bar\alpha}\otimes\mathrm{id})(P_{\mathcal{C}_1}^{\bar\alpha\beta})\otimes 1)=\frac{d([\bar\alpha\beta]_{\mathcal{C}_1})}{d(\beta)}V=\begin{cases}\dfrac{d([\bar\alpha\alpha]_{\mathcal{C}_1})}{d(\alpha)}V & \text{if } \beta\in\alpha\cdot\mathcal{C}_1,\\ 0 & \text{otherwise.}\end{cases}$$ □

Proposition 3.12. *Let $\mathcal{C}_1\subset\mathcal{C}$ be a finite-index inclusion of rigid C^*-tensor categories. Denote by $\mathcal{A}$ the tube algebra of $\mathcal{C}$ and define its subalgebra $\widetilde{\mathcal{A}}_1$ as in (3-6). Then $\mathcal{A}$ is projective as a left $\widetilde{\mathcal{A}}_1$-module and as a right $\widetilde{\mathcal{A}}_1$-module. Moreover, the associated inclusion of von Neumann algebras $\widetilde{\mathcal{A}}_1''\subset\mathcal{A}''$ is λ-Markov with $\lambda=[\mathcal{C}:\mathcal{C}_1]^{-1}$ in the sense of Definition 3.4.*

Proof. By symmetry, it suffices to prove that $\mathcal{A}$ is a projective right $\widetilde{\mathcal{A}}_1$-module.

For each $\alpha \in \mathrm{Irr}(\mathcal{C})$, define the subspace $\mathcal{A}_{\alpha\cdot\mathcal{C}_1} \subset \mathcal{A}$ spanned by all $(i\beta, \beta j)$ with $i, j \in \mathrm{Irr}(\mathcal{C})$ and $\beta \in \alpha \cdot \mathcal{C}_1$. Note that $\mathcal{A}_{\alpha\cdot\mathcal{C}_1} \subset \mathcal{A}$ is a right $\widetilde{\mathcal{A}}_1$-submodule. As in Lemma 3.11, denote by $e_{\mathcal{C}_1}$ the orthogonal projection of $L^2(\mathcal{A})$ onto $L^2(\widetilde{\mathcal{A}}_1)$. Note that $e_{\mathcal{C}_1}(\mathcal{A}) = \widetilde{\mathcal{A}}_1$.

Fix $i, \alpha \in \mathrm{Irr}(\mathcal{C})$ and define the projective right $\widetilde{\mathcal{A}}_1$-module

$$\mathcal{V} := \bigoplus_{j \in \mathrm{Irr}(\mathcal{C})} \big((i\alpha, \alpha j) \otimes p_j \cdot \widetilde{\mathcal{A}}_1\big).$$

The maps

$$\theta_1 : p_i \cdot \mathcal{A}_{\alpha\cdot\mathcal{C}_1} \to \mathcal{V}, \quad \theta_1(V) = \bigoplus_{j \in \mathrm{Irr}(\mathcal{C})} \Bigg(\sum_{W \in \mathrm{onb}(i\alpha, \alpha j)} d(j) W \otimes e_{\mathcal{C}_1}(W^\# \cdot V) \Bigg),$$

$$\theta_2 : \mathcal{V} \to p_i \cdot \mathcal{A}_{\alpha\cdot\mathcal{C}_1}, \quad \theta_2(W \otimes V) = W \cdot V$$

are right $\widetilde{\mathcal{A}}_1$-linear. By Lemma 3.11, we have $\theta_2 \circ \theta_1$ equals a multiple of the identity map on $p_i \cdot \mathcal{A}_{\alpha\cdot\mathcal{C}_1}$. It follows that $p_i \cdot \mathcal{A}_{\alpha\cdot\mathcal{C}_1}$ is a projective right $\widetilde{\mathcal{A}}_1$-module.

Taking the (direct) sum over all $i \in \mathrm{Irr}(\mathcal{C})$ and over a set of representatives $\alpha_1, \ldots, \alpha_\kappa$ for the $\mathcal{C}_1$-orbits in $\mathrm{Irr}(\mathcal{C})$, we conclude that also $\mathcal{A}$ is projective as a right $\widetilde{\mathcal{A}}_1$-module.

By Lemma 3.11, we have

$$\left\{ \sqrt{\frac{d(j)\, d(\alpha_s)}{d([\bar{\alpha}_s \alpha_s]_{\mathcal{C}_1})}}\, W \;\middle|\; i, j \in \mathrm{Irr}(\mathcal{C}),\ s = 1, \ldots, \kappa,\ W \in \mathrm{onb}(i\alpha_s, \alpha_s j) \right\}$$

is a Pimsner–Popa basis for the inclusion $\widetilde{\mathcal{A}}_1'' \subset \mathcal{A}''$. Applying Lemma 3.11 in the case $\mathcal{C}_1 = \mathcal{C}$ (and this literally is [Popa et al. 2017, Lemma 3.9]), we get

$$\sum_{s=1}^{\kappa} \sum_{i,j \in \mathrm{Irr}(\mathcal{C})} \sum_{W \in \mathrm{onb}(i\alpha_s, \alpha_s j)} \frac{d(j)\, d(\alpha_s)}{d([\bar{\alpha}_s \alpha_s]_{\mathcal{C}_1})} W \cdot W^\# = \sum_{s=1}^{\kappa} \sum_{i \in \mathrm{Irr}(\mathcal{C})} \frac{d(\alpha_s)^2}{d([\bar{\alpha}_s \alpha_s]_{\mathcal{C}_1})} p_i = [\mathcal{C} : \mathcal{C}_1]\, 1.$$

So, $\widetilde{\mathcal{A}}_1'' \subset \mathcal{A}''$ is λ-Markov with $\lambda = [\mathcal{C} : \mathcal{C}_1]^{-1}$. □

Proof of Theorem 3.9. By Lemma 3.10, we have

$$\beta_n^{(2)}(\mathcal{C}_1) = \dim_{\widetilde{\mathcal{A}}_1''} \mathrm{Tor}_n^{\widetilde{\mathcal{A}}_1}(L^2(\widetilde{\mathcal{A}}_1)^0, \mathbb{C}).$$

By Proposition 3.12, the left $\widetilde{\mathcal{A}}_1$-module $\mathcal{A}$ is projective. We can thus apply the base change formula for Tor, see for example [Weibel 1994, Proposition 3.2.9], and obtain the isomorphism of left $\widetilde{\mathcal{A}}_1''$-modules

$$\mathrm{Tor}_n^{\widetilde{\mathcal{A}}_1}(L^2(\widetilde{\mathcal{A}}_1)^0, \mathbb{C}) \cong \mathrm{Tor}_n^{\mathcal{A}}(L^2(\widetilde{\mathcal{A}}_1)^0 \otimes_{\widetilde{\mathcal{A}}_1} \mathcal{A}, \mathbb{C}).$$

The left counterpart of Proposition 3.12 provides an inverse for the natural right $\mathcal{A}$-linear map

$$L^2(\widetilde{\mathcal{A}}_1)^0 \otimes_{\widetilde{\mathcal{A}}_1} \mathcal{A} \to L^2(\mathcal{A})^0,$$

which is thus bijective. We conclude that

$$\mathrm{Tor}_n^{\widetilde{\mathcal{A}}_1}(L^2(\widetilde{\mathcal{A}}_1)^0, \mathbb{C}) \cong \mathrm{Tor}_n^{\mathcal{A}}(L^2(\mathcal{A})^0, \mathbb{C})$$

as left $\widetilde{\mathcal{A}}_1''$-modules.

By Proposition 3.12, the inclusion $\widetilde{\mathcal{A}}_1'' \subset \mathcal{A}''$ is λ-Markov with $\lambda = [\mathcal{C} : \mathcal{C}_1]^{-1}$ and trace preserving. Using Proposition 3.6, we conclude that

$$\beta_n^{(2)}(\mathcal{C}_1) = \dim_{\widetilde{\mathcal{A}}_1''} \operatorname{Tor}_n^{\mathcal{A}}(L^2(\mathcal{A})^0, \mathbb{C}) = [\mathcal{C} : \mathcal{C}_1] \dim_{\mathcal{A}''} \operatorname{Tor}_n^{\mathcal{A}}(L^2(\mathcal{A})^0, \mathbb{C}) = [\mathcal{C} : \mathcal{C}_1]\, \beta_n^{(2)}(\mathcal{C}). \qquad \square$$

Using our results on Markov inclusions, we give the following short proof of Proposition 3.3.

Proof of Proposition 3.3. Let $\mathcal{C}$ be a rigid C^*-tensor category with full C^*-tensor subcategories $\mathcal{C}_2 \subset \mathcal{C}_1 \subset \mathcal{C}$. Note that $[\mathcal{C} : \mathcal{C}_2] < \infty$ if and only if $\operatorname{Irr}(\mathcal{C})$ has finitely many $\mathcal{C}_2$-orbits in the sense of Definition 3.1. Since $\operatorname{Irr}(\mathcal{C})$ has finitely many $\mathcal{C}_2$-orbits if and only if $\operatorname{Irr}(\mathcal{C})$ has finitely many $\mathcal{C}_1$-orbits and $\operatorname{Irr}(\mathcal{C}_1)$ has finitely many $\mathcal{C}_2$-orbits, we may assume that the indices $[\mathcal{C} : \mathcal{C}_1]$, $[\mathcal{C} : \mathcal{C}_2]$ and $[\mathcal{C}_1 : \mathcal{C}_2]$ are all finite.

Define the $*$-subalgebras $\widetilde{\mathcal{A}}_1 \subset \mathcal{A}$ and $\widetilde{\mathcal{A}}_2 \subset \mathcal{A}$ given by (3-6), associated with $\mathcal{C}_1 \subset \mathcal{C}$ and $\mathcal{C}_2 \subset \mathcal{C}$, respectively. Note that $\widetilde{\mathcal{A}}_2 \subset \widetilde{\mathcal{A}}_1$. By Proposition 3.12, the inclusion $\widetilde{\mathcal{A}}_i'' \subset \mathcal{A}''$ is $[\mathcal{C} : \mathcal{C}_i]^{-1}$-Markov for $i = 1, 2$. We claim that $\widetilde{\mathcal{A}}_2'' \subset \widetilde{\mathcal{A}}_1''$ is $[\mathcal{C}_1 : \mathcal{C}_2]^{-1}$-Markov. This does not literally follow from Proposition 3.12, but the proof is identical because, choosing representatives $\alpha_1, \dots, \alpha_\kappa$ for the $\mathcal{C}_2$-orbits in $\operatorname{Irr}(\mathcal{C}_1)$, Lemma 3.11 implies

$$\left\{ \sqrt{\frac{d(j)\, d(\alpha_s)}{d([\bar{\alpha}_s \alpha_s]_{\mathcal{C}_2})}}\, W \;\middle|\; i, j \in \operatorname{Irr}(\mathcal{C}),\ s = 1, \dots, \kappa,\ W \in \operatorname{onb}(i\alpha_s, \alpha_s j) \right\}$$

is a Pimsner–Popa basis for the inclusion $\widetilde{\mathcal{A}}_2'' \subset \widetilde{\mathcal{A}}_1''$.

Since $\dim_{\mathcal{A}''}(p_\varepsilon L^2(\mathcal{A})) = 1$, a repeated application of Proposition 3.6 gives

$$\dim_{\widetilde{\mathcal{A}}_2''}(p_\varepsilon L^2(\mathcal{A})) = [\mathcal{C} : \mathcal{C}_2] \dim_{\mathcal{A}''}(p_\varepsilon L^2(\mathcal{A})) = [\mathcal{C} : \mathcal{C}_2],$$
$$\dim_{\widetilde{\mathcal{A}}_2''}(p_\varepsilon L^2(\mathcal{A})) = [\mathcal{C}_1 : \mathcal{C}_2] \dim_{\widetilde{\mathcal{A}}_1''}(p_\varepsilon L^2(\mathcal{A})) = [\mathcal{C}_1 : \mathcal{C}_2]\,[\mathcal{C} : \mathcal{C}_1] \dim_{\mathcal{A}''}(p_\varepsilon L^2(\mathcal{A})) = [\mathcal{C} : \mathcal{C}_1]\,[\mathcal{C}_1 : \mathcal{C}_2].$$

So, the equality $[\mathcal{C} : \mathcal{C}_2] = [\mathcal{C} : \mathcal{C}_1]\,[\mathcal{C}_1 : \mathcal{C}_2]$ is proved.

When $\mathcal{C}$ has only finitely many irreducible objects and $\mathcal{C}_1 \subset \mathcal{C}$ is a full C^*-tensor subcategory, we apply this formula to $\mathcal{C}_2 = \{\varepsilon\}$ and obtain

$$d(\mathcal{C}) = [\mathcal{C} : \mathcal{C}_2] = [\mathcal{C} : \mathcal{C}_1]\,[\mathcal{C}_1 : \mathcal{C}_2] = [\mathcal{C} : \mathcal{C}_1]\, d(\mathcal{C}_1).$$

So, $[\mathcal{C} : \mathcal{C}_1] = d(\mathcal{C})/d(\mathcal{C}_1)$. $\square$

4. L^2-Betti numbers for discrete quantum groups

Following Woronowicz [1998], a compact quantum group $\mathbb{G}$ is given by a unital C^*-algebra B, often suggestively denoted as $B = C(\mathbb{G})$, together with a unital $*$-homomorphism $\Delta : B \to B \otimes_{\min} B$ to the minimal C^*-tensor product satisfying

- coassociativity: $(\Delta \otimes \mathrm{id})\Delta = (\mathrm{id} \otimes \Delta)\Delta$, and
- the density conditions: $\Delta(B)(1 \otimes B)$ and $\Delta(B)(B \otimes 1)$ span dense subspaces of $B \otimes_{\min} B$.

A compact quantum group $\mathbb{G}$ admits a unique Haar state, i.e., a state h on B satisfying $(\mathrm{id} \otimes h)\Delta(b) = (h \otimes \mathrm{id})\Delta(b) = h(b)1$ for all $b \in B$.

An n-dimensional unitary representation U of $\mathbb{G}$ is a unitary element $U \in M_n(\mathbb{C}) \otimes B$ satisfying $\Delta(U_{ij}) = \sum_{k=1}^n U_{ik} \otimes U_{kj}$. The category of finite-dimensional unitary representations, denoted as $\mathrm{Rep}(\mathbb{G})$, naturally is a rigid C^*-tensor category. The coefficients $U_{ij} \in B$ of all finite-dimensional unitary representations of $\mathbb{G}$ span a dense $*$-subalgebra of B, denoted as $\mathrm{Pol}(\mathbb{G})$. We have $\Delta(\mathrm{Pol}(\mathbb{G})) \subset \mathrm{Pol}(\mathbb{G}) \otimes \mathrm{Pol}(\mathbb{G})$, which provides the comultiplication of the Hopf $*$-algebra $\mathrm{Pol}(\mathbb{G})$.

The compact quantum group $\mathbb{G}$ is said to be of Kac type if the Haar state is a trace. This is equivalent with the requirement that for every finite-dimensional unitary representation $U \in M_n(\mathbb{C}) \otimes B$, the contragredient $\overline{U} \in M_n(\mathbb{C}) \otimes B$ defined by $(\overline{U})_{ij} = U_{ij}^*$ is still unitary.

The counit of the Hopf $*$-algebra $\mathrm{Pol}(\mathbb{G})$ is the homomorphism $\varrho : \mathrm{Pol}(\mathbb{G}) \to \mathbb{C}$ given by $\varrho(U_{ij}) = 0$ whenever $i \neq j$ and $\varrho(U_{ii}) = 1$ for all unitary representations $U \in M_n(\mathbb{C}) \otimes B$ of $\mathbb{G}$.

We denote by $L^2(\mathbb{G})$ the Hilbert space completion of $B = C(\mathbb{G})$ with respect to the Haar state h. The von Neumann algebra generated by the left action of B on $L^2(\mathbb{G})$ is denoted as $L^\infty(\mathbb{G})$. The Haar state h extends to a faithful normal state on $L^\infty(\mathbb{G})$, which is a trace in the Kac case.

Definition 4.1 [Kyed 2008, Definition 1.1]. Let $\mathbb{G}$ be a compact quantum group of Kac type. The L^2-Betti numbers of the dual discrete quantum group $\widehat{\mathbb{G}}$ are defined as

$$\beta_n^{(2)}(\widehat{\mathbb{G}}) = \dim_{L^\infty(\mathbb{G})} \mathrm{Tor}_n^{\mathrm{Pol}(\mathbb{G})}(L^2(\mathbb{G}), \mathbb{C}).$$

The main result of this section is the following.

Theorem 4.2. *Let $\mathbb{G}$ be a compact quantum group of Kac type. Then $\beta_n^{(2)}(\widehat{\mathbb{G}}) = \beta_n^{(2)}(\mathrm{Rep}(\mathbb{G}))$ for all $n \geq 0$.*

The equality of L^2-Betti numbers in Theorem 4.2 is surprising. There is no general identification of (co)homology of $\widehat{\mathbb{G}}$ with (co)homology of $\mathrm{Rep}(\mathbb{G})$. Indeed, by [Collins et al. 2009, Theorem 3.2], homology with trivial coefficients distinguishes between the quantum groups $A_o(k)$, but it does not distinguish between their representation categories $\mathrm{Rep}(A_o(k))$ by Corollary 6.2 below. Secondly, for the definition of the L^2-Betti numbers of a discrete quantum group, the Kac assumption is essential, since we need a trace to measure dimensions. By Theorem 4.2, we now also have L^2-Betti numbers for discrete quantum groups that are not of Kac type.

Proof of Theorem 4.2. Define the $*$-algebra

$$c_c(\widehat{\mathbb{G}}) = \bigoplus_{U \in \mathrm{Irr}(\mathbb{G})} M_{d(U)}(\mathbb{C}).$$

Drinfeld's quantum double algebra of $\mathbb{G}$ is the $*$-algebra $\mathcal{A}$ with underlying vector space $\mathrm{Pol}(\mathbb{G}) \otimes c_c(\widehat{\mathbb{G}})$ and product determined as follows. We view $c_c(\widehat{\mathbb{G}}) \subset \mathrm{Pol}(\mathbb{G})^*$ in the usual way: the components of $\omega \in c_c(\widehat{\mathbb{G}})$ are given by $\omega_{U,ij} = \omega(U_{ij})$ for all $U \in \mathrm{Irr}(\mathbb{G})$ and $i, j \in \{1, \dots, d(U)\}$. We write $a\omega$ instead of $a \otimes \omega$ for all $a \in \mathrm{Pol}(\mathbb{G})$ and $\omega \in c_c(\widehat{\mathbb{G}})$. The product on $\mathcal{A}$ is then determined by the formula

$$\omega\, U_{ij} = \sum_{k,l=1}^{n} U_{kl}\, \omega(U_{ik} \cdot U_{jl}^*)$$

for every unitary representation $U \in M_n(\mathbb{C}) \otimes B$. The counit on $\mathcal{A}$ is given by $\varrho(a\omega) = \varrho(a)\omega(1)$ for all $a \in \mathrm{Pol}(\mathbb{G})$ and $\omega \in c_c(\widehat{\mathbb{G}}) \subset \mathrm{Pol}(\mathbb{G})^*$.

Since $\mathbb{G}$ is of Kac type, the Haar weight τ on $\mathcal{A}$ is a trace and it is given by

$$\tau(a\omega) = h(a) \sum_{U \in \mathrm{Irr}(\mathbb{G})} \sum_{i=1}^{d(U)} d(U)\, \omega_{U,ii}.$$

We denote by $\mathcal{A}''$ the von Neumann algebra completion of $\mathcal{A}$ acting on $L^2(\mathcal{A}, \tau)$. By [Neshveyev and Yamashita 2015, Theorem 2.4], the tube algebra of $\mathrm{Rep}(\mathbb{G})$ is strongly Morita equivalent with the quantum double algebra $\mathcal{A}$ defined in the previous paragraph. This strong Morita equivalence respects the counit and the traces on both algebras. Therefore,

$$\beta_n^{(2)}(\mathrm{Rep}(\mathbb{G})) = \dim_{\mathcal{A}''} \mathrm{Tor}_n^{\mathcal{A}}(L^2(\mathcal{A})^0, \mathbb{C}),$$

where $L^2(\mathcal{A})^0$ equals the span of $L^2(\mathcal{A}) \cdot c_c(\widehat{\mathbb{G}})$.

On the other hand,

$$\beta_n^{(2)}(\widehat{\mathbb{G}}) = \dim_{L^\infty(\mathbb{G})} \mathrm{Tor}_n^{\mathrm{Pol}(\mathbb{G})}(L^2(\mathbb{G}), \mathbb{C}).$$

Since $\mathcal{A}$ is a free left $\mathrm{Pol}(\mathbb{G})$-module, the base change formula for Tor again applies and gives the isomorphism of left $L^\infty(\mathbb{G})$-modules

$$\mathrm{Tor}_n^{\mathrm{Pol}(\mathbb{G})}(L^2(\mathbb{G}), \mathbb{C}) \cong \mathrm{Tor}_n^{\mathcal{A}}(L^2(\mathbb{G}) \otimes_{\mathrm{Pol}(\mathbb{G})} \mathcal{A}, \mathbb{C}).$$

Since $L^2(\mathbb{G}) \otimes_{\mathrm{Pol}(\mathbb{G})} \mathcal{A} = L^2(\mathbb{G}) \otimes c_c(\widehat{\mathbb{G}}) = L^2(\mathcal{A})^0$, we conclude that

$$\beta_n^{(2)}(\widehat{\mathbb{G}}) = \dim_{L^\infty(\mathbb{G})} \mathrm{Tor}_n^{\mathcal{A}}(L^2(\mathcal{A})^0, \mathbb{C}).$$

Denoting by $E_{U,ij}$ the natural matrix units for $c_c(\widehat{\mathbb{G}})$, we see that the elements

$$\{d(U)^{-1/2} E_{U,ij} \mid U \in \mathrm{Irr}(\mathbb{G}),\ i, j = 1, \ldots, d(U)\}$$

form a Pimsner–Popa basis for the inclusion $L^\infty(\mathbb{G}) \subset \mathcal{A}''$, which is not trace-preserving. It follows that this inclusion is 1-Markov. Since the left $\mathcal{A}''$-module $L^2(\mathcal{A})^0$ is locally finite (in the sense of Definition 3.5 and using the example given after Definition 3.5), using a bar resolution, one gets that also the left $\mathcal{A}''$-module $\mathrm{Tor}_n^{\mathcal{A}}(L^2(\mathcal{A})^0, \mathbb{C})$ is locally finite. Proposition 3.6 then implies

$$\dim_{L^\infty(\mathbb{G})} \mathrm{Tor}_n^{\mathcal{A}}(L^2(\mathcal{A})^0, \mathbb{C}) = \dim_{\mathcal{A}''} \mathrm{Tor}_n^{\mathcal{A}}(L^2(\mathcal{A})^0, \mathbb{C}). \qquad \square$$

Given a compact quantum group $\mathbb{G}$, all Hopf $*$-subalgebras of $\mathrm{Pol}(\mathbb{G})$ are of the form $\mathrm{Pol}(\mathbb{H}) \subset \mathrm{Pol}(\mathbb{G})$, where $\mathrm{Rep}(\mathbb{H}) \subset \mathrm{Rep}(\mathbb{G})$ is a full C^*-tensor subcategory. We say that $\mathrm{Pol}(\mathbb{H}) \subset \mathrm{Pol}(\mathbb{G})$ is of finite index if $\mathrm{Rep}(\mathbb{H}) \subset \mathrm{Rep}(\mathbb{G})$ is of finite index in the sense of Definition 3.1 and we define the index

$$[\mathrm{Pol}(\mathbb{G}) : \mathrm{Pol}(\mathbb{H})] := [\mathrm{Rep}(\mathbb{G}) : \mathrm{Rep}(\mathbb{H})]$$

using Definition 3.1.

For special types of finite-index Hopf $*$-subalgebras $\mathrm{Pol}(\mathbb{H}) \subset \mathrm{Pol}(\mathbb{G})$, the scaling formula between $\beta_n^{(2)}(\widehat{\mathbb{H}})$ and $\beta_n^{(2)}(\widehat{\mathbb{G}})$ was proved in [Bichon et al. 2017, Theorem D]. Combining Theorems 4.2 and 3.9, it holds in general.

Corollary 4.3. *Let $\mathbb{G}$ be a compact quantum group of Kac type. Let* $\mathrm{Pol}(\mathbb{H}) \subset \mathrm{Pol}(\mathbb{G})$ *be a finite-index Hopf $*$-subalgebra. Then,*

$$\beta_n^{(2)}(\widehat{\mathbb{H}}) = [\mathrm{Pol}(\mathbb{G}) : \mathrm{Pol}(\mathbb{H})]\, \beta_n^{(2)}(\widehat{\mathbb{G}}) \quad \textit{for all } n \geq 0.$$

Remark 4.4. Of course, Corollary 4.3 can be proven directly, using the same methods as in the proof of Theorem 3.9. Choosing representatives $U_1, \dots, U_\kappa$ for the right $\mathrm{Rep}(\mathbb{H})$-orbits in $\mathrm{Irr}(\mathbb{G})$, the appropriate multiples of $(U_s)_{ij}$ form a Pimsner–Popa basis for the inclusion $L^\infty(\mathbb{H}) \subset L^\infty(\mathbb{G})$. As in the proof of Proposition 3.12, it follows that $\mathrm{Pol}(\mathbb{G})$ is a projective $\mathrm{Pol}(\mathbb{H})$-module and that $L^\infty(\mathbb{H}) \subset L^\infty(\mathbb{G})$ is a λ-Markov inclusion with $\lambda = [\mathrm{Pol}(\mathbb{G}) : \mathrm{Pol}(\mathbb{H})]^{-1}$.

5. Computing L^2-Betti numbers of representation categories

For any invertible matrix $F \in \mathrm{GL}_m(\mathbb{C})$, the *free unitary quantum group* $A_u(F)$ is the universal C^*-algebra with generators U_{ij}, $1 \leq i, j \leq m$, and relations making the matrices U and $F\overline{U}F^{-1}$ unitary representations of $A_u(F)$; see [Van Daele and Wang 1996]. Here $(\overline{U})_{ij} = (U_{ij})^*$. We denote by $A_u(m)$ the free unitary quantum group given by the $m \times m$ identity matrix. The following is the main result of this section.

Theorem 5.1. *Let* $F \in \mathrm{GL}_m(\mathbb{C})$ *be an invertible matrix and* $\mathcal{C} = \mathrm{Rep}(A_u(F))$ *the representation category of the free unitary quantum group* $A_u(F)$*. Then,*

$$\beta_1^{(2)}(\mathcal{C}) = 1 \qquad \textit{and} \qquad \beta_n^{(2)}(\mathcal{C}) = 0 \quad \textit{for all } n \neq 1.$$

For $F \in \mathrm{GL}_m(\mathbb{C})$ with $F\overline{F} \in \mathbb{R}1$, the *free orthogonal quantum group* $A_o(F)$ is the universal C^*-algebra with generators U_{ij}, $1 \leq i, j \leq m$, and relations such that U is unitary and $U = F\overline{U}F^{-1}$. We denote by $A_o(m)$ the free orthogonal quantum group given by the $m \times m$ identity matrix. Also note that

$$\mathrm{SU}_q(2) = A_o\begin{pmatrix} 0 & -q \\ 1 & 0 \end{pmatrix}$$

for all $q \in [-1, 1] \setminus \{0\}$.

Using Theorem 4.2 in combination with several results of [Popa et al. 2017], we get the following computations of L^2-Betti numbers of discrete quantum groups.

Theorem 5.2. (i) [Bichon et al. 2017, Theorem A; Kyed and Raum 2017, Theorem A] *For all $m \geq 2$, we have that $\beta_n^{(2)}(\widehat{A_u(m)})$ is equal to 1 if $n = 1$ and equal to 0 if $n \neq 1$.*

(ii) [Collins et al. 2009, Theorem 1.2; Vergnioux 2012, Corollary 5.2] *We have that $\beta_n^{(2)}(\widehat{A_o(m)}) = 0$ for all $m \geq 2$ and $n \geq 0$.*

(iii) *Let (B, τ) be a finite-dimensional C^*-algebra with its Markov trace. Assume that* $\dim B \geq 4$ *and let $A_{\mathrm{aut}}(B, \tau)$ be the quantum automorphism group. Then, $\beta_n^{(2)}(\widehat{A_{\mathrm{aut}}(B, \tau)}) = 0$ for all $n \geq 0$. In particular, all L^2-Betti numbers vanish for the duals of the quantum symmetry groups S_m^+ with $m \geq 4$.*

(iv) *Let $\mathbb{G} = \mathbb{H} \wr_* \mathbb{F}$ be the free wreath product of a nontrivial Kac-type compact quantum group $\mathbb{H}$ and a quantum subgroup $\mathbb{F}$ of S_m^+ that is acting ergodically on m points, $m \geq 2$ (see Remark 5.3 for* *definitions and comments). Then* $\widehat{\mathbb{G}}$ *has the same* L^2*-Betti numbers as the free product* $\widehat{\mathbb{H}} * \widehat{\mathbb{F}}$*, namely*

$$\beta_n^{(2)}(\widehat{\mathbb{G}}) = \begin{cases} \beta_n^{(2)}(\widehat{\mathbb{H}}) + \beta_n^{(2)}(\widehat{\mathbb{F}}) & \text{if } n \geq 2, \\ \beta_1^{(2)}(\widehat{\mathbb{H}}) + \beta_1^{(2)}(\widehat{\mathbb{F}}) + 1 - (\beta_0^{(2)}(\widehat{\mathbb{H}}) + \beta_0^{(2)}(\widehat{\mathbb{F}})) & \text{if } n = 1, \\ 0 & \text{if } n = 0. \end{cases}$$

(v) *In particular, for the duals of the hyperoctahedral quantum group* H_m^+*,* $m \geq 4$*, and the series of quantum reflection groups* H_m^{s+}*,* $s \geq 2$ (see [Banica and Vergnioux 2009])*, all* L^2*-Betti numbers vanish, except* $\beta_1^{(2)}$*, which is resp. equal to* $1/2$ *and* $1 - 1/s$*.*

Proof of Theorem 5.1. By [Bichon et al. 2006, Theorem 6.2], the rigid C^*-tensor category $\operatorname{Rep}(A_u(F))$ only depends on the quantum dimension of the fundamental representation U. We may therefore assume that $F\bar{F} = \pm 1$. In [Bichon et al. 2016a, Examples 2.18 and 3.6] and [Bichon et al. 2016b, Proposition 1.2], it is shown that there are exact sequences of Hopf $*$-algebras

$$\mathbb{C} \to \operatorname{Pol}(\mathbb{H}) \to \operatorname{Pol}(A_o(F) * A_o(F)) \to \mathbb{C}[\mathbb{Z}/2\mathbb{Z}] \to \mathbb{C},$$
$$\mathbb{C} \to \operatorname{Pol}(\mathbb{H}) \to \operatorname{Pol}(A_u(F)) \to \mathbb{C}[\mathbb{Z}/2\mathbb{Z}] \to \mathbb{C}$$

for the same compact quantum group $\mathbb{H}$. At the categorical level, this means that $\operatorname{Rep}(A_u(F))$ and the free product $\operatorname{Rep}(A_o(F)) * \operatorname{Rep}(A_o(F))$ both contain the same index-2 subcategory; see also [Bichon et al. 2017, Section 2].

By the scaling formula in Theorem 3.9, this implies that $\operatorname{Rep}(A_u(F))$ and the free product $\operatorname{Rep}(A_o(F)) * \operatorname{Rep}(A_o(F))$ have the same L^2-Betti numbers. From the free product formula for L^2-Betti numbers in [Popa et al. 2017, Corollary 9.5] and the vanishing of the L^2-Betti numbers of $\operatorname{Rep}(A_o(F))$ proved in Theorem 9.9 of the same paper, the theorem follows. □

Proof of Theorem 5.2. Using Theorem 4.2, (i) follows from Theorem 5.1 and (ii) follows from [Popa et al. 2017, Theorem 9.9]. The representation categories of the quantum automorphism groups $A_{\text{aut}}(B, \tau)$ are monoidally equivalent with the natural index-2 full C^*-tensor subcategory of $\operatorname{Rep}(\mathrm{SU}_q(2))$. So (iii) follows from [Popa et al. 2017, Theorem 9.9] and the scaling formula in Theorem 3.9.

To prove (iv), let $\mathbb{G} = \mathbb{H} \wr_* \mathbb{F}$ be a free wreath product as in the formulation of the theorem. We use the notion of Morita equivalence of rigid C^*-tensor categories; see [Müger 2003, Section 4] and also [Popa et al. 2017, Definition 7.3]. By [Tarrago and Wahl 2016, Theorem B and Remark 7.6], $\operatorname{Rep}(\mathbb{G})$ is Morita equivalent in this sense with a free product C^*-tensor category $\mathcal{C} = \mathcal{C}_1 * \mathcal{C}_2$, where $\mathcal{C}_1$ is Morita equivalent with $\operatorname{Rep}(\mathbb{H})$ and $\mathcal{C}_2$ is Morita equivalent with $\operatorname{Rep}(\mathbb{F})$. To see this, one uses the observation in [Popa et al. 2017, Proposition 9.8] that for the Jones tower $N \subset M \subset M_1 \subset \cdots$ of a finite index subfactor $N \subset M$ and for arbitrary intermediate subfactors

$$M_a \subset P \subset M_n \subset M_{n+1} \subset Q \subset M_b$$

with $a \leq n < b$, the C^*-tensor category of P-P-bimodules generated by $P \subset Q$ is Morita equivalent with the C^*-tensor category of N-N-bimodules generated by the original subfactor $N \subset M$. Then, combining

[Popa et al. 2017, Proposition 7.4 and Corollary 9.5], we find that

$$\beta_n^{(2)}(\widehat{\mathbb{G}}) = \beta_n^{(2)}(\mathcal{C}_1 * \mathcal{C}_2) = \begin{cases} \beta_n^{(2)}(\mathcal{C}_1) + \beta_n^{(2)}(\mathcal{C}_2) & \text{if } n \geq 2, \\ \beta_1^{(2)}(\mathcal{C}_1) + \beta_1^{(2)}(\mathcal{C}_2) + 1 - (\beta_0^{(2)}(\mathcal{C}_1) + \beta_0^{(2)}(\mathcal{C}_2)) & \text{if } n = 1, \\ 0 & \text{if } n = 0. \end{cases}$$

Since $\beta_n^{(2)}(\mathcal{C}_1) = \beta_n^{(2)}(\operatorname{Rep}(\mathbb{H})) = \beta_n^{(2)}(\widehat{\mathbb{H}})$ for all $n \geq 0$, and similarly with $\beta_n^{(2)}(\mathcal{C}_2)$, statement (iv) is proved.

Finally, by [Banica and Vergnioux 2009, Theorem 3.4], the compact quantum groups H_m^{s+} can be viewed as the free wreath product $(\mathbb{Z}/s\mathbb{Z}) \wr_* S_m^+$ and H_m^+ corresponds to the case $s = 2$. So (v) follows from (iv). □

Remark 5.3. The free wreath products $\mathbb{G} = \mathbb{H} \wr_* \mathbb{F}$ were introduced in [Bichon 2004]. We recall the definition here. Denote by $U \in M_m(\mathbb{C}) \otimes C(\mathbb{F})$ the fundamental representation of the quantum group $\mathbb{F}$ acting on m points, so that the action of $\mathbb{F}$ on $\mathbb{C}^m$ is given by the $*$-homomorphism

$$\alpha : \mathbb{C}^m \to \mathbb{C}^m \otimes C(\mathbb{F}), \quad \alpha(e_j) = \sum_{i=1}^m e_i \otimes U_{ij}.$$

Then $C(\mathbb{G})$ is defined as the universal C^*-algebra generated by m copies of $C(\mathbb{H})$, denoted by $\pi_i(C(\mathbb{H}))$, $i = 1, \dots, m$, together with $C(\mathbb{F})$, and the relations saying that $\pi_i(C(\mathbb{H}))$ commutes with U_{ij} for all $i, j \in \{1, \dots, m\}$. The comultiplication Δ on $C(\mathbb{G})$ is defined by

$$\Delta(\pi_i(a)) = \sum_{j=1}^m ((\pi_i \otimes \pi_j)\Delta(a))\,(U_{ij} \otimes 1) \quad \text{and} \quad \Delta(U_{ij}) = \sum_{k=1}^m U_{ik} \otimes U_{kj}.$$

Now observe that it is essential to assume in Theorem 5.2(iv) that the action of $\mathbb{F}$ on $\mathbb{C}^m$ is ergodic, in the same way as it is essential to make this hypothesis in [Tarrago and Wahl 2016, Theorem B]. Indeed, in the extreme case where $\mathbb{F}$ is the trivial one-element group, we find that $C(\mathbb{H} \wr_* \mathbb{F})$ is the m-fold free product of $C(\mathbb{H})$, so that

$$\beta_1^{(2)}(\widehat{\mathbb{G}}) = \beta_1^{(2)}(\underbrace{\widehat{\mathbb{H}} * \cdots * \widehat{\mathbb{H}}}_{m \text{ times}}) = m(\beta_1^{(2)}(\widehat{\mathbb{H}}) - \beta_0^{(2)}(\widehat{\mathbb{H}})) + m - 1,$$

which is different from the value given by Theorem 5.2(iv), namely $\beta_1^{(2)}(\widehat{\mathbb{H}}) - \beta_0^{(2)}(\widehat{\mathbb{H}})$.

Remark 5.4. Let (B, τ) be a finite-dimensional C^*-algebra with its Markov trace and assume that $\mathbb{F}$ is a quantum subgroup of $A_{\text{aut}}(B, \tau)$ that is acting centrally ergodically on (B, τ). Given any Kac-type compact quantum group $\mathbb{H}$, [Tarrago and Wahl 2016, Definition 7.5 and Remark 7.6] provides an implicit definition of the free wreath product $\mathbb{H} \wr_* \mathbb{F}$. The formula in Theorem 5.2(iv) remains valid and gives the L^2-Betti numbers of $\mathbb{H} \wr_* \mathbb{F}$.

Remark 5.5. The fusion $*$-algebra $\mathbb{C}[\mathcal{C}]$ of a rigid C^*-tensor category $\mathcal{C}$ has a natural trace τ and counit $\varrho : \mathbb{C}[\mathcal{C}] \to \mathbb{C}$ and these coincide with the restriction of the trace and the counit of the tube algebra $\mathcal{A}$ to its corner $p_\varepsilon \cdot \mathcal{A} \cdot p_\varepsilon = \mathbb{C}[\mathcal{C}]$. The GNS construction provides a von Neumann algebra completion $L(\operatorname{Irr} \mathcal{C})$ of

$\mathbb{C}[\mathcal{C}]$ acting on $\ell^2(\mathrm{Irr}(\mathcal{C}))$ and having a natural faithful normal tracial state τ. So also the fusion $*$-algebra $\mathbb{C}[\mathcal{C}]$ admits L^2-Betti numbers defined by

$$\begin{aligned}\beta_n^{(2)}(\mathrm{Irr}(\mathcal{C})) &:= \dim_{L(\mathrm{Irr}(\mathcal{C}))} \mathrm{Tor}_n^{\mathbb{C}[\mathcal{C}]}(\ell^2(\mathrm{Irr}(\mathcal{C})), \mathbb{C}) \\ &= \dim_{L(\mathrm{Irr}(\mathcal{C}))} \mathrm{Ext}^n_{\mathbb{C}[\mathcal{C}]}(\mathbb{C}, \ell^2(\mathrm{Irr}(\mathcal{C}))).\end{aligned}$$

Answering a question posed by Dimitri Shlyakhtenko, we will show that the computation in Theorem 5.2(iv) provides the first examples of rigid C^*-tensor categories where $\beta_n^{(2)}(\mathrm{Irr}(\mathcal{C})) \neq \beta_n^{(2)}(\mathcal{C})$. Note that it was already observed in [Popa et al. 2017, comments after Proposition 9.13] that for the Temperley–Lieb–Jones category $\mathcal{C}$, the C^*-tensor category and the fusion $*$-algebra have different homology with trivial coefficients.

The first L^2-Betti number $\beta_1^{(2)}(\mathrm{Irr}(\mathcal{C}))$ can be computed as follows. Write $H = \ell^2(\mathrm{Irr}(\mathcal{C}))$. A linear map $d : \mathbb{C}[\mathcal{C}] \to H$ is called a 1-cocycle if $d(xy) = d(x)\, y + \varrho(x)\, d(y)$ for all $x, y \in \mathbb{C}[\mathcal{C}]$. A 1-cocycle d is called inner if there exists a vector $\xi \in H$ such that $d(x) = \xi\, x - \varrho(x)\, \xi$ for all $x \in \mathbb{C}[\mathcal{C}]$. Two 1-cocycles d_1 and d_2 are called cohomologous if $d_1 - d_2$ is inner. The space of 1-cocycles $Z^1(\mathbb{C}[\mathcal{C}], H)$ is a left $L(\mathrm{Irr}(\mathcal{C}))$-module and when $\mathrm{Irr}(\mathcal{C})$ is infinite, the subspace of inner 1-cocycles has $L(\mathrm{Irr}(\mathcal{C}))$-dimension equal to 1. In that case, one has

$$\beta_1^{(2)}(\mathrm{Irr}(\mathcal{C})) = -1 + \dim_{L(\mathrm{Irr}(\mathcal{C}))} Z^1(\mathbb{C}[\mathcal{C}], H).$$

Let Γ be any countable group and define $\mathbb{G} = \widehat{\Gamma} \wr_* \mathbb{Z}/2\mathbb{Z}$. The fusion rules on $\mathrm{Irr}(\mathbb{G})$ were determined in [Lemeux 2014] and are given as follows. Denote by $v_1 \in \widehat{\mathbb{Z}/2\mathbb{Z}}$ the unique nontrivial element and define $W \subset \Gamma * \widehat{\mathbb{Z}/2\mathbb{Z}}$ as the set of reduced words

$$g_0 v_1 g_1 v_1 \cdots v_1 g_{n-1} v_1 g_n, \quad n \geq 0,\ g_0, \ldots, g_n \in \Gamma \setminus \{e\},$$

that start and end with a letter from $\Gamma \setminus \{e\}$. Then $\mathrm{Irr}(\mathbb{G})$ can be identified with the set consisting of the trivial representation v_0, the 1-dimensional representation v_1 and a set of 2-dimensional representations $v(\varepsilon, g, \delta)$ for $\varepsilon, \delta \in \{\pm\}$ and $g \in \Gamma \setminus \{e\}$. The fusion rules are given by

$$\begin{aligned} v_1 \otimes v(\varepsilon, g, \delta) &= v(-\varepsilon, g, \delta), \\ v(\varepsilon, g, \delta) \otimes v_1 &= v(\varepsilon, g, -\delta), \\ v(\varepsilon, g, \delta) \otimes v(\varepsilon', h, \delta') &= \begin{cases} v(\varepsilon, g v_1 h, \delta') \oplus v(\varepsilon, gh, \delta') & \text{if } gh \neq e, \\ v(\varepsilon, g v_1 h, \delta') \oplus v_1 \oplus v_0 & \text{if } gh = e. \end{cases}\end{aligned}$$

Write $H = \ell^2(\mathrm{Irr}(\mathbb{G}))$. Given an arbitrary family of vectors $(\xi_g)_{g \in \Gamma\setminus\{e\}}$ in H, one checks that there is a uniquely defined 1-cocycle $d : \mathbb{C}[\mathcal{C}] \to H$ satisfying $d(v_0) = d(v_1) = 0$ and $d(v(\varepsilon, g, \delta)) = \xi_g$ for all $g \in \Gamma \setminus \{e\}$ and $\varepsilon, \delta \in \{\pm\}$. Moreover, this provides exactly the 1-cocycles that vanish on v_0 and v_1. Every 1-cocycle is cohomologous to a 1-cocycle vanishing on v_0, v_1, and the inner 1-cocycles vanishing on v_0, v_1 have $L(\mathrm{Irr}(\mathbb{G}))$-dimension $1/2$. It follows that

$$\beta_1^{(2)}(\mathrm{Irr}(\mathbb{G})) = |\Gamma| - 1 - \tfrac{1}{2} = |\Gamma| - \tfrac{3}{2}.$$

On the other hand, by Theorem 5.2(iv), we have

$$\beta_1^{(2)}(\mathrm{Rep}(\mathbb{G})) = \beta_1^{(2)}(\Gamma) - \beta_0^{(2)}(\Gamma) + \tfrac{1}{2}.$$

Taking $\Gamma = \mathbb{Z}$, we find an example where $\beta_1^{(2)}(\mathrm{Irr}(\mathbb{G})) = \infty$, while $\beta_1^{(2)}(\mathrm{Rep}(\mathbb{G})) = 1/2$. Taking $\Gamma = \mathbb{Z}/2\mathbb{Z}$, we find an example where $\mathrm{Rep}(\mathbb{G})$ is an amenable C^*-tensor category, but yet $\beta_1^{(2)}(\mathrm{Irr}(\mathbb{G})) = 1/2 \neq 0$. Although amenability can be expressed as a property of the fusion rules together with the counit (which provides the dimensions of the irreducible objects), amenability does not ensure that the fusion $*$-algebra has vanishing L^2-Betti numbers. In particular, the Cheeger–Gromov argument given in [Popa et al. 2017, Theorem 8.8] does not work on the level of the fusion $*$-algebra. In the above example, $\mathrm{Rep}(\mathbb{G})$ is Morita equivalent to the group $\Gamma * \mathbb{Z}/2\mathbb{Z}$. So also invariance of L^2-Betti numbers under Morita equivalence does not work on the level of the fusion $*$-algebra. All in all, this illustrates that it is not very natural to consider L^2-Betti numbers for fusion algebras.

6. Projective resolution for the Temperley–Lieb–Jones category

Fix $q \in [-1, 1] \setminus \{0\}$ and realize the Temperley–Lieb–Jones category $\mathcal{C}$ as the representation category $\mathcal{C} = \mathrm{Rep}(\mathrm{SU}_q(2))$. Denote by $\mathcal{A}$ the tube algebra of $\mathcal{C}$ together with its counit $\varrho : \mathcal{A} \to \mathbb{C}$.

Although it was proved in [Popa et al. 2017, Theorem 9.9] that $\beta_n^{(2)}(\mathcal{C}) = 0$ for all $n \geq 0$, an easy projective resolution of $\varrho : \mathcal{A} \to \mathbb{C}$ was not given in that paper. On the other hand, [Bichon 2013, Theorem 5.1] provides a length-3 projective resolution for the counit of $\mathrm{Pol}(\mathrm{SU}_q(2))$. In the case of $\mathrm{Pol}(A_o(m))$, this projective resolution was already found in [Collins et al. 2009, Theorem 1.1], but the proof of its exactness was very involved and ultimately relied on a long, computer-assisted Gröbner base calculation. The proof in [Bichon 2013] is much simpler and moreover gives a resolution by so-called Yetter–Drinfeld modules. This means that it is actually a length-3 projective resolution for the quantum double algebra of $\mathrm{SU}_q(2)$. By [Neshveyev and Yamashita 2015, Theorem 2.4], this quantum double algebra is strongly Morita equivalent with the tube algebra $\mathcal{A}$. The following is thus an immediate consequence of [Bichon 2013, Theorem 5.1].

Theorem 6.1. *Label by $(v_n)_{n\in\mathbb{N}}$ the irreducible objects of $\mathcal{C} = \mathrm{Rep}(\mathrm{SU}_q(2))$ and denote by $(p_n)_{n\in\mathbb{N}}$ the corresponding projections in $\mathcal{A}$.*

Decomposing $v_1 v_1 = v_0 \oplus v_2$, the identity operator $1 \in ((v_1 v_1) v_1, v_1(v_1 v_1))$ defines a unitary element $V \in (p_0 + p_2) \cdot \mathcal{A} \cdot (p_0 + p_2)$. Denoting by $\tau \in \{\pm 1\}$ the sign of q, the sequence

$$0 \to \mathcal{A} \cdot p_0 \xrightarrow{W \mapsto W \cdot p_0 \cdot (V+\tau)} \mathcal{A} \cdot (p_0 + p_2) \xrightarrow{W \mapsto W \cdot (V-\tau)} \mathcal{A} \cdot (p_0 + p_2) \xrightarrow{W \mapsto W \cdot (V+\tau) \cdot p_0} \mathcal{A} \cdot p_0 \xrightarrow{\varrho} \mathbb{C}$$

is a resolution of $\varrho : \mathcal{A} \to \mathbb{C}$ by projective left $\mathcal{A}$-modules.

As a consequence of Theorem 6.1, we immediately find the (co)homology of $\mathcal{C} = \mathrm{Rep}(\mathrm{SU}_q(2))$ with trivial coefficients $\mathbb{C}$, which was only computed up to degree 2 in [Popa et al. 2017, Proposition 9.13]. The same result was found in an unpublished note of Y. Arano using different methods.

Corollary 6.2. *For $\mathcal{C} = \mathrm{Rep}(\mathrm{SU}_q(2))$, the homology $H_n(\mathcal{C}, \mathbb{C})$ and cohomology $H^n(\mathcal{C}, \mathbb{C})$ with trivial coefficients are given by $\mathbb{C}$ when $n = 0, 3$ and are 0 when $n \notin \{0, 3\}$.*

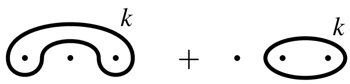
Remark 6.3. It is straightforward to check that inside $\mathcal{A}$, we have $p_0 \cdot V \cdot V = p_0$ and $\varrho(V) = -\tau$. Therefore, the composition of two consecutive arrows in Theorem 6.1 indeed gives the zero map. Using the diagrammatic representation of the tube algebra $\mathcal{A}$ given in [Ghosh and Jones 2016, Section 5.2], there are natural vector space bases for $\mathcal{A} \cdot p_0$ and $\mathcal{A} \cdot (p_0 + p_2)$. It is then quite straightforward to check that the sequence in Theorem 6.1 is indeed exact.

Using the same bases, one also checks that the tensor product of this resolution with $L^2(\mathcal{A}) \otimes_{\mathcal{A}} \cdot$ stays dimension exact. This then provides a slightly more elementary proof that $\beta_n^{(2)}(\mathcal{C}) = 0$ for all $n \geq 0$, as was already proved in [Popa et al. 2017, Theorem 9.9].

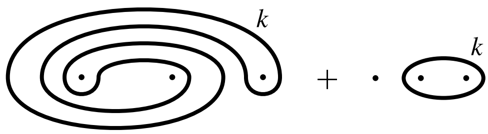
Remark 6.4. Section 9.5 of [Popa et al. 2017] provides a diagrammatic complex to compute the homology $H_n(\mathcal{C}, \mathbb{C})$ with trivial coefficients. In the particular case where $\mathcal{C}$ is the Temperley–Lieb–Jones category $\mathrm{TLJ}(\delta) = \mathrm{Rep}(\mathrm{SU}_q(2))$ with $-1 < q < 0$ and $\delta = -q - 1/q$, the space of n-chains is given by the linear span of all configurations of nonintersecting circles embedded into the plane with n points removed. Using Theorem 6.1, one computes that the 3-homology is spanned by the 3-cycle

$c_1 =$ 

It is however less clear how to write effectively a generating 3-cocycle in this diagrammatic language. For instance, for every integer $k \geq 1$, indicating by k the number of parallel strings,

$c_k =$ 

and

$d_k =$ 

are 3-cycles and ad hoc computations show that in 3-homology, we have $c_k = k\delta^{k-1}\, c_1$ and $d_k = 3k\delta^{k-1}\, c_1$. It would be interesting to have a geometric procedure to identify a given 3-cycle with a multiple of c_1 and to prove geometrically that homology vanishes in higher degrees.

7. Derivations on rigid C^*-tensor categories

7A. ***A Drinfeld-type central element in the tube algebra.*** To describe the first cohomology of a rigid C^*-tensor category $\mathcal{C}$ by a space of derivations, a natural element in the center of the tube algebra (more precisely, in the center of its multiplier algebra) plays a crucial role. In the case where $\mathcal{C}$ has only finitely many irreducible objects and hence, the tube algebra $\mathcal{A}$ is a direct sum of matrix algebras, this Drinfeld-type central element was introduced in [Izumi 2000, Theorem 3.3]. When $\mathcal{C}$ has infinitely many irreducible objects, the same definition applies and yields the following central unitary U in the multiplier algebra $M(\mathcal{A})$ defined by unitary elements $U_i \in p_i \cdot \mathcal{A} \cdot p_i$.

Fix a rigid C^*-tensor category $\mathcal{C}$. For every $i \in \mathrm{Irr}(\mathcal{C})$, denote by $U_i \in p_i \cdot \mathcal{A} \cdot p_i$ the element defined by the identity map in (ii, ii).

Proposition 7.1. *Fix $i, j \in \mathrm{Irr}(\mathcal{C})$. Then $U_i^\# = s_i t_i^*$ and $U_i \cdot U_i^\# = U_i^\# \cdot U_i = p_i$. In other words, U_i is unitary in $p_i \cdot \mathcal{A} \cdot p_i$. Moreover, for any $\alpha \in \mathrm{Irr}(\mathcal{C})$ and $V \in (i\alpha, \alpha j)$, the following relation holds:*

$$U_i \cdot V = V \cdot U_j = \sum_{\gamma \in \mathrm{Irr}(\mathcal{C})} d(\gamma) \sum_{W \in \mathrm{onb}(i\alpha,\gamma),\, W' \in \mathrm{onb}(\alpha j,\gamma)} \langle V, WW'^* \rangle\, (1 \otimes W'^*)(W \otimes 1). \tag{7-1}$$

So, $U := \sum_{i \in I} U_i$ is a central unitary element in the multiplier algebra $M(\mathcal{A})$.

Proof. By definition of the involution on $\mathcal{A}$, we have

$$U_i^\# = (t_i^* \otimes 1 \otimes 1)(1 \otimes 1 \otimes s_i) = s_i t_i^* \in (i\bar{i}, \bar{i}i).$$

Given this, one finds that

$$\begin{aligned} U_i^\# \cdot U_i &= \sum_{\gamma \in \mathrm{Irr}(\mathcal{C})} d(\gamma) \sum_{W \in \mathrm{onb}(\bar{i}i,\gamma)} (1 \otimes W^*)(U_i^\# \otimes 1)(1 \otimes U_i)(W \otimes 1) \\ &= \sum_{\gamma \in \mathrm{Irr}(\mathcal{C})} d(\gamma) \sum_{W \in \mathrm{onb}(\bar{i}i,\gamma)} (1 \otimes W^*)(s_i t_i^* W \otimes 1). \end{aligned}$$

Note that all terms with $\gamma \neq \varepsilon$ vanish. Hence, to conclude the computation, it suffices to note that $\{d(i)^{-1/2} t_i\}$ is an orthonormal basis for $(\bar{i}i, \varepsilon)$. Similarly, one checks that $U_i \cdot U_i^\# = p_i$.

Choose $V \in p_i \cdot \mathcal{A} \cdot p_j$ arbitrarily. Then

$$U_i \cdot V = \sum_{\gamma \in \mathrm{Irr}(\mathcal{C})} \sum_{W \in \mathrm{onb}(i\alpha,\gamma)} d(\gamma)(1 \otimes W^* V)(W \otimes 1).$$

On the other hand,

$$V \cdot U_j = \sum_{\gamma \in \mathrm{Irr}(\mathcal{C})} \sum_{W' \in \mathrm{onb}(j\alpha,\gamma)} d(\gamma)(1 \otimes W'^*)(VW' \otimes 1).$$

From these identities, one readily deduces (7-1) by expanding W^*V (resp. VW') in terms of the other orthonormal basis and using that the scalar products are given by the categorical traces. □

Note that (7-1), along with the fact that $U_\varepsilon = p_\varepsilon$ in particular implies

$$U_i \cdot V = V \quad \text{and} \quad W \cdot U_i = W \tag{7-2}$$

for $V \in p_i \cdot \mathcal{A} \cdot p_\varepsilon$ and $W \in p_\varepsilon \cdot \mathcal{A} \cdot p_i$. As another corollary of (7-1), we find that $U = \sum_{i \in \mathrm{Irr}(\mathcal{C})} U_i$ belongs to the center of the von Neumann algebra $\mathcal{A}''$.

7B. *Properties of 1-cocycles.* Let $\mathcal{C}$ be a rigid C^*-tensor category with tube algebra $\mathcal{A}$. Fix a nondegenerate right Hilbert $\mathcal{A}$-module $\mathcal{K}$. As in [Popa et al. 2017], define the bar complex for Hochschild (co)homology as follows. Denote by $\mathcal{B}$ the linear span of the projections p_i, $i \in \mathrm{Irr}(\mathcal{C})$. Then define

$$C_n = p_\varepsilon \cdot \mathcal{A} \otimes_{\mathcal{B}} \underbrace{\mathcal{A} \otimes_{\mathcal{B}} \cdots \otimes_{\mathcal{B}} \mathcal{A}}_{n \text{ factors}}$$

with boundary maps $\partial : C_n \to C_{n-1}$ given by $\partial = \sum_{k=0}^{n}(-1)^k \partial_k$, where

$$\partial_k(V_0 \otimes \cdots \otimes V_n) = \begin{cases} \varrho(V_0) p_\varepsilon \cdot V_1 \otimes \cdots \otimes V_n, & k = 0, \\ V_0 \otimes \cdots \otimes V_{k-1} \cdot V_k \otimes \cdots \otimes V_n, & 1 \le k \le n. \end{cases}$$

This is a resolution of the trivial right $\mathcal{A}$-module $\mathbb{C}$ by projective right $\mathcal{A}$-modules. So $H^n(\mathcal{C}, \mathcal{K}^0)$ is the n-th cohomology of the dual complex

$$\mathrm{Hom}_{\mathcal{A}}(C_n, \mathcal{K}^0) = \mathrm{Hom}_{\mathcal{A}}(p_\varepsilon \cdot \mathcal{A} \otimes_{\mathcal{B}} \underbrace{\mathcal{A} \otimes_{\mathcal{B}} \cdots \otimes_{\mathcal{B}} \mathcal{A}}_{n \text{ factors}}, \mathcal{K}^0). \tag{7-3}$$

The complex in (7-3) is isomorphic with the complex

$$\widetilde{C}^n = \mathrm{Hom}_{\mathcal{B}}(p_\varepsilon \cdot \mathcal{A} \otimes_{\mathcal{B}} \underbrace{\mathcal{A} \otimes_{\mathcal{B}} \cdots \otimes_{\mathcal{B}} \mathcal{A}}_{n-1 \text{ factors}}, \mathcal{K}^0),$$

where $\widetilde{C}^0 = \mathcal{K} \cdot p_\varepsilon$. For $n \ge 1$, the coboundary maps $\partial : \widetilde{C}^n \to \widetilde{C}^{n+1}$ of this complex are given by $\partial = \sum_{k=0}^{n+1}(-1)^k \partial_k$, where

$$\partial_k(D)(V_0 \otimes \cdots \otimes V_n) = \begin{cases} \varrho(V_0) D(p_\varepsilon \cdot V_1 \otimes \cdots \otimes V_n), & k = 0, \\ D(V_0 \otimes \cdots \otimes V_{k-1} \cdot V_k \otimes \cdots \otimes V_n), & 1 \le k \le n, \\ D(V_0 \otimes \cdots \otimes V_{n-1}) \cdot V_n, & k = n+1. \end{cases}$$

The zeroth coboundary map of $\widetilde{C}^\bullet$ is given by

$$\mathcal{K} \cdot p_\varepsilon \to \mathrm{Hom}_{\mathcal{B}}(p_\varepsilon \cdot \mathcal{A}, \mathcal{K}^0), \quad \xi \mapsto [D_\xi : V \mapsto \varrho(V)\xi - \xi \cdot V]. \tag{7-4}$$

In this picture, the 1-cocycles are precisely the maps $D \in \mathrm{Hom}_{\mathcal{B}}(p_\varepsilon \cdot \mathcal{A}, \mathcal{K}^0)$ that satisfy

$$D(V \cdot W) = D(V) \cdot W + \varrho(V) D(W) \tag{7-5}$$

for all $V \in p_\varepsilon \cdot \mathcal{A} \cdot p_i$ and $W \in p_i \cdot \mathcal{A} \cdot p_j$. We associate a cocycle D_ξ to every vector $\xi \in \mathcal{K} \cdot p_\varepsilon$ via (7-4). These are the *inner* 1-cocycles.

By analogy with the first L^2-Betti number for groups, we want to express how a 1-cocycle D is determined by its values on a generating set of objects of $\mathcal{C}$. So, we first need to specify how D can actually be evaluated on objects $\alpha \in \mathcal{C}$.

By the correspondence theorem from [Popa et al. 2017] discussed in Section 2B, we may suppose that the right Hilbert $\mathcal{A}$-module $\mathcal{K}$ arises from a unitary half-braiding $(X, \sigma) \in \mathcal{Z}(\text{ind-}\mathcal{C})$, where $X \in \text{ind-}\mathcal{C}$ satisfies $(X, i) = \mathcal{K} \cdot p_i$ for all $i \in \mathrm{Irr}(\mathcal{C})$.

For every $\alpha \in \mathrm{Irr}(\mathcal{C})$, we consider the vector subspace $\mathcal{A}_\alpha \subset \mathcal{A}$,

$$\mathcal{A}_\alpha = \bigoplus_{i,j \in \mathrm{Irr}(\mathcal{C})} (i\alpha, \alpha j). \tag{7-6}$$

Note that each $\mathcal{A}_\alpha$ is a $\mathcal{B}$-bimodule. We can then define the natural bijection

$$\mathrm{Hom}_{\mathcal{B}}(p_\varepsilon \cdot \mathcal{A}_\alpha, \mathcal{K}^0) \cong (\alpha X, \alpha)$$

identifying $D \in \mathrm{Hom}_{\mathcal{B}}(p_\varepsilon \cdot \mathcal{A}_\alpha, \mathcal{K}^0)$ with $D_\alpha \in (\alpha X, \alpha)$ through the formulae

$$D(V) = (\mathrm{Tr}_\alpha \otimes \mathrm{id})(D_\alpha V) \quad \text{and} \quad D_\alpha = \sum_{j \in \mathrm{Irr}(\mathcal{C})} \sum_{W \in \mathrm{onb}(\alpha, \alpha j)} d(j)(1 \otimes D(W))W^*$$

for all $i \in \mathrm{Irr}(\mathcal{C})$ and $V \in (\alpha, \alpha i)$. Putting all $\alpha \in \mathrm{Irr}(\mathcal{C})$ together, we find a bijection

$$\mathrm{Hom}_{\mathcal{B}}(p_\varepsilon \cdot \mathcal{A}, \mathcal{K}^0) \cong \prod_{\alpha \in \mathrm{Irr}(\mathcal{C})} (\alpha X, \alpha)$$

identifying $D \in \mathrm{Hom}_{\mathcal{B}}(p_\varepsilon \cdot \mathcal{A}, \mathcal{K}^0)$ with the family $(D_\alpha)_{\alpha \in \mathrm{Irr}(\mathcal{C})}$.

Given a family of elements $D_\alpha \in (\alpha X, \alpha)$ for all $\alpha \in \mathrm{Irr}(\mathcal{C})$, we uniquely define $D_\beta \in (\beta X, \beta)$ for arbitrary objects $\beta \in \mathcal{C}$ by the formula

$$D_\beta = \sum_{\alpha \in \mathrm{Irr}(\mathcal{C})} \sum_{V \in \mathrm{onb}(\alpha, \beta)} d(\alpha)(V^* \otimes 1) D_\alpha V.$$

Note that the naturality condition

$$D_\alpha V = (V \otimes 1) D_\beta \tag{7-7}$$

holds for all $\alpha, \beta \in \mathcal{C}$ and all $V \in (\alpha, \beta)$.

Definition 7.2. Let $\mathcal{C}$ be a rigid C^*-tensor category. We say that a subset $\mathcal{G} \subset \mathrm{Irr}(\mathcal{C})$ *generates* $\mathcal{C}$ when every irreducible object in $\mathcal{C}$ arises as a subobject of some tensor product of elements in $\mathcal{G} \cup \overline{\mathcal{G}}$.

The following proposition implies that a 1-cocycle $D \in \mathrm{Hom}_{\mathcal{B}}(p_\varepsilon \cdot \mathcal{A}, \mathcal{K}^0)$ is completely determined by its "values" $D_\alpha \in (\alpha X, \alpha)$ for α belonging to a generating set $\mathcal{G} \subset \mathrm{Irr}(\mathcal{C})$.

Proposition 7.3. *Consider a morphism $D \in \mathrm{Hom}_{\mathcal{B}}(p_\varepsilon \cdot \mathcal{A}, \mathcal{K}^0)$ with corresponding values $D_\alpha \in (\alpha X, \alpha)$, $\alpha \in \mathcal{C}$. Then D is a 1-cocycle if and only if*

$$D_{\alpha\beta} = (1 \otimes \sigma_\beta^*)(D_\alpha \otimes 1) + (1 \otimes D_\beta) \tag{7-8}$$

for all $\alpha, \beta \in \mathcal{C}$. In particular, any 1-cocycle D satisfies $D_\varepsilon = 0$ and

$$D_{\bar\alpha} = -\sigma_{\bar\alpha}^*(t_\alpha^* \otimes 1 \otimes 1)(1 \otimes D_\alpha \otimes 1)(1 \otimes s_\alpha) = -(1 \otimes s_\alpha^* \otimes 1)(1 \otimes 1 \otimes \sigma_{\bar\alpha}^*)(1 \otimes D_\alpha \otimes 1)(t_\alpha \otimes 1) \tag{7-9}$$

for all $\alpha \in \mathrm{Irr}(\mathcal{C})$.

Proof. Choose arbitrary morphisms $V \in (\alpha, \alpha i)$ and $W \in (i\beta, \beta j)$. The following identities can be verified by direct computation:

$$\begin{aligned} D(V \cdot W) &= (\mathrm{Tr}_{\alpha\beta} \otimes \mathrm{id})\big(D_{\alpha\beta}(V \otimes 1)(1 \otimes W)\big), \\ D(V) \cdot W &= (\mathrm{Tr}_{\alpha\beta} \otimes \mathrm{id})\big((1 \otimes \sigma_\beta^*)(D_\alpha \otimes 1)(V \otimes 1)(1 \otimes W)\big), \\ \varrho(V) D(W) &= (\mathrm{Tr}_{\alpha\beta} \otimes \mathrm{id})\big((1 \otimes D_\beta)(V \otimes 1)(1 \otimes W)\big). \end{aligned}$$

By Frobenius reciprocity, for every fixed $\alpha, \beta, j \in \mathrm{Irr}(\mathcal{C})$, the linear span of all $(V \otimes 1)(1 \otimes W)$ with $i \in \mathrm{Irr}(\mathcal{C})$, $V \in (\alpha, \alpha i)$, $W \in (i\beta, \beta j)$, equals $(\alpha\beta, \alpha\beta j)$. So it follows that D is a 1-cocycle if and only if (7-8) holds for all $\alpha, \beta \in \mathcal{C}$.

Finally, assume that D is a 1-cocycle. By (7-8), we get $D_\varepsilon = 0$. The naturality property of the D_α implies $D_{\alpha\bar\alpha} s_\alpha = 0$ for all $\alpha \in \mathcal{C}$. So,

$$(1 \otimes 1 \otimes D_{\bar\alpha})(1 \otimes s_\alpha) = -(1 \otimes 1 \otimes \sigma_{\bar\alpha}^*)(1 \otimes D_\alpha \otimes 1)(1 \otimes s_\alpha),$$

which yields one half of (7-9) after multiplying by $(t_\alpha^* \otimes 1)$ on both sides. The other identity is proven similarly, by observing that $(s_\alpha^* \otimes 1) D_{\alpha\bar\alpha} = 0$. □

The following lemma shows that the constraint (7-9) on D_α can be succinctly restated in terms of the special unitaries $U_i \in \mathcal{A}$ introduced in the previous section.

Lemma 7.4. *Fix $\alpha \in \operatorname{Irr}(\mathcal{C})$ and consider $\mathcal{A}_\alpha \subset \mathcal{A}$ as in (7-6). Let $D \in \operatorname{Hom}_{\mathcal{B}}(p_\varepsilon \cdot \mathcal{A}_\alpha, \mathcal{K}^0)$ with corresponding $D_\alpha \in (\alpha X, \alpha)$. Then D_α satisfies the relation*

$$\sigma_{\bar\alpha}^*(t_\alpha^* \otimes 1 \otimes 1)(1 \otimes D_\alpha \otimes 1)(1 \otimes s_\alpha) = (1 \otimes s_\alpha^* \otimes 1)(1 \otimes 1 \otimes \sigma_{\bar\alpha}^*)(1 \otimes D_\alpha \otimes 1)(t_\alpha \otimes 1) \tag{7-10}$$

if and only if $D(V) = D(V) \cdot U_i$ for all $i \in \operatorname{Irr}(\mathcal{C})$ and $V \in (\alpha, \alpha i)$.

Proof. By definition of the $\mathcal{A}$-module structure on $\mathcal{K}$, for every $i \in \operatorname{Irr}(\mathcal{C})$ and $V \in (\alpha, \alpha i)$, we have

$$\begin{aligned} D(V) \cdot U_i &= (\operatorname{Tr}_\alpha \otimes \mathrm{id})(D_\alpha V) \cdot U_i = (\operatorname{Tr}_{\alpha i} \otimes \mathrm{id})\big((1 \otimes \sigma_i^*)(D_\alpha V \otimes 1)\big) \\ &= (\operatorname{Tr}_\alpha \otimes \mathrm{id})\big((V \otimes 1)(1 \otimes \sigma_i^*)(D_\alpha \otimes 1)\big) = (\operatorname{Tr}_\alpha \otimes \mathrm{id})\big((V \otimes 1)\sigma_{\alpha i}^*(\sigma_\alpha D_\alpha \otimes 1)\big) \\ &= (\operatorname{Tr}_\alpha \otimes \mathrm{id})\big(\sigma_\alpha^*(1 \otimes V)(\sigma_\alpha D_\alpha \otimes 1)\big), \end{aligned}$$

where the final two equalities follow from the half-braiding property and the naturality of σ, respectively. Writing V as $V = (s_\alpha^* \otimes 1)(1 \otimes W)$ with $W \in (\bar\alpha\alpha, i)$, we then find that

$$\begin{aligned} D(V) \cdot U_i &= (\operatorname{Tr}_\alpha \otimes \mathrm{id})\big(\sigma_\alpha^*(1 \otimes s_\alpha^* \otimes 1)(1 \otimes 1 \otimes W)(\sigma_\alpha D_\alpha \otimes 1)\big) \\ &= (\operatorname{Tr}_\alpha \otimes \mathrm{id})\big(\sigma_\alpha^*(1 \otimes s_\alpha^* \otimes 1)(\sigma_\alpha D_\alpha \otimes 1 \otimes 1)\big) W. \end{aligned}$$

Since $D(V) = (\operatorname{Tr}_\alpha \otimes \mathrm{id})(D_\alpha V) = (t_\alpha^* \otimes 1)(1 \otimes D_\alpha) W$, we conclude that the equality $D(V) = D(V) \cdot U_i$ for all $i \in \operatorname{Irr}(\mathcal{C})$ and $V \in (\alpha, \alpha i)$ is equivalent with the equality

$$(t_\alpha^* \otimes 1)(1 \otimes D_\alpha) = (t_\alpha^* \otimes 1)(1 \otimes \sigma_\alpha^*)(1 \otimes 1 \otimes s_\alpha^* \otimes 1)(1 \otimes \sigma_\alpha D_\alpha \otimes 1 \otimes 1)(t_\alpha \otimes 1 \otimes 1).$$

Applying the transformation $Y \mapsto \sigma_{\bar\alpha}^*(Y \otimes 1)(1 \otimes s_\alpha)$ to the left- and the right-hand sides, this equality becomes equivalent with (7-10). □

We can then formalize how a 1-cocycle is determined by its values on a generating set of a rigid C^*-tensor category as follows.

Proposition 7.5. *Let $\mathcal{C}$ be a rigid C^*-tensor category with finite generating set $\mathcal{G} \subset \operatorname{Irr}(\mathcal{C})$. Denote by $\mathcal{A}$ the tube algebra of $\mathcal{C}$ and let $\mathcal{K}$ be a nondegenerate right Hilbert $\mathcal{A}$-module. For every $i \in \operatorname{Irr}(\mathcal{C})$, define the subspace $\mathcal{K}_i^{\mathrm{fix}} \subset \mathcal{K} \cdot p_i$ given by*

$$\mathcal{K}_i^{\mathrm{fix}} := \{\xi \in \mathcal{K} \cdot p_i \mid \xi \cdot U_i = \xi\}.$$

Define

$$\widetilde{Z}^1(\mathcal{C}, \mathcal{K}^0) = \bigoplus_{\alpha \in \mathcal{G}} \bigoplus_{i \in \operatorname{Irr}(\mathcal{C})} \mathcal{K}_i^{\mathrm{fix}} \otimes (\alpha i, \alpha).$$

Then, the linear map

$$\Phi : Z^1(\mathcal{C}, \mathcal{K}^0) \to \widetilde{Z}^1(\mathcal{C}, \mathcal{K}^0), \quad D \mapsto \bigoplus_{\alpha\in\mathcal{G}} \bigoplus_{i\in\mathrm{Irr}(\mathcal{C})} \sum_{W\in\mathrm{onb}(\alpha,\alpha i)} D(W)\otimes W^*,$$

is injective. In particular, if $\mathcal{C}$ has infinitely many irreducible objects, we find the estimate

$$\beta_1^{(2)}(\mathcal{C}) \le -1 + \sum_{\alpha\in\mathcal{G}} \sum_{i\in\mathrm{Irr}(\mathcal{C})} \mathrm{mult}(i, \bar{\alpha}\alpha)\, \tau(q_i), \tag{7-11}$$

where $q_i \in p_i \cdot \mathcal{A}'' \cdot p_i$ denotes the projection onto the kernel of $U_i - p_i$.

Proof. By Proposition 7.3 and Lemma 7.4, the map Φ is well defined and injective. In the case where $\mathcal{K} = L^2(\mathcal{A})$, the map Φ is left $\mathcal{A}''$-linear. Since $\dim_{\mathcal{A}''} L^2(\mathcal{A})\cdot q_i = \tau(q_i)$, the proposition follows once we have proved that the space of inner 1-cocycles has $\mathcal{A}''$-dimension equal to 1, assuming that $\mathcal{C}$ has infinitely many irreducible objects.

In that case, $\beta_0^{(2)}(\mathcal{C}) = 0$ by [Popa et al. 2017, Corollary 9.2], meaning that the coboundary map

$$L^2(\mathcal{A})\cdot p_\varepsilon \to \mathrm{Hom}_{\mathcal{B}}(p_\varepsilon\cdot\mathcal{A}, L^2(\mathcal{A})^0)$$

is injective. The space of inner 1-cocycles is thus isomorphic with $L^2(\mathcal{A})\cdot p_\varepsilon$ and so, has $\mathcal{A}''$-dimension equal to 1. □

8. Derivations on $\mathrm{Rep}(A_u(F))$

In this section, we again specialize to the case of free unitary quantum groups. Let $F \in \mathrm{GL}_m(\mathbb{C})$. The methods of the previous section allow for a direct and explicit proof that $\beta_1^{(2)}(\mathrm{Rep}(A_u(F))) = 1$. More generally, we determine the first cohomology of $\mathrm{Rep}(A_u(F))$ with arbitrary coefficients.

By [Banica 1997], the category $\mathrm{Rep}(A_u(F))$ is freely generated by the fundamental representation u and the irreducible representations can be labeled by words in u and $\bar{u}$. To avoid confusion between words and tensor products, we explicitly write $\otimes$ to denote the tensor product of two representations. The tensor product $\bar{u}\otimes u$ decomposes as the sum of the trivial representation ε and the irreducible representation with label $\bar{u}u$. Similarly, $u\otimes\bar{u} \cong \varepsilon \oplus u\bar{u}$. Moreover, the standard solutions of the conjugate equations for u, given by $t_u \in (\bar{u}\otimes u, \varepsilon)$ and $s_u \in (u\otimes\bar{u}, \varepsilon)$, generate all intertwiners between tensor products of u and $\bar{u}$.

Proposition 8.1. *Let $F\in\mathrm{GL}_m(\mathbb{C})$ and $\mathcal{C} = \mathrm{Rep}(A_u(F))$, with tube algebra $\mathcal{A}$. Let $\mathcal{K}$ be any nondegenerate right Hilbert $\mathcal{A}$-module. Using the notation of Proposition 7.5, we find an isomorphism*

$$Z^1(\mathrm{Rep}(A_u(F)), \mathcal{K}^0) \cong \mathcal{K}\cdot p_\varepsilon \oplus \mathcal{K}^{\mathrm{fix}}_{\bar{u}u}. \tag{8-1}$$

Proof. As explained in Section 2B, we consider $\mathcal{K}$ as the nondegenerate right Hilbert $\mathcal{A}$-module given by a unitary half-braiding σ on some ind-object X. A vector on the right-hand side of (8-1) then corresponds to an element in $(u\otimes X, u)$ satisfying the conditions of Lemma 7.4, by Frobenius reciprocity. Fix such a morphism $D_u \in (u\otimes X, u)$. We have to show that D_u comes from a 1-cocycle $D\in\mathrm{Hom}_{\mathcal{B}}(p_\varepsilon\cdot\mathcal{A}, \mathcal{K}^0)$, which we will construct as a family of morphisms $(D_\alpha)_{\alpha\in\mathcal{C}}$ satisfying the naturality condition (7-7). The identity (7-9) forces us to define $D_{\bar{u}}$ by

$$D_{\bar{u}} = -\sigma_{\bar{u}}^*(t_u^* \otimes 1 \otimes 1)(1 \otimes D_u \otimes 1)(1 \otimes s_u).$$

This is unambiguous because $u \neq \bar{u}$ in $\mathrm{Rep}(A_u(F))$. By Lemma 7.4, we also have

$$D_{\bar{u}} = -(1 \otimes s_u^* \otimes 1)(1 \otimes 1 \otimes \sigma_{\bar{u}}^*)(1 \otimes D_u \otimes 1)(t_u \otimes 1).$$

The cocycle identity (7-8) imposes the definition

$$D_{\alpha_1 \otimes \cdots \otimes \alpha_n} = \sum_{k=1}^{n} (1^{\otimes k} \otimes \sigma^*_{\alpha_{k+1} \otimes \cdots \otimes \alpha_n})(1^{\otimes(k-1)} \otimes D_{\alpha_k} \otimes 1^{\otimes(n-k)}), \tag{8-2}$$

where $\alpha_k \in \{u, \bar{u}\}$. We also must set $D_\varepsilon = 0$. Since every irreducible object in $\mathrm{Rep}(A_u(F))$ is a subobject of some tensor product of u and $\bar{u}$, these relations fix D_α for all $\alpha \in \mathcal{C}$. Concretely, if $\alpha \in \mathrm{Irr}(\mathcal{C})$ and $w : \alpha_1 \otimes \cdots \otimes \alpha_n \to \alpha$ is a coisometry, where $\alpha_k \in \{u, \bar{u}\}$, we set

$$D_\alpha = (w \otimes 1) D_{\alpha_1 \otimes \cdots \otimes \alpha_n} w^*. \tag{8-3}$$

Now, since α appears in the decomposition of several different tensor products, it is not immediately clear why this is well defined. To this end, we will show that the naturality relation

$$(V \otimes 1) D_{\alpha_1 \otimes \cdots \otimes \alpha_n} = D_{\alpha'_1 \otimes \cdots \otimes \alpha'_m} V \tag{8-4}$$

holds for all morphisms

$$V : \alpha_1 \otimes \cdots \otimes \alpha_n \to \alpha'_1 \otimes \cdots \otimes \alpha'_m,$$

with $\alpha_i, \alpha'_j \in \{u, \bar{u}\}$. This is where the freeness of $\mathcal{C}$ comes into play. By [Banica 1997, Lemme 6], the intertwiner spaces between tensor products involving u and $\bar{u}$ are generated by maps of the forms $1^{\otimes i} \otimes s_u \otimes 1^{\otimes j}$ and $1^{\otimes i} \otimes t_u \otimes 1^{\otimes j}$ and their adjoints. Appealing to the naturality of σ in (8-2), it is therefore sufficient to verify that

$$D_{u \otimes \bar{u}} s_u = 0, \quad (s_u^* \otimes 1) D_{u \otimes \bar{u}} = 0,$$
$$D_{\bar{u} \otimes u} t_u = 0, \quad (t_u^* \otimes 1) D_{\bar{u} \otimes u} = 0,$$

which follows from the two different expressions for $D_{\bar{u}}$, by retracing the computations made in the proof of Proposition 7.3. We conclude that there exists a unique $D \in \mathrm{Hom}_{\mathcal{B}}(p_\varepsilon \cdot \mathcal{A}, \mathcal{K}^0)$ producing the family of maps $(D_\alpha)_{\alpha \in \mathcal{C}}$. This family satisfies the cocycle relation (7-8) by construction. Therefore D is a 1-cocycle, as required. □

Combining Propositions 7.5 and 8.1, we get

$$\beta_1^{(2)}(\mathrm{Rep}(A_u(F))) = \tau(q_{\bar{u}u}).$$

Calculating $\beta_1^{(2)}(\mathrm{Rep}(A_u(F)))$ therefore boils down to computing the trace of $q_{\bar{u}u}$. By von Neumann's mean ergodic theorem, we have

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \tau(U_{\bar{u}u}^k) = \tau(q_{\bar{u}u}). \tag{8-5}$$

In other words, to find the first L^2-Betti number of $\operatorname{Rep}(A_u(F))$, it is now sufficient to compute the traces $\tau(U_{\bar{u}u}^k)$ for all $k \in \mathbb{N}$, i.e., the sequence of moments of $U_{\bar{u}u}$.

The following lemma translates this problem into a combinatorial one.

Lemma 8.2. *Let $\mathcal{C}$ be an arbitrary rigid C^*-tensor category with tube algebra $\mathcal{A}$. For $\alpha \in \mathcal{C}$ and $k \geq 1$, define the rotation map*

$$\zeta_\alpha^k : (\alpha^k, \varepsilon) \to (\alpha^k, \varepsilon), \quad \xi \mapsto (1^{\otimes k} \otimes s_\alpha^*)(1 \otimes \xi \otimes 1) s_\alpha.$$

Then $\tau(U_i^k) = \overline{\operatorname{Tr}(\zeta_i^k)}$ for all $i \in \operatorname{Irr}(\mathcal{C})$, where Tr is the unnormalized trace on the finite-dimensional matrix algebra of linear transformations of (i^k, ε).

Proof. Fix $i \in \operatorname{Irr}(\mathcal{C})$ and observe that

$$\begin{aligned}\tau(U_i^k) &= \sum_{W \in \operatorname{onb}(i^k, \varepsilon)} \operatorname{Tr}_i((1 \otimes W^*)(W \otimes 1)) = \sum_{W \in \operatorname{onb}(i^k, \varepsilon)} s_i^*(1 \otimes W^* \otimes 1)(W \otimes 1 \otimes 1) s_i \\ &= \sum_{W \in \operatorname{onb}(i^k, \varepsilon)} s_i^*(1 \otimes W^* \otimes 1)(1^{\otimes k} \otimes s_i) W = \sum_{W \in \operatorname{onb}(i^k, \varepsilon)} \langle W, \zeta_i^k(W) \rangle = \overline{\operatorname{Tr}(\zeta_i^k)}.\end{aligned}$$

□

Proposition 8.3. *Consider either $\operatorname{Rep}(A_u(F))$ for $F \in \mathrm{GL}_m(\mathbb{C})$ or $\operatorname{Rep}(A_o(F))$ for $F \in \mathrm{GL}_m(\mathbb{C})$ with $F\overline{F} = \pm 1$. In both cases, denote by u the fundamental representation and let π be the nontrivial irreducible summand of $u \otimes \bar{u}$ or $\bar{u} \otimes u$. Then, for all $k \in \mathbb{Z}$, we have*

$$\tau(U_\pi^k) = \begin{cases} d(u)^2 - 1 & \text{if } k = 0, \\ 0 & \text{if } |k| = 1, \\ 1 & \text{if } |k| \geq 2. \end{cases} \tag{8-6}$$

So, $\tau(q_\pi) = 1$ and the spectral measure of U_π with respect to the (unnormalized) trace τ on $p_\pi \cdot \mathcal{A}'' \cdot p_\pi$ is given by $\delta_1 + (d(u)^2 - 2 - 2\operatorname{Re}(z))\, dz$, where dz denotes the normalized Lebesgue measure on the unit circle S^1.

Proof. We first deduce the result for $A_u(F)$ from the $A_o(F)$ case. Up to monoidal equivalence, we may assume that $F\overline{F} = \pm 1$. Consider the group $\mathbb{Z}$ as a C^*-tensor category with generator z, and denote the fundamental representation of $A_o(F)$ by v. Write π for the nontrivial irreducible summand of $v \otimes v$. We can embed $\operatorname{Rep}(A_u(F))$ into the free product $\mathbb{Z} * \operatorname{Rep}(A_o(F))$ as a full subcategory, by sending the fundamental representation u to zv; see [Banica 1997, Théorème 1(iv)]. Under this identification, we have $\bar{u} \otimes u = v \otimes v$, which implies that also $\bar{u}u = \pi$. By mapping u to vz instead, we similarly get $u \otimes \bar{u} = v \otimes v$.

So it remains to prove the proposition for $\operatorname{Rep}(A_o(F))$, where $F \in \mathrm{GL}_m(\mathbb{C})$ with $m \geq 2$ and $F\overline{F} = \pm 1$. If we choose $q \in [-1, 1] \setminus \{0\}$ such that

$$\operatorname{Tr}(F^*F) = |q| + |q|^{-1} \geq 2 \quad \text{and} \quad F\overline{F} = -\operatorname{sgn}(q) 1,$$

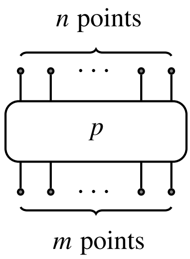
then it follows from [Bichon et al. 2006, Theorem 5.3] that $A_o(F)$ is monoidally equivalent to $\mathrm{SU}_q(2)$. We still denote the fundamental representation by v.

Note that, strictly speaking, the category $\mathrm{Rep}(\mathrm{SU}_q(2))$ depends on the sign of q. However, since we only work in the subcategory generated by $v \otimes v$, all parity issues disappear. More precisely, if v' denotes the fundamental representation of $\mathrm{SU}_{-q}(2)$, then the full C^*-tensor subcategories generated by $v \otimes v$ and $v' \otimes v'$ are monoidally equivalent. To see this, denote the Hopf $*$-subalgebra of $\mathrm{Pol}(\mathrm{SU}_q(2))$ generated by the matrix coefficients of $v \otimes v$ by B. It suffices to remark that in the same way as in [Banica 1999, Corollary 4.1], the adjoint coaction of $\mathrm{SU}_q(2)$ on $M_2(\mathbb{C})$ identifies B with the quantum automorphism group of $(M_2(\mathbb{C}), \phi_{q^2})$, where the state ϕ_{q^2} is given by

$$\phi_{q^2} : M_2(\mathbb{C}) \to \mathbb{C}, \qquad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto \frac{1}{1+q^2}(a_{11} + q^2 a_{22}).$$

In particular, the isomorphism class of B does not depend on the sign of q. By duality, the full C^*-tensor subcategory of $\mathrm{Rep}(\mathrm{SU}_q(2))$ generated by $v \otimes v$ is also independent of the sign of q. We may therefore assume that $q < 0$ without loss of generality. Since U_π is unitary, it suffices to compute $\tau(U_\pi^k)$ for all positive integers k.

In summary, we have reduced the problem to a question about the Temperley–Lieb–Jones category $\mathcal{TL}_{d,-1}$, where $d = |q| + |q|^{-1}$ (cf. [Neshveyev and Tuset 2013, Section 2.5]). This category admits a well-behaved diagram calculus; see, e.g., [Banica and Speicher 2009]. In this view, morphisms from $v^{\otimes n}$ to $v^{\otimes m}$ are given by linear combinations of noncrossing pair partitions $p \in NC_2(n, m)$, which we will represent by diagrams of the following form:



The composition pq of diagrams p and q, whenever meaningful, is defined by vertical concatenation, removing any loops that arise. The tensor product and adjoint operations are given by horizontal concatenation and reflection along the horizontal axis, respectively. We will denote the morphism in $(v^{\otimes m}, v^{\otimes n})$ associated to the partition $p \in NC_2(n, m)$ by T_p. One then has that

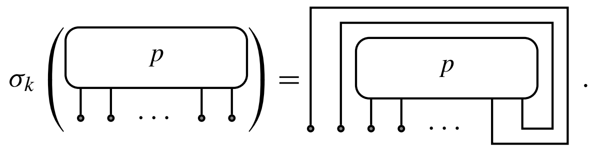
$$T_p = T_{p^*}, \qquad T_{p \otimes q} = T_p \otimes T_q \quad \text{and} \quad T_p T_q = d^{\ell(p,q)} T_{pq},$$

where $\ell(p, q)$ denotes the number of loops removed in the composition of p and q. Moreover, the family $\{T_p \mid p \in NC_2(n, m)\}$ is a basis for $(v^{\otimes m}, v^{\otimes n})$.

In view of Lemma 8.2, we now specialize to noncrossing pair diagrams without upper points, i.e., morphisms in $((v \otimes v)^{\otimes k}, \varepsilon)$. The action of $\zeta^k_{v\otimes v}$ (as defined in Lemma 8.2) on intertwiners of the form T_p for $p \in NC_2(0, 2k)$ has an easy description in terms of the partition calculus discussed above:

$$\zeta^k_{v\otimes v}(T_p) = T_{\sigma_k(p)},$$

where σ_k is the permutation of $NC_2(0, 2k)$ given by



In other words, $\zeta^k_{v\otimes v}$ permutes a basis of $((v \otimes v)^{\otimes k}, \varepsilon)$. In fact, ζ^k_π behaves similarly with respect to a suitable basis of $(\pi^{\otimes k}, \varepsilon)$. Let $Q : v \otimes v \to \pi$ be a coisometry. We proceed to argue that the intertwiners $\{Q^{\otimes k} T_p \mid p \in NC_2^\circ(k)\}$ form a basis of $(\pi^{\otimes k}, \varepsilon)$, where

$$NC_2^\circ(k) = \{p \in NC_2(0, 2k) \mid i \text{ odd} \Longrightarrow \{i, i+1\} \notin p\}.$$

Indeed, it is clear that multiplication by $Q^{\otimes k}$ yields a linear map from $((v\otimes v)^{\otimes k}, \varepsilon)$ to $(\pi^{\otimes k}, \varepsilon)$. Moreover, it is easy to see that T_p lies in the kernel of this map whenever $p \in NC_2(0, 2k) \setminus NC_2^\circ(k)$. Hence, to finish the proof of the claim, it suffices to check that

$$\dim_{\mathbb{C}}(\pi^{\otimes k}, \varepsilon) = \# NC_2^\circ(k). \tag{8-7}$$

This fact is probably well known, but we give a short proof here for completeness. The number of elements of $NC_2^\circ(k)$ is known in the combinatorial literature as the k-th Riordan number. As shown in [Bernhart 1999, Sections 3.2(R2) and 5], the Riordan numbers can be expressed in terms of the Catalan numbers C_i by means of the formula

$$\# NC_2^\circ(k) = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} C_i. \tag{8-8}$$

The left-hand side of (8-7) only depends on the fusion rules of the tensor powers of π, so we can take π to be the 3-dimensional irreducible representation of $SU(2)$ for the purposes of this part of the computation. Making use of the Weyl integration formula for $SU(2)$, see [Hall 2015, Example 11.33], we find that

$$\begin{aligned}\dim_{\mathbb{C}}(\pi^{\otimes k}, \varepsilon) &= \int_{SU(2)} \chi_\pi^k(g)\,\mathrm{d}g = \frac{4}{\pi}\int_0^{\pi/2} (4\cos^2\theta - 1)^k \sin^2\theta\,\mathrm{d}\theta \\ &= \frac{4}{\pi}\int_0^1 (4x^2-1)^k \sqrt{1-x^2}\,\mathrm{d}x = \frac{4}{\pi}\sum_{i=0}^{k} 4^i (-1)^{k-i}\binom{k}{i}\int_0^1 x^{2i}\sqrt{1-x^2}\,\mathrm{d}x.\end{aligned}$$

By the moment formula for the Wigner semicircle distribution, this is precisely (8-8).

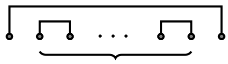
Having shown that the intertwiners of the form $Q^{\otimes k} T_p$ form a basis of $(\pi^{\otimes k}, \varepsilon)$, we now demonstrate that ζ^k_π acts on this basis by permutation. To this end, observe that $s_\pi = (Q \otimes Q) s_{v\otimes v}$. For $\xi \in ((v\otimes v)^{\otimes k}, \varepsilon)$, this yields

$$\begin{aligned}\zeta^k_\pi(Q^{\otimes k}\xi) &= Q^{\otimes k}(1^{\otimes 2k} \otimes s_\pi^*)(1 \otimes \xi \otimes 1)s_\pi \\ &= Q^{\otimes k}(1^{\otimes 2k} \otimes s_{v\otimes v}^*)(1^{\otimes 2k} \otimes Q^*Q \otimes Q^*Q)(1^{\otimes 2} \otimes \xi \otimes 1^{\otimes 2})s_{v\otimes v} = Q^{\otimes k}\zeta^k_{v\otimes v}(\xi),\end{aligned}$$

where the last equality follows by substituting $Q^*Q = 1 - d(v)^{-1}s_v s_v^*$ and noting that all terms involving $s_v s_v^*$ vanish. In summary,

$$\zeta_\pi^k(Q^{\otimes k}T_p) = Q^{\otimes k}T_{\sigma_k(p)}$$

for all $p \in NC_2^\circ(k)$. So ζ_π^k permutes a basis of $(\pi^{\otimes k}, \varepsilon)$, as claimed. It follows that the trace of ζ_π^k is exactly the number of fixed points of σ_k that lie in $NC_2^\circ(k)$. When $k = 1$, this set is empty, but for all $k \geq 2$ there is a unique such fixed point, given by the following partition:



$k-1$ pairs

Since

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \tau(U_\pi^k) = 1,$$

we conclude that $\tau(q_\pi) = 1$. Clearly, the measure on S^1 in the formulation of the proposition has the same moments as U_π and thus is the spectral measure of U_π. □

Remark 8.4. From the computation for $A_o(F)$, one might be tempted to conjecture that the trace of the spectral projection q_i is always less than 1 for all $i \in \mathrm{Irr}(\mathcal{C})$ in any C^*-tensor category. However, this is not the case. Consider the category of finite-dimensional unitary representations of the alternating group A_4. This category has four equivalence classes of irreducible objects, which we will denote by $\varepsilon, \omega_1, \omega_2$ and π. The trivial representation corresponds to ε, and ω_1, ω_2 are 1-dimensional representations that can be thought of as "cube roots of ε", in that $\omega_1 = \bar{\omega}_2$, and $\omega_1 \otimes \omega_1 = \omega_2$. The remaining representation π is 3-dimensional, and satisfies

$$\pi \otimes \pi \cong \varepsilon \oplus \omega_1 \oplus \omega_2 \oplus \pi \oplus \pi.$$

Fix a partition of the identity into pairwise orthogonal projections

$$1_{\pi\otimes\pi} = P_\varepsilon + P_{\omega_1\oplus\omega_2} + P_{\pi\oplus\pi}$$

such that the image of P_α is isomorphic to α. Using numerical methods, we found that

$$\begin{aligned} q_\pi = \tfrac{7}{18}p_\pi \oplus \tfrac{1}{18}1_\pi \oplus \tfrac{1}{18}1_\pi \oplus \big(\tfrac{7}{6}P_\varepsilon + \tfrac{1}{6}P_{\omega_1\oplus\omega_2} + \tfrac{1}{3}P_{\pi\oplus\pi}\big) \\ \in (\pi\varepsilon, \varepsilon\pi) \oplus (\pi\omega_1, \omega_1\pi) \oplus (\pi\omega_2, \omega_2\pi) \oplus (\pi\pi, \pi\pi) = p_\pi \cdot \mathcal{A} \cdot p_\pi. \end{aligned}$$

In particular,

$$\tau(q_\pi) = \tfrac{7}{18}d(\pi) = \tfrac{7}{6} > 1.$$

Acknowledgments

Vaes would like to thank the Isaac Newton Institute for Mathematical Sciences for support and hospitality during the programme *Operator Algebras: Subfactors and their Applications* when work on this paper was undertaken, supported by EPSRC Grant Number EP/K032208/1. Vaes also thanks Dimitri Shlyakhtenko for several helpful remarks.

References

[Arano 2016] Y. Arano, "Unitary spherical representations of Drinfeld doubles", *J. Reine Agnew. Math.* (online publication January 2016).

[Banica 1997] T. Banica, "Le groupe quantique compact libre U(n)", *Comm. Math. Phys.* **190**:1 (1997), 143–172. MR Zbl

[Banica 1999] T. Banica, "Symmetries of a generic coaction", *Math. Ann.* **314**:4 (1999), 763–780. MR Zbl

[Banica and Speicher 2009] T. Banica and R. Speicher, "Liberation of orthogonal Lie groups", *Adv. Math.* **222**:4 (2009), 1461–1501. MR Zbl

[Banica and Vergnioux 2009] T. Banica and R. Vergnioux, "Fusion rules for quantum reflection groups", *J. Noncommut. Geom.* **3**:3 (2009), 327–359. MR Zbl

[Bernhart 1999] F. R. Bernhart, "Catalan, Motzkin, and Riordan numbers", *Discrete Math.* **204**:1-3 (1999), 73–112. MR Zbl

[Bichon 2004] J. Bichon, "Free wreath product by the quantum permutation group", *Algebr. Represent. Theory* **7**:4 (2004), 343–362. MR Zbl

[Bichon 2013] J. Bichon, "Hochschild homology of Hopf algebras and free Yetter–Drinfeld resolutions of the counit", *Compos. Math.* **149**:4 (2013), 658–678. MR Zbl

[Bichon et al. 2006] J. Bichon, A. De Rijdt, and S. Vaes, "Ergodic coactions with large multiplicity and monoidal equivalence of quantum groups", *Comm. Math. Phys.* **262**:3 (2006), 703–728. MR Zbl

[Bichon et al. 2016a] J. Bichon, S. Neshveyev, and M. Yamashita, "Graded twisting of categories and quantum groups by group actions", *Ann. Inst. Fourier* (*Grenoble*) **66**:6 (2016), 2299–2338. MR Zbl

[Bichon et al. 2016b] J. Bichon, S. Neshveyev, and M. Yamashita, "Graded twisting of comodule algebras and module categories", preprint, 2016. arXiv

[Bichon et al. 2017] J. Bichon, D. Kyed, and S. Raum, "Higher ℓ^2-Betti numbers of universal quantum groups", *Canad. Math. Bull.* (online publication June 2017).

[Collins et al. 2009] B. Collins, J. Härtel, and A. Thom, "Homology of free quantum groups", *C. R. Math. Acad. Sci. Paris* **347**:5-6 (2009), 271–276. MR Zbl

[De Commer et al. 2014] K. De Commer, A. Freslon, and M. Yamashita, "CCAP for universal discrete quantum groups", *Comm. Math. Phys.* **331**:2 (2014), 677–701. MR Zbl

[Ghosh and Jones 2016] S. K. Ghosh and C. Jones, "Annular representation theory for rigid C^*-tensor categories", *J. Funct. Anal.* **270**:4 (2016), 1537–1584. MR Zbl

[Hall 2015] B. Hall, *Lie groups, Lie algebras, and representations: an elementary introduction*, 2nd ed., Graduate Texts in Mathematics **222**, Springer, 2015. MR Zbl

[Izumi 2000] M. Izumi, "The structure of sectors associated with Longo–Rehren inclusions, I: General theory", *Comm. Math. Phys.* **213**:1 (2000), 127–179. MR Zbl

[Kyed 2008] D. Kyed, "L^2-homology for compact quantum groups", *Math. Scand.* **103**:1 (2008), 111–129. MR Zbl

[Kyed and Raum 2017] D. Kyed and S. Raum, "On the ℓ^2-Betti numbers of universal quantum groups", *Math. Ann.* (online publication March 2017).

[Kyed et al. 2015] D. Kyed, H. D. Petersen, and S. Vaes, "L^2-Betti numbers of locally compact groups and their cross section equivalence relations", *Trans. Amer. Math. Soc.* **367**:7 (2015), 4917–4956. MR Zbl

[Lemeux 2014] F. Lemeux, "The fusion rules of some free wreath product quantum groups and applications", *J. Funct. Anal.* **267**:7 (2014), 2507–2550. MR Zbl

[Lück 2002] W. Lück, *L^2-invariants: theory and applications to geometry and K-theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete **44**, Springer, 2002. MR Zbl

[Müger 2003] M. Müger, "From subfactors to categories and topology, I: Frobenius algebras in and Morita equivalence of tensor categories", *J. Pure Appl. Algebra* **180**:1-2 (2003), 81–157. MR Zbl

[Neshveyev and Tuset 2013] S. Neshveyev and L. Tuset, *Compact quantum groups and their representation categories*, Cours Spécialisés **20**, Société Mathématique de France, Paris, 2013. MR Zbl

[Neshveyev and Yamashita 2015] S. Neshveyev and M. Yamashita, "A few remarks on the tube algebra of a monoidal category", preprint, 2015. To appear in *Proc. Edinb. Math. Soc.* arXiv

[Neshveyev and Yamashita 2016] S. Neshveyev and M. Yamashita, "Drinfeld center and representation theory for monoidal categories", *Comm. Math. Phys.* **345**:1 (2016), 385–434. MR Zbl

[Ocneanu 1994] A. Ocneanu, "Chirality for operator algebras", pp. 39–63 in *Subfactors* (Kyuzeso, 1993), edited by H. Araki et al., World Sci. Publ., River Edge, NJ, 1994. MR Zbl

[Petersen 2012] H. D. Petersen, *L^2-Betti numbers of locally compact groups*, Ph.D. thesis, University of Copenhagen, 2012, available at http://www.math.ku.dk/noter/filer/phd13hdp.pdf.

[Popa 1994] S. Popa, "Classification of amenable subfactors of type II", *Acta Math.* **172**:2 (1994), 163–255. MR Zbl

[Popa 1995] S. Popa, *Classification of subfactors and their endomorphisms*, CBMS Regional Conference Series in Mathematics **86**, American Math. Soc., Providence, RI, 1995. MR Zbl

[Popa and Vaes 2015] S. Popa and S. Vaes, "Representation theory for subfactors, λ-lattices and C^*-tensor categories", *Comm. Math. Phys.* **340**:3 (2015), 1239–1280. MR Zbl

[Popa et al. 2017] S. Popa, D. Shlyakhtenko, and S. Vaes, "Cohomology and L^2-Betti numbers for subfactors and quasi-regular inclusions", *Int. Math. Res. Not.* (online publication January 2017).

[Tarrago and Wahl 2016] P. Tarrago and J. Wahl, "Free wreath product quantum groups and standard invariants of subfactors", preprint, 2016. arXiv

[Van Daele and Wang 1996] A. Van Daele and S. Wang, "Universal quantum groups", *Internat. J. Math.* **7**:2 (1996), 255–263. MR Zbl

[Vergnioux 2012] R. Vergnioux, "Paths in quantum Cayley trees and L^2-cohomology", *Adv. Math.* **229**:5 (2012), 2686–2711. MR Zbl

[Weibel 1994] C. A. Weibel, *An introduction to homological algebra*, Cambridge Studies in Advanced Mathematics **38**, Cambridge University Press, 1994. MR Zbl

[Woronowicz 1998] S. L. Woronowicz, "Compact quantum groups", pp. 845–884 in *Symétries quantiques* (Les Houches, 1995), edited by A. Connes et al., North-Holland, Amsterdam, 1998. MR Zbl

Received 9 Feb 2017. Revised 3 May 2017. Accepted 11 Jun 2017.

DAVID KYED: dkyed@imada.sdu.dk
Department of Mathematics and Computer Science, University of Southern Denmark, DK-5230 Odense, Denmark

SVEN RAUM: sven.raum@epfl.ch
Sciences de Base, Section de Mathématiques, École polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

STEFAAN VAES: stefaan.vaes@kuleuven.be
Department of Mathematics, KU Leuven, 3001 Leuven, Belgium

MATTHIAS VALVEKENS: matthias.valvekens@kuleuven.be
Department of Mathematics, KU Leuven, 3001 Leuven, Belgium

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.org/apde.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in APDE are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

ANALYSIS & PDE

Volume 10 No. 7 2017