# LOW MACH NUMBER LIMIT OF THE FULL NAVIER-STOKES EQUATIONS

THOMAS ALAZARD

Abstract. The low Mach number limit for classical solutions of the full Navier-Stokes equations is here studied. The combined effects of large temperature variations and thermal conduction are taken into account. In particular, we consider general initial data. The equations lead to a singular problem whose linearized is not uniformly well-posed. Yet, it is proved that the solutions exist and are uniformly bounded for a time interval which is independent of the Mach number $\mathrm{Ma} \in (0,1]$, the Reynolds number $\mathrm{Re} \in [1,+\infty]$ and the Péclet number $\mathrm{Pe} \in [1,+\infty]$. Based on uniform estimates in Sobolev spaces, and using a Theorem of G. Métivier and S. Schochet [30], we next prove that the penalized terms converge strongly to zero. This allows us to rigorously justify, at least in the whole space case, the well-known computations given in the introduction of the P.-L. Lions' book [26].

## Contents



## 1. Introduction

There are five factors that dictate the nature of the low Mach number limit: the equations may be isentropic or non-isentropic; the fluid may be viscous or inviscid; the fluid may be an efficient or poor thermal conductor; the domain may be bounded or unbounded; the temperature variations may be small or large. Yet, there are only two cases in which the mathematical analysis of the low Mach number limit is well developed: first, in the isentropic regime [8, 9, 11, 12, 16, 27]; second, for inviscid and non heat-conductive fluids [1, 30, 32].

Our goal is to start a rigorous analysis of the general case in which the combined effects of large temperature variations and thermal conduction are taken into account. As first anticipated in [28], this case yields some new problems concerning the nonlinear coupling of the equations.

1.1. **Setting up the problem.** The full Navier-Stokes equations are:

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho v) = 0, \\ \partial_t(\rho v) + \operatorname{div}(\rho v \otimes v) + \nabla P = \operatorname{div} \tau, \\ \partial_t(\rho e) + \operatorname{div}(\rho v e) + P \operatorname{div} v = \operatorname{div}(k \nabla \mathcal{T}) + \tau \cdot Dv, \end{cases} \tag{1.1}$$

where $\rho$, $v = (v^1, \dots, v^d)$, $P$, $e$ and $\mathcal{T}$ denote the fluid density, velocity, pressure, energy and temperature, respectively. We consider Newtonian gases with Lamé viscosity coefficients $\zeta$ and $\eta$, so that the viscous strain tensor $\tau$ is given by

$$\tau := 2\zeta Dv + \eta \operatorname{div} v I_d,$$

where $2Dv = \nabla v + (\nabla v)^t$ and $I_d$ is the $d \times d$ identity matrix.

Considerable insight comes from being able to simplify the description of the governing equations (1.1) by introducing clever physical models and the use of judicious mathematical approximations. To reach this goal, a standard strategy is to introduce dimensionless numbers which determine the relative significance of competing physical processes taking place in moving fluids. Not only does this allow us to derive simplified equations of motion, but also to reveal the central feature of the phenomenon considered.

We distinguish three dimensionless parameters:

$$\varepsilon \in (0,1], \qquad \mu \in [0,1], \qquad \kappa \in [0,1].$$

The first parameter $\varepsilon$ is the Mach number, namely the ratio of a characteristic velocity in the flow to the sound speed in the fluid. The parameters $\mu$ and $\kappa$ are essentially the inverses of the Reynolds and Péclet numbers; they measure the importance of viscosity and heat-conduction.

To rescale the equations, there are basically two approaches which are available. The first is to cast equations in dimensionless form by scaling every variable by its characteristic value [28, 34].

The second is to consider one of the three changes of variables:

$$\begin{aligned} &t \to \varepsilon^2 t, && x \to \varepsilon x, && v \to \varepsilon v, && \zeta \to \mu\zeta, && \eta \to \mu\eta, && k \to \kappa k, \\ &t \to \varepsilon t, && x \to x, && v \to \varepsilon v, && \zeta \to \varepsilon\mu\zeta, && \eta \to \varepsilon\mu\eta, && k \to \varepsilon\kappa k, \\ &t \to t, && x \to x/\varepsilon, && v \to \varepsilon v, && \zeta \to \varepsilon^2\mu\zeta, && \eta \to \varepsilon^2\mu\eta, && k \to \varepsilon^2\kappa k. \end{aligned} \tag{1.2}$$

See [26, 44] for comments on the first two changes of variables. The third one is related to large-amplitude high-frequency solutions (these rapid variations are anomalous oscillations in the context of nonlinear geometric optics [5]).

In the end, these two approaches both yield the same result. The full Navier-Stokes equations, written in a non-dimensional way, are:

$$
\begin{cases}
\partial_t \rho + \operatorname{div}(\rho v) = 0, \\
\partial_t(\rho v) + \operatorname{div}(\rho v \otimes v) + \dfrac{\nabla P}{\varepsilon^2} = \mu \operatorname{div} \tau, \\
\partial_t(\rho e) + \operatorname{div}(\rho v e) + P \operatorname{div} v = \kappa \operatorname{div}(k \nabla \mathcal{T}) + \varepsilon^2 \mu \tau \cdot Dv.
\end{cases}
\tag{1.3}
$$

Our study is concerned with the analysis of the low Mach number limit for classical solutions of the full Navier Stokes equations (1.3) in the non-isentropic general case and for general initial data. In particular, the combined effects of large temperature variations and thermal conduction are accounted. We are interested in the limit $\varepsilon \to 0$ and in proving results that are independent of $\mu$ and $\kappa$. The analysis contains two parts. We first prove an existence and uniform boundedness result for a time interval independent of the parameters $\varepsilon$, $\mu$ and $\kappa$. We next study the behavior of the solutions when $\varepsilon$ tends to 0.

Many results have been obtained in the past two decades about the justification of the *incompressible limit*—which is a special case of the *low Mach number approximation*. Concerning the Euler equations ($\mu = \kappa = 0$), the study began in earlies eighties with works of Klainerman and Majda [22, 23], Schochet [36], Isozaki [18, 19], Ukai [42], and others. As regards the isentropic Navier-Stokes equations ($\mu = 1$, $\zeta$ and $\eta$ constants, $\rho = \rho(P)$), the mathematical analysis of the low Mach number limit has come of age since the pioneering works. Recent progress are presented in Danchin [8, 9], Desjardins and Grenier [11], Desjardins, Grenier, Lions and Masmoudi [12], Hoff [16] and Lions and Masmoudi [27]. They are also two very interesting earlier results concerning the group method: Grenier [15] and Schochet [37].

For the non-isentropic Euler equations and general initial data, Métivier and Schochet have recently proved a couple of theorems [30, 31, 32] that supersede a number of earlier results (a part of their study is extended in [1] to the boundary case). In particular they have proved the existence of classical solutions on a time interval independent of $\varepsilon$. The aim of this paper is precisely to start a rigorous analysis of the corresponding problems for the full Navier–Stokes equations.

The study of the low Mach number limit is a vast subject of which we have barely scratched the surface here. To fill in this gap we recommend Danchin [10], Desjardins and Lin [13], Gallagher [14] and Schochet [38] for well written survey papers. For the reader who wishes to learn more about the physics, Majda [28] and Zeytounian [44, 45] are good places to start.

Detailed historical accounts of the subject can be found in [34], along with a broad number of references for further reading. In connection to the stability analysis performed below, let us point out that the research of numerical algorithms valid for all flow speeds is a very active field [24, 35, 43].

1.2. **Uniform stability.** We consider classical solutions, that is, solutions valued in the Sobolev spaces $H^s(\mathbb{D})$ with $s$ large enough, where the domain $\mathbb{D}$ is either the whole space $\mathbb{R}^d$ or the torus $\mathbb{T}^d$. Our main result asserts that, for perfect gases, the classical solutions exist and are uniformly bounded for a time interval independent of $\varepsilon$, $\mu$ and $\kappa$. We mention that the case of general gases involves additional difficulties (see Remarks 1.4 and 1.6) and will be addressed in a separate paper.

We choose to work with the unknowns $P$, $v = (v^1, \dots, v^d)$ and $\mathcal{T}$. In order to be closed, the system must be augmented with two equations of state, prescribing the density $\rho$ and the energy $e$ as given functions of $P$ and $\mathcal{T}$. Here, we restrict ourselves to perfect gases. There exists two positive constants $R$ and $C_V$ such that

$$P = R\rho\mathcal{T} \quad \text{and} \quad e = C_V\mathcal{T}.$$

We begin by rewriting equations (1.3) in terms of $(P, v, \mathcal{T})$. Set $\gamma = 1 + R/C_V$. Performing linear algebra, we find that

$$\begin{cases} \partial_t P + v\cdot\nabla P + \gamma P \operatorname{div} v = (\gamma-1)\kappa \operatorname{div}(k\nabla\mathcal{T}) + (\gamma-1)\varepsilon\mathcal{Q}, \\ \rho(\partial_t v + v\cdot\nabla v) + \dfrac{\nabla P}{\varepsilon^2} = \mu \operatorname{div}\tau, \\ \rho C_V(\partial_t\mathcal{T} + v\cdot\nabla\mathcal{T}) + P\operatorname{div} v = \kappa\operatorname{div}(k\nabla\mathcal{T}) + \varepsilon\mathcal{Q}, \end{cases} \tag{1.4}$$

where $\rho = P/(R\mathcal{T})$ and $\mathcal{Q} := \varepsilon\mu\tau\cdot Dv$.

Equations (1.4) are supplemented with initial data:

$$P_{|t=0} = P_0, \quad v_{|t=0} = v_0 \quad \text{and} \quad \mathcal{T}_{|t=0} = \mathcal{T}_0. \tag{1.5}$$

Finally, it is assumed that $\zeta$, $\eta$ and the coefficient of thermal conductivity $k$ are $C^\infty$ functions of the temperature $\mathcal{T}$, satisfying

$$k(\mathcal{T}) > 0, \quad \zeta(\mathcal{T}) > 0 \quad \text{and} \quad \eta(\mathcal{T}) + 2\zeta(\mathcal{T}) > 0.$$

**Notation 1.1.** Hereafter, $A$ denotes the set of adimensioned parameters:

$$A := \left\{\, a = (\varepsilon,\mu,\kappa) \,|\, \varepsilon\in(0,1],\ \mu\in[0,1],\ \kappa\in[0,1] \,\right\}.$$

**Theorem 1.2.** *Let $d \geqslant 1$ and $\mathbb{D}$ denote either the whole space $\mathbb{R}^d$ or the torus $\mathbb{T}^d$. Consider an integer $s > 1 + d/2$. For all positive $\underline{P}$, $\underline{\mathcal{T}}$ and $M_0$, there is a positive time $T$ such that for all $a = (\varepsilon,\mu,\kappa)\in A$ and all initial data $(P_0^a, v_0^a, \mathcal{T}_0^a)$ such that $P_0^a$ and $\mathcal{T}_0^a$ take positive values and such that*

$$\varepsilon^{-1}\left\| P_0^a - \underline{P}\right\|_{H^{s+1}(\mathbb{D})} + \left\| v_0^a\right\|_{H^{s+1}(\mathbb{D})} + \left\|\mathcal{T}_0^a - \underline{\mathcal{T}}\right\|_{H^{s+1}(\mathbb{D})} \leqslant M_0,$$

*the Cauchy problem for (1.4)–(1.5) has a unique solution $(P^a, v^a, \mathcal{T}^a)$ such that $(P^a - \underline{P}, v^a, \mathcal{T}^a - \underline{\mathcal{T}}) \in C^0([0,T]; H^{s+1}(\mathbb{D}))$ and such that $P^a$ and $\mathcal{T}^a$ take positive values. In addition there exists a positive $M$, depending only on $M_0$, $\underline{P}$ and $\underline{\mathcal{T}}$, such that*

$$\sup_{a\in A}\sup_{t\in[0,T]}\left\{\varepsilon^{-1}\left\|P^a(t)-\underline{P}\right\|_{H^s(\mathbb{D})} + \left\|v^a(t)\right\|_{H^s(\mathbb{D})} + \left\|\mathcal{T}^a(t)-\underline{\mathcal{T}}\right\|_{H^s(\mathbb{D})}\right\} \leqslant M.$$

**Remark 1.3. i.** We will prove a more precise result which, in particular, exhibits some new smoothing effects for $\operatorname{div} v$ and $\nabla P$ (see Theorem 2.7).
**ii.** General initial data are here considered, and allow for large density and temperature variations. The hypothesis $P_0^a(x) - \underline{P} = O(\varepsilon)$ is the natural scaling to balance the acoustic components, see (1.6) and [9, 22, 28, 30].
**iii.** One technical reason why we are uniquely interested in the whole space $\mathbb{R}^d$ or the Torus $\mathbb{T}^d$ is that we will make use of the Fourier transform tools. A more serious obstacle is that, in the boundary case, there should be boundary layers to analyze [3]. For the Euler equations (that is, $\mu = \kappa = 0$), however, Theorem 1.2 remains valid in the boundary case [1].

Before leaving this paragraph, let us say some words about the difficulties involved in the proof of Theorem 1.2. The main obstacle is that the equations lead to a singular problem, depending on the small scaling parameter $\varepsilon$, whose linearized system is not uniformly well-posed in Sobolev spaces. In other words, since we consider large temperature variations, the problem is linearly (uniformly) unstable (see [30] for comments on this unstability). Therefore, we cannot obtain the nonlinear energy estimates by differentiating the equations nor by localizing in the frequency space by means of Littlewood-Paley operators. In particular, the technical aspects are different from those present in the previous studies of the Cauchy problem for strong solutions of (1.1). The latter problem has been widely studied, starting from [29] and culminating in [6, 7] which investigate global strong solutions in spaces invariant by the scaling (1.2) with $\mu = \kappa = 1$, following the approach initiated by Fujita and Kato.

More precisely, after applying the changes of variables given in §2.4, we are led to study a mixed hyperbolic/parabolic system of nonlinear equations of the form:

$$(1.6)\qquad \begin{cases} g_1(\theta, \varepsilon p)\big(\partial_t p + v\cdot\nabla p\big) + \dfrac{1}{\varepsilon}\operatorname{div} v - \dfrac{\kappa}{\varepsilon}\chi_1(\varepsilon p)\operatorname{div}\big(\beta(\theta)\nabla\theta\big) = \Upsilon_1, \\ g_2(\theta, \varepsilon p)\big(\partial_t v + v\cdot\nabla v\big) + \dfrac{1}{\varepsilon}\nabla p - \mu B_2(\theta, \varepsilon p) v = 0, \\ g_3(\theta, \varepsilon p)\big(\partial_t \theta + v\cdot\nabla\theta\big) + \operatorname{div} v - \kappa\chi_3(\varepsilon p)\operatorname{div}\big(\beta(\theta)\nabla\theta\big) = \varepsilon\Upsilon_3, \end{cases}$$

where $B_2$ is a second order elliptic differential operator, and $\Upsilon_i$ $(i = 1, 3)$ are of no consequence. An important feature of system (1.6) is that $g_1$ and $g_2$ depend on $\theta$. As a consequence $\nabla g_1$ and $\nabla g_2$ are of order $O(1)$. System (1.6) does not enter into the classical framework of a singular limit [15, 37] because of this strong coupling between the short time-scale and the fast time-scale. That is why we cannot derive estimates in Sobolev norms by standard methods using differentiation of the equations.

One of the main differences between the Euler equations [1, 30] (with $\mu = 0 = \kappa$) and the full equations is the following. When $\kappa = 0$, it is typically easy to obtain $L^2$ estimates uniform in $\varepsilon$ by a simple integration by parts in which the large terms in $1/\varepsilon$ cancel out. In sharp contrast (as observed in [28]), when $\kappa \neq 0$ and the initial temperature variations are large, the problem is more involved because the penalization operator is no longer skew-symmetric. Several difficulties also specifically arise for proving estimates that are independent of $\mu$ and $\kappa$.

**Remark 1.4.** For general equations of state, we are led to study systems having the form (1.6) where the coefficients $\chi_1$ and $\chi_3$ depend also on $\theta$. As a consequence, the singular operator is *nonlinear*.

1.3. **The low Mach number limit.** We now consider the limit of solutions of (1.3) in $\mathbb{R}^d$ as the Mach number $\varepsilon$ goes to 0. The purpose of the low Mach number approximation is to justify that the compression due to pressure variations can be neglected. This is a common assumption that is made when discussing the fluid dynamics of highly subsonic flows. In particular, provided that the sound propagation is adiabatic, it is the same as saying that the flow is incompressible. On sharp contrast, this is no longer true if the combined effect of large temperature variations and heat conduction is taken into account.

Going back to (1.4) we compute that, formally, the limit system reads

$$
\begin{cases} \gamma \underline{P} \operatorname{div} v = (\gamma - 1)\kappa \operatorname{div}(k\nabla \mathcal{T}), \\ \rho\big(\partial_t v + v\cdot\nabla v\big) + \nabla\pi = \mu \operatorname{div}\tau, \\ \rho C_P\big(\partial_t \mathcal{T} + v\cdot\nabla\mathcal{T}\big) = \kappa\operatorname{div}(k\nabla\mathcal{T}), \end{cases} \tag{1.7}
$$

where $\rho = \underline{P}/(R\mathcal{T})$, $C_P = \gamma C_V$ and, in keeping with the notations of Theorem 1.2, $\underline{P}$ denotes the constant value of the pressure at spatial infinity.

Our next result states that the solutions of the full equations (1.4) converge to the unique solution of (1.7) whose initial velocity is the incompressible part of the original velocity. Again, we consider general initial data which allow very large acceleration of order $O(\varepsilon^{-1})$.

**Theorem 1.5.** *Fix $\mu \in [0,1]$ and $\kappa \in [0,1]$. Assume that $(P^\varepsilon, v^\varepsilon, \mathcal{T}^\varepsilon)$ satisfy* (1.4) *and*

$$
\sup_{\varepsilon\in(0,1]}\sup_{t\in[0,T]} \|\varepsilon^{-1}(P^\varepsilon(t)-\underline{P})\|_{H^s} + \|v^\varepsilon(t)\|_{H^s} + \|\mathcal{T}^\varepsilon(t)-\underline{\mathcal{T}}\|_{H^s} < +\infty,
$$

*for some fixed time $T>0$, reference states $\underline{P},\underline{\mathcal{T}}$ and index $s$ large enough. Suppose in addition that the initial data $\mathcal{T}^\varepsilon_{|t=0} - \underline{\mathcal{T}}$ is compactly supported. Then, for all $s'<s$, the pressure variations $\varepsilon^{-1}(P^\varepsilon - \underline{P})$ converges strongly to 0 in $L^2(0,T;H^{s'}_{loc}(\mathbb{R}^d))$. Moreover, for all $s'<s$, $(v^\varepsilon,\mathcal{T}^\varepsilon)$ converges strongly in $L^2(0,T;H^{s'}_{loc}(\mathbb{R}^d))$ to a limit $(v,\mathcal{T})$ satisfying the system* (1.7).

Note that the convergence is not uniform in time for the oscillations on the acoustic time-scale prevent the convergence of the solutions on a small initial layer in time (see [18, 42]).

The key to proving this convergence result is to prove the decay to zero of the local energy of the acoustic waves. To do so we will consider general systems which include (1.6) as a special case. In particular, the analysis of the general systems considered below will apply for the study of the low

Mach number combustion as described in [28]. We mention that, in view of [1], it seems possible to consider the same problem for exterior domains (which is interesting for aeroacoustics [45]). Yet, we will not address this question. The results proved in [3, 32] indicate that the periodic case involves important additional phenomena.

We conclude this introduction with a remark concerning general gases.

**Remark 1.6.** For perfect gases, the limit constraint is linear in the following sense: it is of the form $\operatorname{div} v_e = 0$ with $v_e := v - C^{te}\kappa k \nabla \mathcal{T}$. In sharp contrast, for general gases the constraint is nonlinear. Indeed, it reads $\operatorname{div} v = f(\underline{P}, \mathcal{T})\kappa \operatorname{div}(k\nabla\mathcal{T})$. As a consequence, it is not immediate that, in this case, the corresponding Cauchy problem for (1.7) is well posed.

## 2. Main results

We will see in §2.4 below that it is possible to transform equations (1.4) into a system of the form

$$
\begin{cases}
g_1(\theta,\varepsilon p)\big(\partial_t p + v\cdot\nabla p\big) + \dfrac{1}{\varepsilon}\operatorname{div} v - \dfrac{\kappa}{\varepsilon}B_1(\theta,\varepsilon p)\theta = \Upsilon_1,\\
g_2(\theta,\varepsilon p)\big(\partial_t v + v\cdot\nabla v\big) + \dfrac{1}{\varepsilon}\nabla p - \mu B_2(\theta,\varepsilon p)v = 0,\\
g_3(\theta,\varepsilon p)\big(\partial_t \theta + v\cdot\nabla \theta\big) + \operatorname{div} v - \kappa B_3(\theta,\varepsilon p)\theta = \varepsilon\Upsilon_3,
\end{cases}
\tag{2.1}
$$

where the unknown $(p, v, \theta)$ is a function of the variables $(t, x) \in \mathbb{R}\times\mathbb{D}$ with values in $\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}$. Recall that $\mathbb{D}$ denotes either the whole space $\mathbb{R}^d$ or the torus $\mathbb{T}^d$. Moreover, the coefficients $g_i$, $i = 1, 2, 3$, are real-valued and the $B_i$'s are second-order differential operators given by:

$$
\begin{aligned}
&B_1(\theta,\varepsilon p) := \chi_1(\varepsilon p)\operatorname{div}(\beta(\theta)\nabla\cdot),\\
&B_2(\theta,\varepsilon p) := \chi_2(\varepsilon p)\operatorname{div}(2\zeta(\theta)D\cdot) + \chi_2(\varepsilon p)\nabla(\eta(\theta)\operatorname{div}\cdot), \quad 2D := \nabla + (\nabla\cdot)^t,\\
&B_3(\theta,\varepsilon p) := \chi_3(\varepsilon p)\operatorname{div}(\beta(\theta)\nabla\cdot).
\end{aligned}
$$

Finally, $\Upsilon_i := \chi_i(\varepsilon p)F(\theta, \sqrt{\mu}\nabla v)$ where $F \in C^\infty$ is such that $F(0) = 0$.

### 2.1. Structural assumptions.

**Assumption 2.1.** To avoid confusion, we denote by $(\vartheta, \wp) \in \mathbb{R}^2$ the place holder of the unknown $(\theta, \varepsilon p)$.

(A1) The $g_i$'s, $i = 1, 2, 3$, are $C^\infty$ positive functions of $(\vartheta, \wp) \in \mathbb{R}^2$.

(A2) The coefficients $\beta$ and $\zeta$ are $C^\infty$ positive functions of $\vartheta \in \mathbb{R}$, and $\eta$ is a $C^\infty$ function of $\vartheta \in \mathbb{R}$ satisfying $\eta(\vartheta) + 2\zeta(\vartheta) > 0$.

(A3) The coefficients $\chi_i$, $i = 1, 2, 3$, are $C^\infty$ positive functions of $\wp \in \mathbb{R}$. Moreover, for all $\wp \in \mathbb{R}$, there holds $\chi_1(\wp) < \chi_3(\wp)$.

**Assumption 2.2.** Use the notation $\mathsf{d}f = (\partial f/\partial\vartheta)\,\mathsf{d}\vartheta + (\partial f/\partial\wp)\,\mathsf{d}\wp$. There exists two $C^\infty$ diffeomorphisms $\mathbb{R}^2 \ni (\vartheta,\wp) \mapsto (S(\vartheta,\wp),\wp) \in \mathbb{R}^2$ and $\mathbb{R}^2 \ni (\vartheta,\wp) \mapsto (\vartheta, \varrho(\vartheta,\wp)) \in \mathbb{R}^2$ such that $S(0, 0) = \varrho(0, 0) = 0$ and

$$
\mathsf{d}S = g_3\,\mathsf{d}\vartheta - g_1\,\mathsf{d}\wp \quad \text{and} \quad \mathsf{d}\varrho = -\frac{\chi_1}{\chi_3}g_3\,\mathsf{d}\vartheta + g_1\,\mathsf{d}\wp.
\tag{2.2}
$$

*Brief discussion of the hypotheses.*

The main hypothesis in Assumption 2.1 is the inequality $\chi_1 < \chi_3$. It plays a crucial role in proving $L^2$ estimates (see Section 4). Moreover, given the assumption $\beta(\vartheta) > 0$, it ensures that the operator $B_1(\theta,0) - B_3(\theta,0)$ [which appears in the last equation of the limit system given below in (2.3)] is positive. This means nothing but the fact that the limit temperature evolves according to the standard equation of heat diffusion!

The identities given in (2.2) are compatibility conditions between the penalization operator and the viscous perturbation. The reason for introducing $S$ (or $\varrho$) will be clear in §6.5 (respectively §7.1).

2.2. **Uniform stability result.** Given $0 \leqslant t \leqslant T$, a normed space $X$ and a function $\varphi$ defined on $[0,T]\times\mathbb{D}$ with values in $X$, we denote by $\varphi(t)$ the function $\mathbb{D} \ni x \mapsto \varphi(t,x) \in X$. The usual Sobolev spaces are denoted $H^\sigma(\mathbb{D})$. Recall that, when $\mathbb{D} = \mathbb{R}^d$, they are equipped with the norms

$$\|u\|_{H^\sigma}^2 := (2\pi)^{-d}\int_{\mathbb{R}^d} \langle\xi\rangle^{2\sigma} |\widehat{u}(\xi)|^2\, d\xi,$$

where $\widehat{u}$ is the Fourier transform of $u$ and $\langle\xi\rangle := \big(1+|\xi|^2\big)^{1/2}$. With regards to the case $\mathbb{D} = \mathbb{T}^d$, we replace the integrals in $\xi \in \mathbb{R}^d$ by sums on $k \in \mathbb{Z}^d$.

Let us introduce a bit of notation which will be used continually in the rest of the paper.

**Notation 2.3.** Given $\sigma \in \mathbb{R}$ and $\varrho \geqslant 0$, set $\|u\|_{H^\sigma_\varrho} := \|u\|_{H^{\sigma-1}} + \varrho\,\|u\|_{H^\sigma}$.

Recall that $A := \{\,(\varepsilon,\mu,\kappa)\,|\,\varepsilon \in (0,1], \mu \in [0,1], \kappa \in [0,1]\,\}$. The norm that we will control is the following.

**Definition 2.4.** *Let $T > 0$, $\sigma \geqslant 0$, $a = (\varepsilon,\mu,\kappa) \in A$ and set $\nu := \sqrt{\mu+\kappa}$. The space $\mathcal{H}^\sigma_a(T)$ consists of these functions $(p,v,\theta)$ defined on $[0,T]\times\mathbb{D}$ with values in $\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}$ and such that $\|(p,v,\theta)\|_{\mathcal{H}^\sigma_a(T)} < +\infty$, where*

$$\|(p,v,\theta)\|_{\mathcal{H}^\sigma_a(T)} := \sup_{t\in[0,T]} \big\{\, \|(p(t),v(t))\|_{H^{\sigma+1}_{\varepsilon\nu}} + \|\theta(t)\|_{H^{\sigma+1}_{\nu}} \,\big\} +$$

$$\left(\int_0^T \mu\,\|\nabla v\|_{H^{\sigma+1}_{\varepsilon\nu}}^2 + \kappa\,\|\nabla\theta\|_{H^{\sigma+1}_{\nu}}^2 + \kappa\,\|\operatorname{div} v\|_{H^\sigma}^2 + (\mu+\kappa)\,\|\nabla p\|_{H^\sigma}^2\,dt\right)^{\frac{1}{2}}.$$

Similarly, the space $\mathcal{H}^\sigma_{a,0}$ consists of these functions $(p,v,\theta)$ defined on $\mathbb{D}$ with values in $\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}$ and such that $\|(p,v,\theta)\|_{\mathcal{H}^\sigma_{a,0}} < +\infty$, where

$$\|(p,v,\theta)\|_{\mathcal{H}^\sigma_{a,0}} := \|(p,v)\|_{H^{\sigma+1}_{\varepsilon\nu}} + \|\theta\|_{H^{\sigma+1}_{\nu}} \quad \text{with} \quad \nu := \sqrt{\mu+\kappa}.$$

**Remark 2.5.** At first we may not expect to have to take these norms into account, but they come up on their own in Section 4. Moreover, not only the norms, but also the spaces depend on $a$ (since we allow $\mu = 0$ or $\kappa = 0$).

**Notation 2.6.** Given a normed space $X$ and a nonnegative $M$, we denote by $B(X;M)$ the ball of center 0 and radius $M$ in $X$.

Here is our main result. In the context of Assumptions 2.1–2.2 we prove that the solutions of (2.1) exist and are uniformly bounded for a time interval which is independent of $a \in A$.

**Theorem 2.7.** *Suppose that system* (2.1) *satisfies Assumptions 2.1–2.2. Let $d \geqslant 1$. For all integer $s > 1 + d/2$ and for all positive $M_0$, there exists a positive $T$ and a positive $M$ such that for all $a \in A$ and all initial data in $B(\mathcal{H}^s_{a,0}; M_0)$, the Cauchy problem for* (2.1) *has a unique classical solution in $B(\mathcal{H}^s_a(T); M)$.*

The proof of this result will occupy us untill §7.2. The crucial part is to obtain estimates in Sobolev norms that are independent of $a \in A$. Notable technical aspects include the proof of an energy estimate for linearized equations [see Theorem 4.3] and the use of new tools to localize in the frequency space [see Propositions 3.2 and 3.4]. With these results in hands, we begin in §5 by analyzing the high frequency regime. The rest of the analysis is devoted to the proof of low frequency estimates. We mention that we do not need specific estimates for medium frequency components.

**Remark 2.8.** Up to numerous changes, a close inspection of the proof of Theorem 2.7 indicates that, in fact: for all $M > M_0 > 0$, there exists $T > 0$ such that for all $a \in A$ and all initial data in $B(\mathcal{H}^s_{a,0}; M_0)$ the Cauchy problem for (2.1) has a unique classical solution in $B(\mathcal{H}^s_a(T); M)$.

2.3. **Convergence toward the solution of the limit system.** We now consider the behavior of the solutions of (2.1) in $\mathbb{R}^d$ as the Mach number $\varepsilon$ tends to zero. Fix $\mu$ and $\kappa$ and consider a family of solutions of system (2.1), $(p^\varepsilon, v^\varepsilon, \theta^\varepsilon)$. It is assumed to be bounded in $C([0,T]; H^\sigma(\mathbb{R}^d))$ with $\sigma$ large enough and $T > 0$. Strong compactness of $\theta^\varepsilon$ is clear from uniform bounds for $\partial_t \theta^\varepsilon$. For the sequence $(p^\varepsilon, v^\varepsilon)$, however, the uniform bounds imply only weak compactness, insufficient to insure that the limits satisfy the limit equations. We remedy this by proving that the penalized terms converge strongly to zero.

**Theorem 2.9.** *Suppose that system* (2.1) *satisfies Assumption 2.1. Fix $\mu \in [0,1]$ and $\kappa \in [0,1]$, and let $d \geqslant 1$. Assume that $(p^\varepsilon, v^\varepsilon, \theta^\varepsilon)$ satisfy* (2.1) *and are uniformly bounded in $\mathcal{H}^s_{(\varepsilon,\mu,\kappa)}(T)$ for some fixed $T > 0$ and $s > 6 + d/2$. Suppose that the initial data $\theta^\varepsilon(0)$ converge in $H^s(\mathbb{R}^d)$ to a function $\theta_0$ decaying sufficiently rapidly at infinity in the sense that $\langle x \rangle^\delta \theta_0 \in H^s(\mathbb{R}^d)$ for some given $\delta > 2$, where $\langle x \rangle := (1 + |x|^2)^{1/2}$. Then, for all indices $s' < s$, $p^\varepsilon \to 0$ strongly in $L^2(0,T; H^{s'}_{loc}(\mathbb{R}^d))$ and $\operatorname{div} v^\varepsilon - \chi_1(\varepsilon p^\varepsilon) \operatorname{div}(\beta(\theta^\varepsilon) \nabla \theta^\varepsilon) \to 0$ strongly in $L^2(0,T; H^{s'-1}_{loc}(\mathbb{R}^d))$.*

The proof is given in §8. It is based on a Theorem of Métivier and Schochet recalled below in Theorem 8.3, about the decay to zero of the local energy for a class of wave operators with time dependent coefficients.

We have just seen that $p^\varepsilon$ converges to 0. The following result states that $(v^\varepsilon, \theta^\varepsilon)$ converges toward the solution of the limit system.

**Theorem 2.10.** *Using the same assumptions and notations as in Theorem 2.9, the family* $\{(v^\varepsilon,\theta^\varepsilon)\,|\,\varepsilon\in(0,1]\}$ *converges weakly in* $L^\infty(0,T;H^s(\mathbb{R}^d))$ *and strongly in* $L^2(0,T;H^{s'}_{loc}(\mathbb{R}^d))$ *for all* $s'<s$ *to a limit* $(v,\theta)$ *satisfying*

$$
(2.3)\qquad \begin{cases} \operatorname{div} v = \kappa\chi_1(0)\operatorname{div}(\beta(\theta)\nabla\theta),\\ g_2(\theta,0)\big(\partial_t v + v\cdot\nabla v\big) + \nabla\pi - \mu B_2(\theta,0)v = 0,\\ g_3(\theta,0)\big(\partial_t\theta + v\cdot\nabla\theta\big) - \kappa(\chi_3(0)-\chi_1(0))\operatorname{div}(\beta(\theta)\nabla\theta) = 0, \end{cases}
$$

*for some* $\pi$ *which can be chosen such that* $\nabla\pi\in C^0([0,T];H^{s-1}(\mathbb{R}^d))$.

Given Theorem 2.9, the proof of Theorem 2.10 follows from a close inspection of the proof of Theorem 1.5 in [30], and so will be omitted.

2.4. **Changes of variables.** Let us rewrite equations (1.3) in terms of the pressure fluctuations $p$, velocity $v$ and temperature fluctuations $\theta$; where $p$ and $\theta$ are defined by

$$
(2.4)\qquad P = \underline{P}e^{\varepsilon p},\quad \mathcal{T} = \underline{\mathcal{T}}e^{\theta}\quad \text{or}\quad p := \frac{1}{\varepsilon}\log\Big(\frac{P}{\underline{P}}\Big),\quad \theta := \log\Big(\frac{\mathcal{T}}{\underline{\mathcal{T}}}\Big),
$$

where $\underline{P}$ and $\underline{\mathcal{T}}$ are given by the statement of Theorem 1.2.

We can convert the pressure and temperature evolution equations into evolution equations for the fluctuations $p$ and $\theta$. Starting from (1.4), it is accomplished most readily by logarithmic differentiation ($\partial_{t,x}P = \varepsilon P\partial_{t,x}p$ and $\partial_{t,x}\mathcal{T} = \mathcal{T}\partial_{t,x}\theta$). By doing so we find that $(p,v,\theta)$ satisfies

$$
(2.5)\qquad \begin{cases} P\big(\partial_t p + v\cdot\nabla p\big) + \dfrac{\gamma P}{\varepsilon}\operatorname{div} v - \dfrac{(\gamma-1)\kappa}{\varepsilon}\operatorname{div}(k\mathcal{T}\nabla\theta) = (\gamma-1)\mathcal{Q},\\ \dfrac{P}{R\mathcal{T}}\big(\partial_t v + v\cdot\nabla v\big) + \dfrac{P}{\varepsilon}\nabla p - \mu\operatorname{div}(2\zeta Dv) - \mu\nabla(\eta\operatorname{div} v) = 0,\\ \dfrac{C_V P}{R}\big(\partial_t\theta + v\cdot\nabla\theta\big) + P\operatorname{div} v - \kappa\operatorname{div}(k\mathcal{T}\nabla\theta) = \varepsilon\mathcal{Q}. \end{cases}
$$

Therefore, the system (2.1) includes (2.5) as a special case where

$$
g_1^* := \frac{1}{\gamma},\ g_2^* := \frac{1}{R\mathcal{T}},\ g_3^* := \frac{C_V}{R},\ \chi_1^* := \frac{\gamma-1}{\gamma P},\ \chi_2^* = \chi_3^* := \frac{1}{P},
$$

where the $*$ indicates that the function is evaluated at $(\theta,\varepsilon p)$. Moreover, for $i=1,3$, $\Upsilon_i := \chi_i(\varepsilon p)F(\theta,\sqrt{\mu}\nabla v)$ where $F(\theta,\sqrt{\mu}\nabla v) := \mathcal{Q}$ is as in (1.4).

We easily verify that the Assumptions 2.1–2.2 are satisfied in this case. Hence, Theorem 1.2 as stated in the introduction is now a consequence of Theorem 2.7 since $P$ and $\mathcal{T}$ (given by (2.4)) are obviously positive functions and since we have $\|\cdot\|_{C^0([0,T];H^s(\mathbb{D}))}\leqslant\|\cdot\|_{\mathcal{H}^s_a(T)}$ (the details are left to the reader). Similarly, Theorem 1.5 follows from Theorem 2.9 and Theorem 2.10.

## 3. Localization in the frequency space

We now develop the analysis needed to localize in the frequency space. The first two paragraphs are a review consisting of various notations and usual results which serve as the requested background for what follows. The core of this section is §§3.3–3.4, in which we prove two technical ingredients needed to localize in the low frequency region. This will not be used before Section 6 and can be omitted before the reader gets there.

3.1. **Notations.** To fix matters, in this section we work on $\mathbb{R}^d$, yet all the results are valid *mutatis mutandis* in the Torus $\mathbb{T}^d$. All functions are assumed to be complex valued unless otherwise specified. The notation $\sigma_0$ always refers to a real number strictly greater than $d/2$, and $h$ stands for a small parameter. We use $K$ to denote a generic constant (independent of $h$). The notation $A \lesssim B$ means that $A \leqslant KB$ for such a constant $K$. Given two normed vector spaces $X_1$ and $X_2$, $\mathcal{L}(X_1, X_2)$ denotes the space of bounded operators from $X_1$ to $X_2$. We denote by $\|\cdot\|_{X_1 \to X_2}$ its norm, $\mathcal{L}(X)$ is a compact notation for $\mathcal{L}(X, X)$ and $I$ always refers to the identity operator. Finally, recall that

$$\langle \xi \rangle := \big(1 + |\xi|^2\big)^{1/2}.$$

3.2. **Preliminaries.** Let $m \in \mathbb{R}$. A function $q$ belongs to the class $S^m$ if $q(\xi)$ is a $C^\infty$ function of $\xi \in \mathbb{R}^d$ and satisfies $|\partial_\xi^\alpha q(\xi)| \leqslant C_\alpha \langle \xi \rangle^{m-|\alpha|}$, for all $\xi \in \mathbb{R}^d$ and for all multi-indices $\alpha \in \mathbb{N}^d$. Such a function is called a symbol. With the best constants $C_\alpha$ as semi-norms, $S^m$ is a Fréchet space. Given a symbol $q \in S^m$, the Fourier multiplier associated with $q$ is given by the operator $\mathcal{Q}$ acting on tempered distribution $u$ by $\widehat{\mathcal{Q}u} := q\widehat{u}$.

We now introduce the first of two families of operators which are used below to localize in the frequency space. Let $\jmath$ be a $C^\infty$ function of $\xi \in \mathbb{R}^d$, satisfying

$$0 \leqslant \jmath \leqslant 1, \quad \jmath(\xi) = 1 \text{ for } |\xi| \leqslant 1, \quad \jmath(\xi) = 0 \text{ for } |\xi| \geqslant 2, \quad \jmath(\xi) = \jmath(-\xi)\,.$$

Set $\jmath_h(\xi) = \jmath(h\xi)$, for $0 \leqslant h \leqslant 1$ and $\xi \in \mathbb{R}^d$; so that $\jmath_h$ is supported in the ball of radius $2/h$ about the origin. Then we define $J_h$ as the Fourier multiplier with symbol $\jmath_h$:

$$J_h = \jmath(hD_x).$$

Let us make a series of remarks on $J_h$. The operator $J_h$ is self-adjoint since $\jmath_h$ is a real-valued function. Using that $\jmath_h$ is even, we deduce that $J_h u$ is real-valued for any real-valued tempered distribution $u$. They are smoothing operators, and the family $\{\, J_h \mid 0 < h \leqslant 1 \,\}$ is an approximate identity.

We will often use the simple observation that for all $r \geqslant 0$, there exists a nonnegative constant $K$ so that for all $h \in (0, 1]$ and $\sigma \in \mathbb{R}$,

$$\|J_h\|_{H^\sigma \to H^{\sigma+r}} \leqslant \frac{K}{h^r} \quad \text{and} \quad \|I - J_h\|_{H^\sigma \to H^{\sigma-r}} \leqslant h^r. \tag{3.1}$$

One reason it is interesting to assume that $\jmath$ has compact support is the following:

$$J_h = J_h J_{ch} \qquad \text{for all } \; 0 \leqslant c \leqslant 2^{-1}. \tag{3.2}$$

In words, the Friedrichs mollifiers $J_h$ are interesting because they are essentially projection operators. Alternatively, it is also interesting to use a family of invertible smoothing operators. A good candidate is the family

$$\left\{ \Lambda_h^m := (I - h^2\Delta)^{m/2} \,|\, h \in (0,1] \right\}.$$

Heuristically, we expect that for nonnegative $m$,

$$J_h \Lambda_h^m \lesssim I, \quad (I - J_h)\Lambda_h^m \lesssim h^m |D_x|^m.$$

With regards to negative powers of $\Lambda_h$, we will use that

$$h^m \Lambda_h^{-m} \lesssim \Lambda^{-m}, \quad \Lambda_h^{-m} \lesssim I.$$

The previous statements are made precise by the following lemma.

**Lemma 3.1.** *Let $(m_1, m_2) \in \mathbb{R}^2$ be such that $0 \leqslant m_1 \leqslant m_2$. Then, for all $h \in [0,1]$ and all $\sigma \in \mathbb{R}$, we have*

$$\left\| h^{m_1} \Lambda_h^{-m_2} \right\|_{H^\sigma \to H^{\sigma + m_1}} \leqslant 1. \tag{3.3}$$

*For all $m \geqslant 0$ and all $c \in (0,1]$, there exists a positive constant $K$ such that for all $h \in [0,1]$, and all $\sigma \in \mathbb{R}$,*

$$\left\| J_{ch} \Lambda_h^m \right\|_{H^\sigma \to H^\sigma} \leqslant K, \tag{3.4}$$

$$\left\| (I - J_{ch}) \Lambda_h^m \right\|_{H^\sigma \to H^{\sigma - m}} \leqslant K h^m. \tag{3.5}$$

To prove these results, we check that, on the symbol level

$$0 \leqslant h^{m_1} \langle h\xi \rangle^{-m_2} \langle \xi \rangle^{m_1} \leqslant 1,$$
$$0 \leqslant \jmath(ch\xi) \langle h\xi \rangle^m \leqslant \langle 2c^{-1} \rangle^m,$$
$$0 \leqslant \big(1 - \jmath(ch\xi)\big) \langle h\xi \rangle^m \langle \xi \rangle^{-m} \leqslant h^m \langle c \rangle^2.$$

3.3. **A product estimate.** We now establish a product estimate, which says that the smoothing effect of the operators $\Lambda_h^{-m}$ is distributive.

**Proposition 3.2.** *Let $\sigma_0 > d/2$, $(\sigma_1, \sigma_2) \in \mathbb{R}_+^2$, $(m_1, m_2) \in \mathbb{R}_+^2$ be such that*

$$\sigma_1 + \sigma_2 + m_1 + m_2 \leqslant 2\sigma_0. \tag{3.6}$$

*There exists $K$ such that for all $h \in [0,1]$ and $u_i \in H^{\sigma_0 - \sigma_i - m_i}$,*

$$\left\| \Lambda_h^{-m_1 - m_2}(u_1 u_2) \right\|_{H^{\sigma_0 - \sigma_1 - \sigma_2}} \leqslant K \left\| \Lambda_h^{-m_1} u_1 \right\|_{H^{\sigma_0 - \sigma_1}} \left\| \Lambda_h^{-m_2} u_2 \right\|_{H^{\sigma_0 - \sigma_2}}.$$

*Proof.* The key point is that the operators $\Lambda_h^{-m}$ are invertible. This allows us to derive the desired product estimates from the corresponding results in the usual setting $m_1 = m_2 = 0$. To do so Proposition 3.2 is better formulated as follows: there exists $K$ such that for all $h \in [0,1]$ and $f_i \in H^{\sigma_0 - \sigma_i}(\mathbb{R}^d)$,

$$\left\| \Lambda_h^{-m_1 - m_2} \left\{ \big(\Lambda_h^{m_1} f_1\big)\big(\Lambda_h^{m_2} f_2\big) \right\} \right\|_{H^{\sigma_0 - \sigma_1 - \sigma_2}} \leqslant K \left\| f_1 \right\|_{H^{\sigma_0 - \sigma_1}} \left\| f_2 \right\|_{H^{\sigma_0 - \sigma_2}}. \tag{3.7}$$

The proof of this claim is based on the decomposition of each function $f_i$ into two pieces: its low wave number part $J_h f_i$, and its high wave number part $(I-J_h)f_i$. This leads to four products that are handled the same way by introducing, for $c\in\{0,1\}$ and $m\in\mathbb{R}$, the Fourier multipliers

$$\Theta^h_{c,m} := h^{-cm}\Lambda_h^m\left(J_h - cI\right).$$

With this notation, the left-hand side of (3.7) is less than

$$\sum_{0\leqslant c_1,c_2\leqslant 1}\left\| h^{c_1m_1+c_2m_2}\Lambda_h^{-m_1-m_2}\big\{\big(\Theta^h_{c_1,m_1}f_1\big)\big(\Theta^h_{c_2,m_2}f_2\big)\big\}\right\|_{H^{\sigma_0-\sigma_1-\sigma_2}}.$$

Hence, to prove (3.7) it suffices now to combine three ingredients:

$$\begin{aligned}
&\left\| h^{c_1m_1+c_2m_2}\Lambda_h^{-m_1-m_2}v\right\|_{H^{\sigma_0-\sigma_1-\sigma_2}} \leqslant \left\|v\right\|_{H^{\sigma_0-\sigma_1-\sigma_2-c_1m_1-c_2m_2}}\ ,\\
&\left\|v_1v_2\right\|_{H^{\sigma_0-\sigma_1-\sigma_2-c_1m_1-c_2m_2}} \lesssim \left\|v_1\right\|_{H^{\sigma_0-\sigma_1-c_1m_1}}\left\|v_2\right\|_{H^{\sigma_0-\sigma_2-c_2m_2}}\ ,\\
&\left\|\Theta^h_{c_i,m_i}v\right\|_{H^{\sigma_0-\sigma_i-c_im_i}} \lesssim \left\|v\right\|_{H^{\sigma_0-\sigma_i}} \quad (i\in\{1,2\})\,.
\end{aligned}$$

The first and last inequalities follow from Lemma 3.1. In order to prove the second one, we first recall the classical rule of product in Sobolev spaces (which is Theorem 8.3.1. in [17]). For $(r_1,r_2)\in\mathbb{R}^2$, the product maps continuously $H^{r_1}(\mathbb{R}^d)\times H^{r_2}(\mathbb{R}^d)$ to $H^r(\mathbb{R}^d)$ whenever

$$r_1+r_2\geqslant 0,\quad r\leqslant\min\{r_1,r_2\}\quad\text{and}\quad r\leqslant r_1+r_2-d/2, \tag{3.8}$$

with the third inequality strict if $r_1$ or $r_2$ or $-r$ is equal to $d/2$. We next verify that (3.8) applies with

$$r:=\sigma_0-\sigma_1-\sigma_2-c_1m_1-c_2m_2,\qquad r_i:=\sigma_0-\sigma_i-c_im_i\quad\text{for } i=1,2.$$

3.4. **A Friedrichs' Lemma.** In this paragraph we present a result which complements the usual Friedrichs' Lemma. To do that we first need a commutator estimate, noting that the commutator of a Fourier multiplier of order $m$ and the multiplication by a function is an operator of order $m-1$.

**Lemma 3.3.** *Let $\sigma_0>d/2+1$ and $m\in[0,+\infty)$. For any bounded subset $\mathcal{B}$ of $S^m$ and all $\sigma\in(-\sigma_0+m,\sigma_0-1]$, there exists a constant $K$ such that for all symbol $q\in\mathcal{B}$, all $f\in H^{\sigma_0}(\mathbb{R}^d)$, and all $u\in H^\sigma(\mathbb{R}^d)$,*

$$\left\|\mathcal{Q}(fu)-f\mathcal{Q}u\right\|_{H^{\sigma-m+1}}\leqslant K\left\|f\right\|_{H^{\sigma_0}}\left\|u\right\|_{H^\sigma}, \tag{3.9}$$

*where $\mathcal{Q}$ is the Fourier multiplier with symbol $q$.*

Lemma 3.3 is classical. Yet, for the convenience of the reader, we include a proof at the end of this paragraph.

A word of caution: the estimate (3.9) carries over to matrix-valued functions and symbols except for one key point. Suppose that $f$ and $q$ are matrix valued. In order for (3.9) to be true the following condition must be fulfilled: $q(\xi)f(x)=f(x)q(\xi)$ for all $(x,\xi)\in\mathbb{R}^{2d}$.

**Proposition 3.4.** *Let $\sigma_0>d/2+1$ and $m\in[0,1]$. For all $\sigma$ in the interval $(-\sigma_0+m,\sigma_0-1]$, there exists a constant $K$, such that for all $h\in(0,1]$, all $f\in H^{\sigma_0}(\mathbb{R}^d)$ and all $u\in H^{-\sigma_0}(\mathbb{R}^d)$,*

$$\left\|J_h(fu)-fJ_hu\right\|_{H^{\sigma-m+1}}\leqslant h^mK\left\|f\right\|_{H^{\sigma_0}}\left\|\Lambda_h^{-(\sigma_0+\sigma)}u\right\|_{H^\sigma}. \tag{3.10}$$

**Remark 3.5.** The thing of interest here is that the precise rate of convergence does not require much on the high wave number part of $u$.

*Proof.* Given $\mu \in \mathbb{R}$ and $g \in H^{\mu}(\mathbb{R}^d)$, we denote by $g^{\flat}$ the multiplication operator $H^{-\mu}(\mathbb{R}^d) \ni u \mapsto gu \in \mathcal{S}'(\mathbb{R}^d)$. Then, Proposition 3.4 can be formulated concisely in the following way. There exists a constant $K$, depending only on $\sigma_0$, $m$ and $\sigma$, such that for all $h \in (0,1]$,

$$\left\| \Lambda^{1-m} \big[ J_h, f^{\flat} \big] \Lambda_h^{\sigma_0+\sigma} \right\|_{H^\sigma \to H^\sigma} \leqslant h^m K \left\| f \right\|_{H^{\sigma_0}} .$$

The proof of this claim makes use of the division of the frequency space into two pieces: the low frequencies region $|\xi| \lesssim 1/h$ and the high frequencies region $|\xi| \gtrsim 1/h$. We write

$$\Lambda^{1-m} \big[ J_h, f^{\flat} \big] \Lambda_h^{\sigma_0+\sigma} = C_0^h + C_\infty^h,$$

where

$$\begin{aligned} C_0^h &:= \Lambda^{1-m} \big[ J_h, f^{\flat} \big] J_h' \Lambda_h^{\sigma_0+\sigma}, \\ C_\infty^h &:= \Lambda^{1-m} \big[ J_h, f^{\flat} \big] (I - J_h') \Lambda_h^{\sigma_0+\sigma}, \end{aligned}$$

and $J_h' = J_{h/5}$.

Firstly, we estimate $C_0^h$. We rewrite $C_0^h$ as

$$C_0^h = h^m \Lambda^{1-m} \big[ h^{-m}(J_h - I), f^{\flat} \big] J_h' \Lambda_h^{\sigma_0+\sigma}. \tag{3.11}$$

The key point to estimate $C_0^h$ is to notice that $\{ h^{-m}(\jmath_h - 1) \mid 0 < h \leqslant 1 \}$ is a bounded family in $S^m$. Indeed, if $q \in S^m$ and $q(\xi)$ vanishes for small $\xi$, then the symbols $q_\lambda(\cdot) = \lambda^{-m} q(\lambda \cdot)$ belong uniformly to $S^m$, for $0 < \lambda < \infty$. Once this is granted, Lemma 3.3 implies that the family

$$\Big\{ \Lambda^{1-m} \big[ h^{-m}(J_h - I), f^{\flat} \big] \mid 0 < h \leqslant 1 \Big\}$$

is bounded in $\mathcal{L}(H^\sigma)$.

On the other hand $\{ J_h' \Lambda_h^{\sigma_0+\sigma} \mid 0 < h \leqslant 1 \}$ is a bounded family in $\mathcal{L}(H^\sigma)$, see (3.4). Consequently, in light of (3.11), we end up with

$$\begin{aligned} \left\| C_0^h \right\|_{H^\sigma \to H^\sigma} &\leqslant h^m \left\| \Lambda^{1-m} \big[ h^{-m}(J_h - I), f^{\flat} \big] \right\|_{H^\sigma \to H^\sigma} \left\| J_h' \Lambda_h^{\sigma_0+\sigma} \right\|_{H^\sigma \to H^\sigma} \\ &\lesssim h^m \left\| f \right\|_{H^{\sigma_0}} . \end{aligned}$$

Our next task is to show similar estimates for $C_\infty^h$. With regards to $C_\infty^h$, the fact that the operators $J_h$ are essentially projection operators is the key to the proof. More precisely, we use the identity (3.2) written in the form $J_h(1 - J_h') = 0$ [recall that $J_h' = J_{h/5}$]. It yields

$$C_\infty^h = \Lambda^{1-m} J_h f^{\flat} \left( I - J_h' \right) \Lambda_h^{\sigma_0+\sigma}. \tag{3.12}$$

It turns out that the situation is even better. Let $v$ belongs to the Schwartz class $\mathcal{S}$. The spectrum (support of the Fourier transform, hereafter denoted by spec ) of $(J_h f)\big((I - J_h')v\big)$ is contained in

$$\operatorname{spec}(J_h f) + \operatorname{spec}(I - J_h')v \subset B\left(0, 2h^{-1}\right) + B\left(0, 10h^{-1}\right)^c \subset B\left(0, 3h^{-1}\right)^c,$$

the exterior of the ball centered at 0 of radius $3h^{-1}$. This results in

$$J_h\big((J_h f)((I-J_h')v)\big)=0,$$

that is, $J_h\big(f((I-J_h')v)\big)=J_h\big((I-J_h)f)((I-J_h')v)\big)$. By combining this identity with (3.12), we are left with

$$C_\infty^h=\Lambda^{1-m}J_h\big((I-J_h)f\big)^\flat\,\big(I-J_h'\big)\,\Lambda_h^{\sigma_0+\sigma}.$$

To estimate $\left\|C_\infty^h\right\|_{H^\sigma\to H^\sigma}$ we prove a dual estimate for the operator adjoint. We write $\big(C_\infty^h\big)^\star$ as a product of two operators:

$$\big(C_\infty^h\big)^\star=\Big\{h^{-\sigma_0-\sigma}(I-J_h')\Lambda_h^{\sigma_0+\sigma}\Big\}\Big\{\big(\overline{(I-J_h)f}\big)^\flat\,\big(h^{\sigma_0+\sigma}J_h\big)\,\Lambda^{1-m}\Big\},$$

where $\bar z$ denotes the complex conjugated of $z$. The problem reduces to establishing that

$$\left\|\big(\overline{(I-J_h)f}\big)^\flat\big(h^{\sigma_0+\sigma}J_h\Lambda^{1-m}\big)\right\|_{H^{-\sigma}\to H^{\sigma_0}}\lesssim h^m\left\|f\right\|_{H^{\sigma_0}}. \tag{3.13}$$

Indeed, since $\{\, h^{-\sigma_0-\sigma}(I-J_h')\Lambda_h^{\sigma_0+\sigma}\mid 0<h\leqslant 1\,\}$ is bounded in $\mathcal{L}(H^{\sigma_0};H^{-\sigma})$ (see (3.5)), the estimate (3.13) implies that

$$\left\|\big(C_\infty^h\big)^\star\right\|_{H^{-\sigma}\to H^{-\sigma}}\lesssim h^m\left\|f\right\|_{H^{\sigma_0}}.$$

Which in turn implies $\left\|C_\infty^h\right\|_{H^\sigma\to H^\sigma}\lesssim h^m\left\|f\right\|_{H^{\sigma_0}}$ and completes the proof.

We now have to prove (3.13). Let $v$ belongs to the Schwartz class $\mathcal{S}$. Since $\sigma_0>d/2$, we can invoke the standard tame estimate for products, which leads to

$$\begin{aligned}&\left\|\overline{(I-J_h)f}\big(h^{\sigma_0+\sigma}J_h\Lambda^{1-m}v\big)\right\|_{H^{\sigma_0}}\lesssim\\ &\qquad\left\|(I-J_h)f\right\|_{L^\infty}\left\|\big(h^{\sigma_0+\sigma}J_h\Lambda^{1-m}\big)\,v\right\|_{H^{\sigma_0}}+\\ &\qquad\left\|(I-J_h)f\right\|_{H^{\sigma_0}}\left\|\big(h^{\sigma_0+\sigma}J_h\Lambda^{1-m}\big)\,v\right\|_{L^\infty}.\end{aligned} \tag{3.14}$$

To estimate these fours terms, we use the embedding of $H^{\sigma_0-1}(\mathbb{R}^d)$ into $L^\infty(\mathbb{R}^d)$ and the bounds given in (3.1), to obtain

$$\begin{aligned}\left\|(I-J_h)f\right\|_{L^\infty}&\lesssim\left\|(I-J_h)f\right\|_{H^{\sigma_0-1}}\lesssim h\left\|f\right\|_{H^{\sigma_0}},\\ \left\|\big(h^{\sigma_0+\sigma}J_h\Lambda^{1-m}\big)\,v\right\|_{H^{\sigma_0}}&=h^{m-1}\left\|\big(h^{\sigma_0+\sigma+1-m}J_h\big)\,v\right\|_{H^{\sigma_0+1-m}}\\ &\lesssim h^{m-1}\|v\|_{H^{-\sigma}},\\ \left\|(I-J_h)f\right\|_{H^{\sigma_0}}&\leqslant\left\|f\right\|_{H^{\sigma_0}},\\ \left\|\big(h^{\sigma_0+\sigma}J_h\Lambda^{1-m}\big)\,v\right\|_{L^\infty}&\lesssim\left\|\big(h^{\sigma_0+\sigma}J_h\Lambda^{1-m}\big)\,v\right\|_{H^{\sigma_0-1}}\\ &=h^m\left\|\big(h^{\sigma_0+\sigma-m}J_h\big)\,v\right\|_{H^{\sigma_0-m}}\\ &\lesssim h^m\left\|v\right\|_{H^{\sigma_0-m-(\sigma_0+\sigma-m)}}=h^m\left\|v\right\|_{H^{-\sigma}}.\end{aligned}$$

Inserting the four estimates we just proved in (3.14), we conclude that

$$\left\|\overline{(I-J_h)f}\big(h^{\sigma_0+\sigma}J_h\Lambda^{1-m}v\big)\right\|_{H^{\sigma_0}}\lesssim h^m\left\|f\right\|_{H^{\sigma_0}}\left\|v\right\|_{H^{-\sigma}}.$$

This completes the proof of (3.13). □

*Proof of Lemma 3.3.* To avoid trivialities ($\sigma \in \emptyset$), assume that $m \leqslant 2\sigma_0 - 1$. We establish the estimate by using the para-differential calculus of Bony [2]. In keeping with the notations of the previous proof, $f^\flat$ denotes the multiplication operator $u \mapsto fu$. We denote by $T_f$ the operator of para-multiplication by $f$. Rewrite the commutator $\left[\mathcal{Q}, f^\flat\right]$ as

$$[\mathcal{Q}, T_f] + \mathcal{Q}(f^\flat - T_f) - (f^\flat - T_f)\mathcal{Q}.$$

The claim then follows from the bounds

$$\forall \sigma \in \mathbb{R}, \qquad \left\| [\mathcal{Q}, T_f] \right\|_{H^\sigma \to H^{\sigma - m + 1}} \leqslant c_1(q, \sigma) \left\| f \right\|_{H^{\sigma_0}}, \tag{3.15}$$

$$\forall \sigma \in (-\sigma_0, \sigma_0 - 1], \quad \left\| f^\flat - T_f \right\|_{H^\sigma \to H^{\sigma+1}} \leqslant c_2(\sigma) \left\| f \right\|_{H^{\sigma_0}}, \tag{3.16}$$

$$\forall \sigma \in \mathbb{R}, \qquad \left\| \mathcal{Q} \right\|_{H^\sigma \to H^{\sigma - m}} \lesssim \sup_\xi \left| \langle \xi \rangle^{-m} q(\xi) \right|, \tag{3.17}$$

where $c_1(\cdot, \sigma) \colon S^m \to \mathbb{R}_+$ maps bounded sets to bounded sets. We refer the reader to [2] and [33] for the proofs of the first two inequalities (see also [17, Prop. 10.2.2] and [17, Th. 9.6.4′] for the proof of (3.15); a detailed proof of (3.16) is given in [17, Prop. 10.2.9]). The estimate (3.17) is obvious. □

## 4. Energy estimates for the linearized system

Many results have been obtained in the past two decades concerning the symmetrization of the Navier Stokes equations (see, e.g., [4, 7, 20, 21]). Yet, the previous works do not include the dimensionless numbers. Here we prove estimates valid for all $a = (\varepsilon, \mu, \kappa)$ in $A$, where $A$ is defined in Notation 1.1. As already written in the introduction, our result improves earlier works [1, 22, 30] on allowing $\kappa \neq 0$. Indeed, when $\kappa = 0$, the penalization operator is skew-symmetric and hence the perturbation terms do not appear in the $L^2$ estimate, so that the classical proof for solutions to the unperturbed equations holds. In sharp contrast (as observed in [28]), when $\kappa \neq 0$ and the initial temperature variations are large, the problem is more involved because the singular operator is no longer skew-symmetric. Several difficulties also specifically arise for the purpose of proving estimates that are independent of $\mu$ and $\kappa$. In this regard we prove some additional smoothing effects for $\operatorname{div} v$ and $\nabla p$.

We consider the following linearized equations:

$$\begin{cases} g_1(\phi)\big(\partial_t \widetilde{p} + v \cdot \nabla \widetilde{p}\big) + \dfrac{1}{\varepsilon} \operatorname{div} \widetilde{v} - \dfrac{\kappa}{\varepsilon} \operatorname{div}(\beta_1(\phi) \nabla \widetilde{\theta}) = f_1, \\ g_2(\phi)\big(\partial_t \widetilde{v} + v \cdot \nabla \widetilde{v}\big) + \dfrac{1}{\varepsilon} \nabla \widetilde{p} - \mu \beta_2(\phi) \Delta \widetilde{v} - \mu \beta_2^\sharp(\phi) \nabla \operatorname{div} \widetilde{v} = f_2, \\ g_3(\phi)\big(\partial_t \widetilde{\theta} + v \cdot \nabla \widetilde{\theta}\big) + \operatorname{div} \widetilde{v} - \kappa \beta_3(\phi) \Delta \widetilde{\theta} = f_3. \end{cases} \tag{4.1}$$

To fix matters, the unknown $(\widetilde{p}, \widetilde{v}, \widetilde{\theta})$ is a function of the variables $(t, x) \in [0, T] \times \mathbb{D}$ ($T$ is a given positive real number and $\mathbb{D}$ denotes either $\mathbb{R}^d$ or $\mathbb{T}^d$) with values in $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$. The coefficients $\phi = \phi(t, x)$ and $v = v(t, x)$ take their values in $\mathbb{R}^N$ and $\mathbb{R}^d$, respectively ($N$ is a given integer).

**Assumption 4.1.** Parallel to Assumption 2.1, we make the following hypotheses. Throughout this section we require $g_1, g_2, g_3$ to be $C^\infty$ positive functions of $\phi \in \mathbb{R}^N$, without recalling this assumption explicitly in the statements. Similarly, it is assumed that $\beta_1$, $\beta_2$, $\beta_2^\sharp$ and $\beta_3$ are $C^\infty$ functions of $\phi \in \mathbb{R}^N$ satisfying

$$\beta_1 > 0, \quad \beta_2 > 0, \quad \beta_2^\sharp + \beta_2 > 0, \quad \beta_3 > 0.$$

Our main assumption reads

$$\beta_1 < \beta_3. \tag{4.2}$$

**Assumption 4.2.** For our purposes, it is sufficient to prove *a priori* estimates. It is always assumed that the unknown $\widetilde{U} := (\widetilde{p}, \widetilde{v}, \widetilde{\theta})$, the coefficients $(v, \phi)$ as well as the source term $f := (f_1, f_2, f_3)$ are in $C^0([0,T]; H^\infty(\mathbb{D}))$.

4.1. **Statements of the results.** We establish estimate on $\|\widetilde{U}\|_{\mathcal{H}^0_a(T)}$ in terms of the norm $\|\widetilde{U}(0)\|_{\mathcal{H}^0_{a,0}}$ of the data and norm of the source term $f$. Recall from Definition 2.4 that

$$\begin{aligned}
\|\widetilde{U}\|_{\mathcal{H}^0_a(T)} &:= \sup_{t\in[0,T]} \left\{ \|(\widetilde{p},\widetilde{v})(t)\|_{H^1_{\varepsilon\nu}} + \|\widetilde{\theta}(t)\|_{H^1_\nu} \right\} \\
&+ \left( \int_0^T \kappa \left\|\nabla\widetilde{\theta}\right\|^2_{H^1_\nu} + \mu \left\|\nabla\widetilde{v}\right\|^2_{H^1_{\varepsilon\nu}} + \kappa \left\|\operatorname{div}\widetilde{v}\right\|^2_{L^2} + (\mu+\kappa)\left\|\nabla\widetilde{p}\right\|^2_{L^2} \, dt \right)^{\frac{1}{2}},
\end{aligned}$$

with $\nu := \sqrt{\mu+\kappa}$. Recall the following notation we use continually in the sequel: given $\sigma \in \mathbb{R}$ and $\varrho \geqslant 0$ we set $\|\cdot\|_{H^\sigma_\varrho} := \|\cdot\|_{H^{\sigma-1}} + \varrho \|\cdot\|_{H^\sigma}$.

**Theorem 4.3.** *There is a smooth non-decreasing function $C$ from $[0,+\infty)$ to $[0,+\infty)$ such that for all $a \in A$, all $T \in (0,1]$, all coefficients $v$ and $\phi$, and all $(\widetilde{U}, f)$ satisfying* (4.1)*, the norm $\|\widetilde{U}\|_{\mathcal{H}^0_a(T)}$ satisfies the estimate*

$$\|\widetilde{U}\|_{\mathcal{H}^0_a(T)} \leqslant C(R_0)e^{TC(R)} \|\widetilde{U}(0)\|_{\mathcal{H}^0_{a,0}} + C(R)\int_0^T \|(f_1,f_2)\|_{H^1_{\varepsilon\nu}} + \|f_3\|_{H^1_\nu} \, dt,$$

*where $\nu := \sqrt{\mu+\kappa}$ and*

$$R_0 := \|\phi(0)\|_{L^\infty(\mathbb{D})}, R := \sup_{t\in[0,T]} \|(\phi, \partial_t\phi, \nabla\phi, \nu\nabla^2\phi, v, \nabla v)(t)\|_{L^\infty(\mathbb{D})}. \tag{4.3}$$

**Remark 4.4.** By Assumption 4.2, all the functions we encounter in the previous statement are in $C^0([0,T]; H^\infty(\mathbb{D}))$. Yet, one can verify that the estimate is valid whenever its two sides are well defined.

Theorem 4.3 is not enough for the purpose of proving high frequency estimates independent of $\kappa$. We will need the following version.

**Theorem 4.5.** *The statement of Theorem 4.3 remains valid if the system* (4.1) *is replaced with*

$$(4.4)\qquad \begin{cases} g_1(\phi)\big(\partial_t\widetilde{p} + v\cdot\nabla\widetilde{p}\big) + \dfrac{1}{\varepsilon}\operatorname{div}\widetilde{v} - \dfrac{\kappa}{\varepsilon}\operatorname{div}(\beta_1(\phi)\nabla\widetilde{\theta}) = f_1, \\ g_2(\phi)\big(\partial_t\widetilde{v} + v\cdot\nabla\widetilde{v}\big) + \dfrac{1}{\varepsilon}\nabla\widetilde{p} - \mu\beta_2(\phi)\Delta\widetilde{v} - \mu\beta_2^\sharp(\phi)\nabla\operatorname{div}\widetilde{v} = f_2, \\ g_3(\phi)\big(\partial_t\widetilde{\theta} + v\cdot\nabla\widetilde{\theta}\big) + G(\phi,\nabla\phi)\cdot\widetilde{v} + \operatorname{div}\widetilde{v} - \kappa\beta_3(\phi)\Delta\widetilde{\theta} = f_3, \end{cases}$$

*where $G$ is smooth in its arguments with values in $\mathbb{R}^d$.*

The proof of Theorem 4.3 relies upon two $L^2$ estimates. Because they may be useful in other circumstances, we give separate statements in §4.3. The proof of Theorem 4.5 follows from a close inspection of the proof of Theorem 4.3, and so will be omitted.

4.2. **Example.** Some important features of the proof of Theorem 4.3 can be revealed by analyzing the following simplified system:

$$(4.5)\qquad \begin{cases} \partial_t p + \dfrac{1}{\varepsilon}\operatorname{div} v - \dfrac{1}{\varepsilon}\Delta\theta = 0, \\ \partial_t v + \dfrac{1}{\varepsilon}\nabla p = 0, \\ \partial_t\theta + \operatorname{div} v - \beta\Delta\theta = 0. \end{cases}$$

For the sake of notational simplicity we abandon the tildes in this subsection. Parallel to (4.2), we suppose

$$(4.6)\qquad \beta > 1.$$

To symmetrize the large terms in $\varepsilon^{-1}$, we introduce $v_e := v - \nabla\theta$. This change of variables transforms (4.5) into

$$\begin{cases} \partial_t p + \dfrac{1}{\varepsilon}\operatorname{div} v_e = 0, \\ \partial_t v_e + \dfrac{1}{\varepsilon}\nabla p - \nabla\operatorname{div} v_e + (\beta-1)\nabla\Delta\theta = 0, \\ \partial_t\theta + \operatorname{div} v_e - (\beta-1)\Delta\theta = 0. \end{cases}$$

We take the $L^2$ scalar product of the first [resp. the second] equation with $p$ [resp. $v_e$]. It yields:

$$(4.7)\qquad \frac{1}{2}\frac{d}{dt}\left\|(p,v_e)\right\|_{L^2}^2 + \left\|\operatorname{div} v_e\right\|_{L^2}^2 - (\beta-1)\langle\, \Delta\theta\,,\operatorname{div} v_e\,\rangle = 0.$$

We take the $L^2$ scalar product of the third equation with $-\eta\Delta\theta$, where $\eta$ is a positive constant to be determined later on. It yields

$$\frac{1}{2}\frac{d}{dt}\left\|(p,v_e,\sqrt{\eta}\nabla\theta)\right\|_{L^2}^2 + \left\|\operatorname{div} v_e\right\|_{L^2}^2 - (\beta-1+\eta)\langle\, \Delta\theta\,,\operatorname{div} v_e\,\rangle + \eta(\beta-1)\left\|\Delta\theta\right\|_{L^2}^2 = 0.$$

Set $\eta := \beta - 1$, which is positive thanks to assumption (4.6). We get

$$\frac{1}{2}\frac{d}{dt}\left\|(p, v_e, \sqrt{\beta-1}\nabla\theta)\right\|_{L^2}^2 + \left\|\operatorname{div} v_e - (\beta-1)\Delta\theta\right\|_{L^2}^2 = 0.$$

Let $t \geqslant 0$. Integrating the previous identity from 0 to $t$, and using the triangle inequality to replace $v_e$ by $v$, we get

$$\left\|(p, v, \nabla\theta)(t)\right\|_{L^2}^2 + \int_0^t \left\|\operatorname{div} v - \beta\Delta\theta\right\|_{L^2}^2 \, d\tau \leqslant K_\beta \left\|(p, v, \nabla\theta)(0)\right\|_{L^2}^2 . \tag{4.8}$$

Here and below, $K_\beta$ denotes a generic constant which depends only on $\beta$.

We thus have proved an $L^2$ estimate independent of $\varepsilon$. To go beyond and obtain smoothing effect on $\operatorname{div} v$ it is sufficient to estimate $\Delta\theta$ independently. The strategy is to incorporate the troublesome term $\operatorname{div} v$ [in the equation for $\theta$] into a skew-symmetric operator.

To do so introduce

$$\zeta := \varepsilon\beta p - \theta \quad \text{and} \quad v_\varepsilon := \varepsilon v.$$

We compute

$$\begin{cases} \partial_t \zeta + \dfrac{\beta-1}{\varepsilon} \operatorname{div} v_\varepsilon = 0, \\ \partial_t v_\varepsilon + \dfrac{1}{\beta\varepsilon}\nabla\zeta + \dfrac{1}{\beta\varepsilon}\nabla\theta = 0, \\ \partial_t \theta + \dfrac{1}{\varepsilon}\operatorname{div} v_\varepsilon - \beta\Delta\theta = 0. \end{cases}$$

We now use the essential feature of the system, that is assumption (4.6). Multiply the first equation [resp. the second] by $1/(\beta-1)$ [resp. $\beta$], to put the penalization operator in symmetric form. The energy estimate thus reads

$$\frac{1}{2}\frac{d}{dt}\left\|(\sqrt{1/(\beta-1)}\zeta, \sqrt{\beta}v_\varepsilon, \theta)\right\|_{L^2}^2 + \beta\left\|\nabla\theta\right\|_{L^2}^2 = 0.$$

Integrate this inequality, to obtain

$$\left\|(\zeta, v_\varepsilon, \theta)(t)\right\|_{L^2}^2 + \int_0^t \left\|\nabla\theta\right\|_{L^2}^2 \, d\tau \leqslant K_\beta \left\|(\zeta, v_\varepsilon, \theta)(0)\right\|_{L^2}^2 . \tag{4.9}$$

Since the coefficients are constants, the same estimate holds true for the first order derivatives of $\zeta$, $v_\varepsilon$ and $\theta$. Namely, one has

$$\left\|\nabla(\zeta, v_\varepsilon, \theta)(t)\right\|_{L^2}^2 + \int_0^t \left\|\nabla^2\theta\right\|_{L^2}^2 \, d\tau \leqslant K_\beta \left\|\nabla(\zeta, v_\varepsilon, \theta)(0)\right\|_{L^2}^2 . \tag{4.10}$$

On applying the triangle inequality, one can replace $\zeta$ with $\varepsilon p$ in the previous estimates. Finally, by combining (4.8), (4.9) and (4.10), we get

$$\begin{aligned} &\left\|(p, v, \theta)(t)\right\|_{L^2} + \left\|\nabla(\theta, \varepsilon p, \varepsilon v)(t)\right\|_{L^2} + \left(\int_0^t \left\|\operatorname{div} v\right\|_{L^2}^2 + \left\|\nabla\theta\right\|_{H^1}^2 \, d\tau\right)^{1/2} \\ &\qquad \leqslant K_\beta \left\|(p, v, \theta)(0)\right\|_{L^2} + K_\beta \left\|\nabla(\theta, \varepsilon p, \varepsilon v)(0)\right\|_{L^2} . \end{aligned} \tag{4.11}$$

It is worth remarking that, to establish (4.11), it is enough to prove (4.7) and (4.10). Let us now compare (4.11) with the estimate given in Theorem 4.3. We see that the previous study fails to convey one feature of the proof, namely the usefulness of the additional smoothing effect for $\nabla p$. The reason is that we worked with constant coefficients. In the general case, an estimate for $\int_0^t \|\nabla p\|_{L^2}^2 \, d\tau$ is needed to control the left hand side of (4.10) (see Lemma 4.11 below).

Let us explain how to estimate $\int_0^t \|\nabla p\|_{L^2}^2 \, d\tau$. Multiply the second equation in (4.5) by $\varepsilon \nabla p$ and integrate over the strip $[0,t] \times \mathbb{D}$, to obtain

$$\int_0^t \|\nabla p\|_{L^2}^2 \, d\tau = -\int_0^t \langle\, \partial_t v \,,\, \varepsilon \nabla p \,\rangle \, d\tau.$$

Integrating by parts both in space and time yields

$$\begin{aligned}\int_0^t \langle\, \varepsilon \partial_t v \,,\, \nabla p \,\rangle \, d\tau &= -\int_0^t \langle\, v \,,\, \varepsilon \partial_t \nabla p \,\rangle \, d\tau + \varepsilon \big[ \langle\, v(\tau) \,,\, \nabla p(\tau) \,\rangle \big]_{\tau=0}^{\tau=t} \\ &= \int_0^t \langle\, \operatorname{div} v \,,\, \varepsilon \partial_t p \,\rangle \, d\tau + \varepsilon \big[ \langle\, v(\tau) \,,\, \nabla p(\tau) \,\rangle \big]_{\tau=0}^{\tau=t}.\end{aligned}$$

In view of the first equation in (4.5), we are left with

$$\int_0^t \|\nabla p\|_{L^2}^2 \, d\tau = \int_0^t \|\operatorname{div} v\|_{L^2}^2 - \langle\, \operatorname{div} v \,,\, \Delta \theta \,\rangle \, d\tau - \varepsilon \big[ \langle\, v(\tau) \,,\, \nabla p(\tau) \,\rangle \big]_{\tau=0}^{\tau=t}.$$

All the terms that appear in the previous identity have been estimated previously. As a consequence the estimate (4.11) holds true if we include $\int_0^t \|\nabla p\|_{L^2}^2 \, d\tau$ in its left hand side. By so doing, we obtain the exact analogue of the estimate given in Theorem 4.3 (for $\mu = 0$ and $\kappa = 1$).

4.3. $L^2$ **estimates.** Guided by the previous example, we want to prove $L^2$ estimates for $(\widetilde{p}, \widetilde{v})$ and $(\widetilde{\theta}, \varepsilon \widetilde{p}, \varepsilon \widetilde{v})$. The strategy for proving both estimates is the same: transform the system (4.1) so as to obtain $L^2$ estimates uniform in $\varepsilon$ by a simple integration by parts in which the large terms in $1/\varepsilon$ cancel out. Namely, we want to obtain systems having the form

$$\mathcal{L}_1(v,\phi)\mathcal{U} - \mathcal{L}_2(\mu,\kappa,\phi)\mathcal{U} + \frac{1}{\varepsilon} S(\phi)\mathcal{U} = F, \tag{4.12}$$

where $\mathcal{L}_1(v,\phi) - \mathcal{L}_2(\mu,\kappa,\phi)$ is a mixed hyperbolic/parabolic system of equations:

$$L_0(\phi)\partial_t + \sum_{1 \leqslant j \leqslant d} L_j(v,\phi)\partial_j - \sum_{1 \leqslant j,k \leqslant d} L_{jk}(\mu,\kappa,\phi)\partial_j \partial_k,$$

and the singular perturbation $S(\phi)$ is a differential operator in the space variable which is skew-symmetric (with not necessarily constant coefficients).

We first prove an estimate parallel to (4.9).

**Proposition 4.6.** *Using the same notations as in Theorem 4.3, we have*

$$
\begin{aligned}
&\sup_{t\in[0,T]} \|(\widetilde{\theta},\varepsilon\widetilde{p},\varepsilon\widetilde{v})(t)\|_{L^2} + \left(\int_0^T \kappa \|\nabla\widetilde{\theta}\|_{L^2}^2 + \mu \|\varepsilon\nabla\widetilde{v}\|_{L^2}^2 \, dt\right)^{1/2} \\
&\leqslant C(R_0)e^{TC(R)} \|(\widetilde{\theta},\varepsilon\widetilde{p},\varepsilon\widetilde{v})(0)\|_{L^2} + C(R)\left(\int_0^T B_\varepsilon(f,\widetilde{U})\, dt\right)^{1/2},
\end{aligned}
\tag{4.13}
$$

*where*

$$
B_\varepsilon(f,\widetilde{U}) := \|(f_3,\varepsilon f_1,\varepsilon f_2)\|_{L^2} \|(\widetilde{\theta},\varepsilon\widetilde{p},\varepsilon\widetilde{v})\|_{L^2}. \tag{4.14}
$$

**Corollary 4.7.** *The estimate* (4.13) *holds true with* $\left(\int_0^T B_\varepsilon(f,\widetilde{U})\, dt\right)^{1/2}$ *replaced by*

$$
\int_0^T \varepsilon \|(f_1,f_2)\|_{L^2} + \|f_3\|_{L^2} \, dt.
$$

*Proof of Corollary 4.7 given Proposition 4.6.* One has, for all $\lambda > 0$,

$$
\left(\int_0^T B_\varepsilon(f,\widetilde{U})\, dt\right)^{1/2} \leqslant \frac{1}{\lambda}\sup_{t\in[0,T]} \|(\widetilde{\theta},\varepsilon\widetilde{p},\varepsilon\widetilde{v})\|_{L^2} + \lambda\int_0^T \|(f_3,\varepsilon f_1,\varepsilon f_2)\|_{L^2}\, dt.
$$

Using (4.13) we obtain the desired result by taking $\lambda$ large enough. □

**Notation 4.8.** Within the proofs of Proposition 4.6 and 4.9: Rather than writing $C(R)$ and $C(R_0)$ in full, we will use the following abbreviations: $C = C(R)$ and $C_0 = C(R_0)$. So that $C$ denotes generic constants which depend only on $R := \left\|\left(\phi,\partial_t\phi,\nabla\phi,\nu\nabla^2\phi,v,\nabla v\right)\right\|_{L^\infty([0,T]\times\mathbb{D})}$ and $C_0$ denotes generic constants which depend only on $R_0 := \|\phi(0)\|_{L^\infty(\mathbb{D})}$. As usual, the values of $C_0$ and $C$ may vary from relation to relation.

*Proof of Proposition 4.6.* All the computations given below are meaningful since we concentrate on regular solutions of (4.1). Introduce

$$
\widetilde{\zeta} := \varepsilon g_1\beta_3\widetilde{p} - g_3\beta_1\widetilde{\theta}, \quad \widetilde{v}_\varepsilon := \varepsilon\widetilde{v} \quad \text{and} \quad \widetilde{\mathcal{U}} := (\widetilde{\zeta},\widetilde{v}_\varepsilon,\widetilde{\theta})^t.
$$

where we simply write $g_i$, $\beta_i$ instead of $g_i(\phi)$, $\beta_i(\phi)$. We first show that $\widetilde{\mathcal{U}}$ solves a system of the form (4.12). Expressing $\widetilde{p}$ in terms of $\widetilde{\zeta}$ and $\widetilde{\theta}$, replacing $\widetilde{v}$ by $\varepsilon^{-1}\widetilde{v}_\varepsilon$ and performing a little algebra, yields

$$
\begin{aligned}
&\partial_t\widetilde{\zeta} + v\cdot\nabla\widetilde{\zeta} + \frac{\beta_3-\beta_1}{\varepsilon}\operatorname{div}\widetilde{v}_\varepsilon = f_1', \\
&g_2\big(\partial_t\widetilde{v}_\varepsilon + v\cdot\nabla\widetilde{v}_\varepsilon\big) + \frac{1}{\varepsilon}\nabla\big(\gamma_1\widetilde{\zeta}\big) + \frac{1}{\varepsilon}\nabla\big(\gamma_2\widetilde{\theta}\big) - \mu\beta_2\Delta\widetilde{v}_\varepsilon - \mu\beta_2^\sharp\nabla\operatorname{div}\widetilde{v}_\varepsilon = \varepsilon f_2, \\
&g_3\big(\partial_t\widetilde{\theta} + v\cdot\nabla\widetilde{\theta}\big) + \frac{1}{\varepsilon}\operatorname{div}\widetilde{v}_\varepsilon - \kappa\beta_3\Delta\widetilde{\theta} = f_3,
\end{aligned}
$$

where $\gamma_1 := \dfrac{1}{g_1\beta_3}$, $\gamma_2 := \dfrac{g_3\beta_1}{g_1\beta_3}$ and

$$
\begin{aligned}
f_1' :=\ & \varepsilon\beta_3 f_1 - \beta_1 f_3 + \kappa\beta_3\nabla\beta_1\cdot\nabla\widetilde{\theta} \\
&+ \varepsilon\widetilde{p}\big(\partial_t(g_1\beta_3) + v\cdot\nabla(g_1\beta_3)\big) - \widetilde{\theta}\big(\partial_t(g_3\beta_1) + v\cdot(\nabla g_3\beta_1)\big).
\end{aligned}
$$

In order to symmetrize these equations, we make use of the assumption (4.2), that is $\beta_3 - \beta_1 > 0$. Multiply the first equation [resp. the third] by $\gamma_1(\beta_3-\beta_1)^{-1}$ [resp. $\gamma_2$]. By so doing, we obtain that $\widetilde{\mathcal{U}}$ solves

$$\mathcal{L}_1(v,\phi)\widetilde{\mathcal{U}} - \mathcal{L}_2(\mu,\kappa,\phi)\widetilde{\mathcal{U}} + \frac{1}{\varepsilon}S(\phi)\widetilde{\mathcal{U}} = F, \tag{4.15}$$

with $F := \big(\gamma_1(\beta_3-\beta_1)^{-1}f_1', \varepsilon f_2, \gamma_2 f_3\big)^T$, $\mathcal{L}_1(v,\phi) := L_0(\phi)\partial_t + \sum L_j(v,\phi)\partial_j$ where

$$L_0 := \begin{pmatrix} \frac{\gamma_1}{\beta_3-\beta_1} & 0 & 0 \\ 0 & g_2 I_d & 0 \\ 0 & 0 & \gamma_2 g_3 \end{pmatrix} \quad \text{and} \qquad L_j := v_j L_0.$$

Moreover, the singular and viscous perturbations are defined by

$$S = \begin{pmatrix} 0 & \gamma_1 \operatorname{div} & 0 \\ \nabla(\gamma_1\,\cdot) & 0 & \nabla(\gamma_2\,\cdot) \\ 0 & \gamma_2\operatorname{div} & 0 \end{pmatrix}, \mathcal{L}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \mu\beta_2\Delta + \mu\beta_2^\sharp\nabla\operatorname{div} & 0 \\ 0 & 0 & \kappa\gamma_2\beta_3\Delta \end{pmatrix}.$$

The end of the proof proceeds by multiplying by $\widetilde{\mathcal{U}}$ and integration by parts. The point is that the large terms in $\varepsilon^{-1}$ cancel out for $S$ is skew-symmetric. Let $t \in (0,T]$. If we multiply (4.15) by $\widetilde{\mathcal{U}}$ and integrate from 0 to $t$, we obtain

$$\big\langle L_0(\phi)\widetilde{\mathcal{U}}, \widetilde{\mathcal{U}}\big\rangle(t) - 2\int_0^t \big\langle \mathcal{L}_2(\phi)\widetilde{\mathcal{U}}, \widetilde{\mathcal{U}}\big\rangle\, d\tau = \big\langle L_0(\phi)\widetilde{\mathcal{U}}, \widetilde{\mathcal{U}}\big\rangle(0) + Y(t), \tag{4.16}$$

where $\langle\,\cdot\,,\,\cdot\,\rangle$ denotes the scalar product in $L^2(\mathbb{D})$ and

$$Y(t) := \int_0^t \Big\langle \Big\{\partial_t L_0(\phi) + \sum_{1\leqslant j\leqslant d} \partial_j L_j(v,\phi)\Big\}\widetilde{\mathcal{U}}, \widetilde{\mathcal{U}}\Big\rangle\, d\tau + 2\int_0^t \langle\, F\,, \widetilde{\mathcal{U}}\,\rangle\, dt.$$

Here we used the symmetry of the matrices $L_0$ and $L_j$.

We begin by estimating the right-hand side. One easily gathers

$$\Big\|\partial_t L_0(\phi) + \sum_{1\leqslant j\leqslant d}\partial_j L_j(v,\phi)\Big\|_{L^\infty} \leqslant C,$$

(where $C$ is as in Notation 4.8) and

$$|\langle\, F\,, \widetilde{\mathcal{U}}\,\rangle| \leqslant \|F\|_{L^2}\,\|\widetilde{\mathcal{U}}\|_{L^2} \leqslant C\,\|(f_1', \varepsilon f_2, f_3)\|_{L^2}\,\|\widetilde{\mathcal{U}}\|\,. \tag{4.17}$$

Furthermore, since by definition $f_1'$ is a linear combination of $\varepsilon f_1$, $f_3$, $\varepsilon\widetilde{p}$, $\widetilde{\theta}$ and $\kappa\nabla\widetilde{\theta}$ whose coefficients are estimated in $L^\infty$ norm by a constant depending only on $R$, we get

$$\|f_1'\| \leqslant C\,\|(f_3,\varepsilon f_1)\|_{L^2} + C\,\|(\widetilde{\theta},\varepsilon\widetilde{p})\|_{L^2} + C\,\|\kappa\nabla\widetilde{\theta}\|_{L^2}\,. \tag{4.18}$$

Moreover, observe that

$$(\widetilde{\theta}, \varepsilon\widetilde{p}, \varepsilon\widetilde{v})^t = M(\phi)\widetilde{\mathcal{U}}, \tag{4.19}$$

where $M(\phi)$ is an $n \times n$ invertible matrix. As a consequence, there exists a constant $C$ such that $C^{-1} \|\widetilde{\mathcal{U}}\|_{L^2} \leqslant \|(\widetilde{\theta}, \varepsilon\widetilde{p}, \varepsilon v)\|_{L^2} \leqslant C \|\widetilde{\mathcal{U}}\|_{L^2}$. Hence, the estimates (4.17) and (4.18) result in

$$|\langle F, \widetilde{\mathcal{U}} \rangle| \leqslant C B_\varepsilon(f, \widetilde{U}) + C \|\widetilde{\mathcal{U}}\|_{L^2}^2 + C \|\kappa \nabla \widetilde{\theta}\|_{L^2} \|\widetilde{\mathcal{U}}\|_{L^2}, \tag{4.20}$$

where $B_\varepsilon(f, \widetilde{U})$ is given by (4.14). For all $\lambda \geqslant 1$, the last term in the right hand-side of the previous estimate is bounded by

$$\lambda C \|\widetilde{\mathcal{U}}\|_{L^2}^2 + \frac{\kappa^2}{\lambda} \|\nabla \widetilde{\theta}\|_{L^2}^2 .$$

Consequently the right-hand side of (4.16) is less than

$$C_0 \|\widetilde{\mathcal{U}}(0)\|_{L^2}^2 + \lambda C \int_0^t \|\widetilde{\mathcal{U}}\|_{L^2}^2 \, d\tau + \frac{\kappa^2}{\lambda} \int_0^t \|\nabla \widetilde{\theta}\|_{L^2}^2 \, d\tau + C \int_0^t B_\varepsilon(f, \widetilde{U}) \, d\tau. \tag{4.21}$$

We want to estimate now the left-hand side of (4.16). In this regard, the most direct estimates show that

$$\langle L_0(\phi)\widetilde{\mathcal{U}}, \widetilde{\mathcal{U}} \rangle \geqslant \|L_0(\phi)^{-1}\|_{L^\infty}^{-1} \|\widetilde{\mathcal{U}}\|_{L^2}^2 .$$

Similarly, using Lemma 9.2 in Appendix, one has

$$-\langle \mathcal{L}_2(\phi)\widetilde{\mathcal{U}}, \widetilde{\mathcal{U}} \rangle \geqslant K\mu m \|\nabla \widetilde{v}_\varepsilon\|^2 + K\kappa m \|\nabla \widetilde{\theta}\|^2 - C \|\widetilde{\mathcal{U}}\|_{L^2}^2 ,$$

where $m := \min\{ \|\beta_2^{-1}\|_{L^\infty}^{-1}, \|(\beta_2 + \beta_2^\sharp)^{-1}\|_{L^\infty}^{-1}, \|\beta_3^{-1}\|_{L^\infty}^{-1} \}$ and $K$ is a generic constant depending only on the dimension.

To estimate the $L^\infty$ norm of $L_0^{-1}$, $\beta_2^{-1}$, $(\beta_2 + \beta_2^\sharp)^{-1}$ and $\beta_3^{-1}$, write

$$\begin{aligned} \|F(\phi(t))\|_{L^\infty} &\leqslant \|F(\phi(0))\|_{L^\infty} + \int_0^t \|\partial_t F(\phi(\tau))\|_{L^\infty} \, d\tau \\ &\leqslant C_0 + t \sup_{\tau \in [0,t]} \|F'(\phi(\tau))\|_{L^\infty} \|\partial_t \phi(\tau)\|_{L^\infty} \\ &\leqslant C_0 + TC \leqslant C_0 e^{TC}. \end{aligned} \tag{4.22}$$

Consequently, the left hand-side of (4.16) is greater than

$$C_0 e^{-TC} \Big\{ \|\widetilde{\mathcal{U}}(t)\|_{L^2}^2 + \int_0^t \mu \|\nabla \widetilde{v}_\varepsilon\|^2 + \kappa \|\nabla \widetilde{\theta}\|^2 \, d\tau \Big\} - C \int_0^t \|\widetilde{\mathcal{U}}\|_{L^2}^2 \, d\tau. \tag{4.23}$$

From this and the assumption $\kappa \leqslant 1$, we see that one can absorb the third term in (4.21) by taking $\lambda$ large enough. Then, Gronwall's Lemma implies

$$\|\widetilde{\mathcal{U}}(t)\|_{L^2}^2 + \int_0^t \mu \|\nabla \widetilde{v}_\varepsilon\|^2 + \kappa \|\nabla \widetilde{\theta}\|^2 \, d\tau \leqslant C_0 e^{TC} \|\widetilde{\mathcal{U}}(0)\|_{L^2}^2 + C \int_0^t B_\varepsilon(f, \widetilde{U}) \, d\tau.$$

We claim that the previous estimate holds true with $\widetilde{\mathcal{U}}$ replaced by $(\widetilde{\theta}, \varepsilon\widetilde{p}, \varepsilon\widetilde{v})$. Indeed, by (4.19) and (4.22), one has $\|(\widetilde{\theta}, \varepsilon\widetilde{p}, \varepsilon\widetilde{v})\|_{L^2} \leqslant C_0 e^{TC} \|\widetilde{\mathcal{U}}\|_{L^2}$.

To complete the proof, take the square root of the inequality thus obtained and take the supremum over $t \in [0, T]$. □

We next prove an estimate parallel to (4.7).

**Proposition 4.9.** *Using the same notations as in Theorem 4.3, we have*

$$\sup_{t\in[0,T]} \|(\widetilde{p},\widetilde{v})(t)\|_{L^2} + \left(\int_0^T \mu\|\nabla\widetilde{v}\|_{L^2}^2 + \kappa\|\operatorname{div}\widetilde{v}\|_{L^2}^2\,dt\right)^{1/2}$$
$$\leqslant C(R_0)e^{TC(R)}\|(\widetilde{p},\widetilde{v})(0)\|_{L^2} + C(R)\int_0^T \|(f_1,f_2)\|_{L^2} + \kappa\|f_3\|_{H^1}\,dt$$
$$+ C(R_0)e^{TC(R)}\Big\{\sup_{t\in[0,T]}\|\widetilde{\theta}(t)\|_{H^1_\nu} + \left(\int_0^T \kappa\|\nabla\widetilde{\theta}\|_{H^1_\nu}^2\,dt\right)^{1/2}\Big\}.$$

**Remark 4.10.** From now on, we make intensive and implicit uses of the assumptions $\mu\leqslant 1$ and $\kappa\leqslant 1$. In particular we freely use obvious estimates like $\mu\leqslant\sqrt{\mu}\leqslant\nu$. Similarly, we use the estimate $\nu\|\nabla u\|_{L^2}\leqslant\|u\|_{H^1_\nu}$ without mentioning it explicitly in the proofs.

*Proof.* Introduce $\widetilde{v}_e := \widetilde{v} - \kappa\beta_1\nabla\widetilde{\theta}$. Performing a little algebra we find that $\widetilde{\mathcal{V}} := (\widetilde{p},\widetilde{v}_e)^t$ satisfies

$$\begin{pmatrix} g_1 & 0\\ 0 & g_2 I_d\end{pmatrix}\big(\partial_t\widetilde{\mathcal{V}} + v\cdot\nabla\widetilde{\mathcal{V}}\big) + \frac{1}{\varepsilon}\begin{pmatrix} 0 & \operatorname{div}\\ \nabla & 0\end{pmatrix}\widetilde{\mathcal{V}} - \begin{pmatrix} 0 & 0\\ 0 & \mathcal{L}_{22}\end{pmatrix}\widetilde{\mathcal{V}} = F,$$

where

$$\mathcal{L}_{22} := \kappa g_2\beta_1\nabla(g_3^{-1}\operatorname{div}\cdot\,) + \mu\beta_2\Delta + \mu\beta_2^\sharp\nabla\operatorname{div},$$

and $F = (f_1, f_2 + f_2')^t$ with

(4.24)
$$f_2' := \mu\kappa\beta_2\Delta(\beta_1\nabla\widetilde{\theta}) + \mu\kappa\beta_2^\sharp\nabla\operatorname{div}(\beta_1\nabla\widetilde{\theta}) - \kappa g_2\big(\partial_t\beta_1 + v\cdot\nabla\beta_1\big)\nabla\widetilde{\theta}$$
$$+ \kappa g_2\beta_1\nabla\big(\kappa g_3^{-1}\operatorname{div}(\beta_1\nabla\widetilde{\theta}) - \kappa g_3^{-1}\beta_3\Delta\widetilde{\theta} - g_3^{-1}f_3\big) + \kappa g_2\beta_1(\nabla v)^t\nabla\widetilde{\theta}.$$

The thing of greatest interest here is that $-\mathcal{L}_{22}$ is a differential operator whose leading symbol is greater than $C\kappa\xi^t\xi + C\mu|\xi|^2 I_d$.

As before, the proof proceeds by multiplying by $\widetilde{\mathcal{V}}$ and integrating on the strip $[0,t]\times\mathbb{D}$. Then the analysis establishing that the right [resp. left] hand side of (4.16) is smaller than (4.21) [resp. greater than (4.23)] also gives:

$$\|\widetilde{\mathcal{V}}(t)\|_{L^2}^2 + \int_0^t \mu\|\nabla\widetilde{v}_e\|_{L^2}^2 + \kappa\|\operatorname{div}\widetilde{v}_e\|_{L^2}^2\,d\tau$$
$$\leqslant C_0e^{TC}\|\widetilde{\mathcal{V}}(0)\|_{L^2}^2 + C\int_0^t\|\widetilde{\mathcal{V}}\|_{L^2}^2\,d\tau + C\int_0^t|\langle F,\widetilde{\mathcal{V}}\rangle|\,d\tau. \qquad (4.25)$$

Let us estimate the last term in the right-hand side of (4.25). Using (4.24) one can decompose $f_2'$ as $f_{2,1}' + \sqrt{\kappa}\nabla f_{2,2}'$ where

$$\big\|f_{2,1}'\big\|_{L^2} \leqslant C\kappa\|f_3\|_{H^1} + C\kappa\|\nabla\widetilde{\theta}\|_{L^2} + C\sqrt{\kappa}(\mu+\kappa)\|\nabla^2\widetilde{\theta}\|_{L^2},$$
$$\big\|f_{2,2}'\big\|_{L^2} \leqslant C\kappa\|\nabla\widetilde{\theta}\|_{L^2} + C\sqrt{\kappa}(\mu+\kappa)\|\nabla^2\widetilde{\theta}\|_{L^2}.$$

By notation, $\langle F, \widetilde{\mathcal{V}} \rangle = \langle f_1, \widetilde{p} \rangle + \langle f_2, \widetilde{v}_e \rangle + \langle f'_{2,1}, \widetilde{v}_e \rangle - \langle f'_{2,2}, \sqrt{\kappa}\operatorname{div} \widetilde{v}_e \rangle$. We thus get, for all $\lambda \geqslant 1$,

$$
\begin{aligned}
\int_0^t \left|\langle F, \widetilde{\mathcal{V}} \rangle\right| d\tau \leqslant \lambda C\Big(\int_0^t \|(f_1, f_2)\|_{L^2} + \kappa \|f_3\|_{H^1} \, d\tau\Big)^2 \\
+ \frac{1}{\lambda}\int_0^t \kappa \|\operatorname{div} \widetilde{v}_e\|_{L^2}^2 \, d\tau + \frac{1}{\lambda}\sup_{\tau\in[0,t]} \|\widetilde{\mathcal{V}}(t)\|_{L^2}^2 \\
+ \lambda C \int_0^t \|(\kappa\nabla\widetilde{\theta}, \sqrt{\kappa}(\mu+\kappa)\nabla^2\widetilde{\theta}\|_{L^2}^2 \, d\tau.
\end{aligned}
\tag{4.26}
$$

Replacing $\widetilde{v}_e$ with $\widetilde{v}$ in (4.25)–(4.26), taking $\lambda$ large enough and applying Gronwall's lemma leads to the expected result. □

4.4. **End of the proof of Theorem 4.3.** From now on, we consider a time $0 < T \leqslant 1$, a fixed triple of parameter $a = (\varepsilon, \mu, \kappa) \in A$ and a solution $\widetilde{U} = (\widetilde{p}, \widetilde{v}, \widetilde{\theta})$ of the system (4.1). We denote by $R_0$ and $R$ the norms defined in the statement of Theorem 4.3 (see (4.3)). We also set $\nu := \sqrt{\mu+\kappa}$.

Introduce the functions $N, w, y, Y, Z \colon [0, T] \to \mathbb{R}_+$ given by

$$
N(t) := w(t) + \nu z(t) + y(t) + \nu Y(t),
\tag{4.27}
$$

where

$$
w(t) := \sup_{\tau\in[0,t]} \|(\widetilde{p}, \widetilde{v})(\tau)\|_{L^2} + \Big(\int_0^t \mu \|\nabla\widetilde{v}\|_{L^2}^2 + \kappa \|\operatorname{div}\widetilde{v}\|_{L^2}^2 \, d\tau\Big)^{1/2},
$$

$$
y(t) := \sup_{\tau\in[0,t]} \|(\widetilde{\theta}, \varepsilon\widetilde{p}, \varepsilon\widetilde{v})(\tau)\|_{L^2} + \Big(\int_0^t \kappa \|\nabla\widetilde{\theta}\|_{L^2}^2 + \mu \|\varepsilon\nabla\widetilde{v}\|_{L^2}^2 \, d\tau\Big)^{1/2},
$$

$$
Y(t) := \sup_{\tau\in[0,t]} \|\nabla(\widetilde{\theta}, \varepsilon\widetilde{p}, \varepsilon\widetilde{v})(\tau)\|_{L^2} + \Big(\int_0^t \kappa \|\nabla^2\widetilde{\theta}\|_{L^2}^2 + \mu \|\varepsilon\nabla^2\widetilde{v}\|_{L^2}^2 \, d\tau\Big)^{1/2},
$$

$$
z(t) := \Big(\int_0^t \|\nabla\widetilde{p}\|_{L^2} \, d\tau\Big)^{1/2}.
$$

Note that $\|\widetilde{U}\|_{\mathcal{H}_a^0(T)} \leqslant N(T)$. Since we have estimated $w(T)$ and $y(T)$, it remains only to estimate $\nu Y(T)$ and $\nu z(T)$. Parallel to (4.10), we begin by establishing an estimate for $Y$.

**Lemma 4.11.** *There exists a constant $C_0$ depending only on $R_0$ and a constant $C$ depending only on $R$ such that for all $t \in [0, T]$ and all $\lambda \geqslant 1$,*

$$
\begin{aligned}
Y(t)^2 \leqslant C_0 e^{TC} Y(0)^2 + \lambda C \int_0^t Y(\tau)^2 \, d\tau + \frac{C}{\lambda}\int_0^t \|(\operatorname{div}\widetilde{v}, \nabla\widetilde{p})\|_{L^2}^2 \, d\tau \\
+ \Big(C \int_0^t \|f_3\|_{H^1} + \|(f_1, f_2)\|_{H^1_\varepsilon} \, d\tau\Big)^2.
\end{aligned}
\tag{4.28}
$$

*Proof.* The proof, if tedious, is elementary. Indeed, the strategy for proving (4.28) consists in differentiating the system (4.1) so as to apply Proposition 4.6 with $(\widetilde{\theta}, \varepsilon\widetilde{p}, \varepsilon\widetilde{v})$ replaced by $\nabla(\widetilde{\theta}, \varepsilon\widetilde{p}, \varepsilon\widetilde{v})$.

Let $1 \leqslant j \leqslant d$ and set

$$\widetilde{U}_j := (\widetilde{p}_j, \widetilde{v}_j, \widetilde{\theta}_j) := (\partial_j \widetilde{p}, \partial_j \widetilde{v}, \partial_j \widetilde{\theta}).$$

We commute[1] the $i$th equation in (4.1) with $g_i(\partial_j g_i^{-1} \cdot)$. It yields

$$\begin{cases} g_1(\phi)\big(\partial_t \widetilde{p}_j + v \cdot \nabla \widetilde{p}_j\big) + \dfrac{1}{\varepsilon} \operatorname{div} \widetilde{v}_j - \dfrac{\kappa}{\varepsilon} \operatorname{div}(\beta_1(\phi) \nabla \widetilde{\theta}_j) = F_1, \\ g_2(\phi)\big(\partial_t \widetilde{v}_j + v \cdot \nabla \widetilde{v}_j\big) + \dfrac{1}{\varepsilon} \nabla \widetilde{p}_j - \mu\beta_2(\phi)\Delta \widetilde{v}_j - \mu\beta_2^\sharp(\phi)\nabla \operatorname{div} \widetilde{v}_j = F_2, \\ g_3(\phi)\big(\partial_t \widetilde{\theta}_j + v \cdot \nabla \widetilde{\theta}_j\big) + \operatorname{div} \widetilde{v}_j - \kappa\beta_3(\phi)\Delta \widetilde{\theta}_j = F_3, \end{cases}$$

where, for $i = 1, 2, 3$, the source term $F_i$ is given by

$$F_i := F_i^1 + F_i^2 := g_i \partial_j (g_i^{-1} f_i) + g_i \tilde{F}_i^2,$$

with

$$\begin{aligned} \tilde{F}_1^2 &:= -\,\partial_j v \cdot \nabla \widetilde{p} - \frac{1}{\varepsilon} \partial_j g_1^{-1} \big\{ \operatorname{div} \widetilde{v} - \kappa \operatorname{div}(\beta_1 \nabla \widetilde{\theta}) \big\} + \frac{\kappa}{\varepsilon} g_1^{-1} \operatorname{div}(\partial_j \beta_1 \nabla \widetilde{\theta}), \\ \tilde{F}_2^2 &:= -\,\partial_j v \cdot \nabla \widetilde{v} - \frac{1}{\varepsilon} \partial_j g_2^{-1} \nabla \widetilde{p} + \mu \partial_j (g_2^{-1} \beta_2) \Delta \widetilde{v} + \mu \partial_j (g_2^{-1} \beta_2^\sharp) \nabla \operatorname{div} \widetilde{v}, \\ \tilde{F}_3^2 &:= -\,\partial_j v \cdot \nabla \widetilde{\theta} - \partial_j g_3^{-1} \operatorname{div} \widetilde{v} + \kappa \partial_j (g_3^{-1} \beta_3) \Delta \widetilde{\theta}. \end{aligned}$$

Proposition 4.6 implies that

$$Y(t)^2 \leqslant C_0 e^{TC} \big\|\widetilde{\mathcal{U}}(0)\big\|_{L^2}^2 + C \sum_{1 \leqslant j \leqslant d} \int_0^t B_\varepsilon(F, \widetilde{U}_j)\, d\tau. \tag{4.29}$$

where $\widetilde{\mathcal{U}} := (\nabla\widetilde{\theta}, \varepsilon\nabla\widetilde{p}, \varepsilon\nabla\widetilde{v})$ and $B_\varepsilon$ is defined by (4.14). We now have to estimate $B_\varepsilon(F, \widetilde{U}_j)$. The source terms are directly estimated by

$$\begin{aligned} &\big\| \big(F_3^1, \varepsilon F_2^1, \varepsilon F_3^1\big) \big\|_{L^2} \leqslant C \|f_3\|_{H^1} + C \|(f_1, f_2)\|_{H^1_\varepsilon}, \\ &\big\| \big(\tilde{F}_3^2, \varepsilon \tilde{F}_2^2, \varepsilon \tilde{F}_3^2\big) \big\|_{L^2} \leqslant C \big( \|\widetilde{\mathcal{U}}\|_{L^2} + \|(\operatorname{div} \widetilde{v}, \nabla \widetilde{p})\|_{L^2} + \|(\varepsilon\mu\nabla^2 \widetilde{v}, \kappa \nabla^2 \widetilde{\theta})\|_{L^2} \big). \end{aligned}$$

The last estimate implies that, for all positive $\lambda$ and $\lambda'$, one has

$$\begin{aligned} B_\varepsilon(F^2, \widetilde{U}_j) \leqslant{}& (1 + \lambda + \lambda') C \|\widetilde{\mathcal{U}}\|_{L^2}^2 \\ &+ \frac{C}{\lambda} \|(\operatorname{div} \widetilde{v}, \nabla \widetilde{p})\|_{L^2}^2 + \frac{C}{\lambda'} \|(\varepsilon\mu\nabla^2 \widetilde{v}, \kappa \nabla^2 \widetilde{\theta})\|_{L^2}^2. \end{aligned}$$

Moreover, the term $\int_0^t |B_\varepsilon(F^1, \widetilde{U}_j)|\, d\tau$ is estimated as in the proof of Corollary 4.7. With these estimates in hands, we find that the last term in the

[1] Which means that we commute $\partial_j$ with the $i$th equation ($i = 1, 2, 3$) premultiplied by $g_i^{-1}$ and we next multiply the result by $g_i$.

right-side of (4.29) is less then

$$
\begin{aligned}
&\int_0^t (1+\lambda+\lambda')C\,\|\widetilde{\mathcal{U}}\|_{L^2}^2\,d\tau+\frac{C}{\lambda''}\sup_{\tau\in[0,t]}\|\widetilde{\mathcal{U}}\|_{L^2}^2\\
&+\int_0^t \frac{C}{\lambda}\,\|(\operatorname{div}\widetilde{v},\nabla\widetilde{p})\|_{L^2}^2\,d\tau+\frac{C}{\lambda'}\int_0^t\|(\varepsilon\mu\nabla^2\widetilde{v},\kappa\nabla^2\widetilde{\theta})\|_{L^2}^2\,d\tau\\
&+\lambda''\Big(C\int_0^t\|f_3\|_{H^1}+\|(f_1,f_2)\|_{H^1_\varepsilon}\,d\tau\Big)^2.
\end{aligned}
\tag{4.30}
$$

By definition, $\sup_{\tau\in[0,t]}\|\widetilde{\mathcal{U}}\|_{L^2}^2+\int_0^t\|(\varepsilon\mu\nabla^2\widetilde{v},\kappa\nabla^2\widetilde{\theta})\|_{L^2}^2\,d\tau\leqslant Y(t)^2$. Hence, taking $\lambda'=\lambda''$ large enough, one can absorb the second and fourth terms in (4.30) in the left hand side of (4.29), thereby obtaining the desired estimate (4.28). □

To obtain a closed set of inequalities it remains to estimate $z(t)$.

**Lemma 4.12.** *There exists a constant $C_0$ depending only on $R_0$ and a constant $C$ depending only on $R$ such that for all $t\in[0,T]$,*

$$
\nu^2 z(t)^2\leqslant C_0e^{TC}\big\{\nu^2Y(t)^2+w(t)^2\big\}+C\Big(\int_0^T\|(f_1,f_2)\|_{L^2}\,dt\Big)^2.
\tag{4.31}
$$

*Proof.* Set $\Omega_t:=[0,t]\times\mathbb{D}$ and denote by $\langle\!\langle\ ,\ \rangle\!\rangle$ the scalar product in $L^2\big(\Omega_t\big)$. Multiplying the second equation in (4.1) by $\varepsilon\nu^2g_2^{-1}\nabla\widetilde{p}$, and integrating over $\Omega_t$ yields

$$
\begin{aligned}
&\nu^2\langle\!\langle\, g_2^{-1}\nabla\widetilde{p},\,\nabla\widetilde{p}\,\rangle\!\rangle=\\
&\qquad -\varepsilon\nu^2\langle\!\langle\,(\partial_t+v\cdot\nabla)\widetilde{v},\,\nabla\widetilde{p}\,\rangle\!\rangle+\varepsilon\mu\nu^2\langle\!\langle\, g_2^{-1}\beta_2\Delta\widetilde{v},\,\nabla\widetilde{p}\,\rangle\!\rangle\\
&\qquad +\varepsilon\mu\nu^2\langle\!\langle\, g_2^{-1}\beta_2^\sharp\nabla\operatorname{div}\widetilde{v},\,\nabla\widetilde{p}\,\rangle\!\rangle+\varepsilon\nu^2\langle\!\langle\, g_2^{-1}f_2\,,\,\nabla\widetilde{p}\,\rangle\!\rangle.
\end{aligned}
\tag{4.32}
$$

The most direct estimates show that

$$
\begin{aligned}
\langle\!\langle\, g_2^{-1}\nabla\widetilde{p},\,\nabla\widetilde{p}\,\rangle\!\rangle&\geqslant\|g_2\|_{L^\infty(\Omega_t)}^{-1}\|\nabla\widetilde{p}\|_{L^2(\Omega_t)}^2\,,\\
\varepsilon\mu\nu^2\langle\!\langle\, g_2^{-1}\beta_2\Delta\widetilde{v},\,\nabla\widetilde{p}\,\rangle\!\rangle&\leqslant\|g_2^{-1}\beta_2\|_{L^\infty(\Omega_t)}\|\varepsilon\mu\nu\nabla^2\widetilde{v}\|_{L^2(\Omega_t)}\|\nu\nabla\widetilde{p}\|_{L^2(\Omega_t)}\,,\\
\varepsilon\mu\nu^2\langle\!\langle\, g_2^{-1}\beta_2^\sharp\nabla\operatorname{div}\widetilde{v},\,\nabla\widetilde{p}\,\rangle\!\rangle&\leqslant\|g_2^{-1}\beta_2^\sharp\|_{L^\infty(\Omega_t)}\|\varepsilon\mu\nu\nabla^2\widetilde{v}\|_{L^2(\Omega_t)}\|\nu\nabla\widetilde{p}\|_{L^2(\Omega_t)}\,,\\
\varepsilon\nu^2\langle\!\langle\, g_2^{-1}f_2\,,\,\nabla\widetilde{p}\,\rangle\!\rangle&\leqslant\|g_2^{-1}\|_{L^\infty(\Omega_t)}\|f_2\|_{L^1_tL^2}\|\varepsilon\nu^2\nabla\widetilde{p}\|_{C^0_tL^2}\,,
\end{aligned}
$$

where we used the shorthand notations

$$
\|\cdot\|_{C^0_tX}:=\sup_{\tau\in[0,t]}\|\cdot\|_X\quad\text{and}\quad\|\cdot\|_{L^p_tX}:=\Big(\int_0^t\|\cdot\|_X^p\,d\tau\Big)^{1/p}.
$$

With the bound (4.22), the previous four estimates and (4.32) imply (4.33)

$$
\begin{aligned}
\|\nu\nabla\widetilde{p}\|_{L^2(\Omega_t)}^2&\leqslant C_0e^{TC}\left|\varepsilon\nu^2\langle\!\langle\,(\partial_t+v\cdot\nabla)\widetilde{v},\,\nabla\widetilde{p}\,\rangle\!\rangle\right|\\
&\quad+C_0e^{TC}\Big(\|\varepsilon\mu\nu\nabla^2\widetilde{v}\|_{L^2(\Omega_t)}+\|f_2\|_{L^1_tL^2}+\|\varepsilon\nu^2\nabla\widetilde{p}\|_{C^0_tL^2}\Big)^2.
\end{aligned}
$$

Note that the second term in the right-hand side of the previous estimate is bounded by the right-hand side of (4.31). It thus remains to estimate the first term. To do so integrate by parts in both the space and time variables to obtain

$$
\begin{aligned}
&\varepsilon\nu^2\big\langle\!\big\langle(\partial_t + v\cdot\nabla)\widetilde{v},\, \nabla\widetilde{p}\big\rangle\!\big\rangle\\
&\qquad = \varepsilon\nu^2\big\langle\!\big\langle \operatorname{div}\widetilde{v},\, (\partial_t + v\cdot\nabla)\widetilde{p}\big\rangle\!\big\rangle\\
&\qquad\quad - \varepsilon\nu^2\big\langle \operatorname{div}\widetilde{v}(t)\,,\, \widetilde{p}(t)\big\rangle + \varepsilon\nu^2\big\langle \operatorname{div}\widetilde{v}(0)\,,\, \widetilde{p}(0)\big\rangle\\
&\qquad\quad + \varepsilon\nu^2\langle\!\langle\, \nabla v\,,\, \nabla\widetilde{p}\otimes\widetilde{v}\,\rangle\!\rangle - \varepsilon\nu^2\langle\!\langle\, \operatorname{div} v\,,\, \widetilde{v}\cdot\nabla\widetilde{p}\,\rangle\!\rangle.
\end{aligned}
\tag{4.34}
$$

The last two terms in the right-hand side of (4.34) are estimated by

$$
Kt\,\|\nabla v\|^2_{C^0_tL^\infty}\,\|\widetilde{v}\|^2_{C^0_tL^2} + \|\varepsilon\nu\nabla\widetilde{p}\|^2_{C^0_tL^2} \leqslant TCw(t)^2 + \nu^2 Y(t)^2,
$$

and the second and third terms are estimated by

$$
\|\varepsilon\nu\nabla\widetilde{v}\|^2_{C^0_tL^2} + \|\widetilde{p}\|^2_{C^0_tL^2} \leqslant \nu^2 Y(t)^2 + w(t)^2.
$$

In particular, the sum of the last four terms in the right-hand side of (4.34) is estimated by $e^{TC}\{w(t)^2 + \nu^2 Y(t)^2\}$. This brings us to estimate the first term. Using the first equation in (4.1), we get

$$
\begin{aligned}
\varepsilon\nu^2\langle\!\langle\, \operatorname{div}\widetilde{v},\, (\partial_t + v\cdot\nabla)\widetilde{p}\,\rangle\!\rangle &= -\nu^2\big\langle\!\big\langle \operatorname{div}\widetilde{v},\, g_1^{-1}\operatorname{div}\widetilde{v}\big\rangle\!\big\rangle\\
&+ \nu^2\kappa\langle\!\langle\, \operatorname{div}\widetilde{v},\, g_1^{-1}\operatorname{div}(\beta_1\nabla\widetilde{\theta})\,\rangle\!\rangle + \varepsilon\nu^2\langle\!\langle\, \operatorname{div}\widetilde{v},\, g_1^{-1}f_1\,\rangle\!\rangle.
\end{aligned}
$$

Then, the analysis establishing (4.33) also gives

$$
\begin{aligned}
&\varepsilon\nu^2\,|\langle\!\langle\, \operatorname{div}\widetilde{v},\, (\partial_t + v\cdot\nabla)\widetilde{p}\,\rangle\!\rangle|\\
&\quad\leqslant C_0e^{TC}\Big\{\nu^2\,\|(\operatorname{div}\widetilde{v}, \kappa\nabla^2\widetilde{\theta})\|^2_{L^2(\Omega_t)} + \nu^2\kappa\,\|(\nabla\widetilde{\theta}, \varepsilon\nabla\widetilde{v})\|^2_{C^0_tL^2} + \|f_1\|^2_{L^1_tL^2}\Big\},
\end{aligned}
$$

which in turn is bounded by the right-hand side of (4.31). The proof is complete. □

We thus have proved a closed system of estimates. Indeed, by taking appropriate combinations of the previous estimates, we see that there exists a constant $C_0$ depending only on $R_0$ and a constant $C$ depending only on $R$ such that for all $t\in[0,T]$ and all $\lambda\geqslant 1$, the norm $N(t)$ (as defined in (4.27)) satisfies the estimate

$$
N(t)^2 \leqslant C_0e^{TC}N(0)^2 + \lambda C\int_0^t N(\tau)^2\,d\tau + \frac{C}{\lambda}N(t)^2 + CF(T)^2,
$$

where $F(T) := \int_0^T \|(f_1,f_2)\|_{H^1_{\varepsilon\nu}} + \|f_3\|_{H^1_\nu}\,dt$.

Again, take $\lambda$ large enough and apply the Gronwall's lemma. The proof of Theorem 4.3 is complete since $\|\widetilde{U}\|_{\mathcal{H}^0_a(T)} \leqslant N(T)$.

## 5. High frequency regime

To establish Theorem 2.7, the crucial part consists of obtaining *a priori* estimates in Sobolev norms independent of $a \in A$. Theorem 4.3 provides the basic $L^2$ estimates. However, the estimates of the derivatives also require a careful analysis. Indeed, the classical approach, which consists of differentiating the equations, certainly fails since it reveals unbounded terms in $\varepsilon^{-1}$. However, one can follow this strategy in the high frequency regime where the parabolic behavior prevails.

More precisely, given a smooth solution $U = (p, v, \theta)$ to (2.1), we will estimate the $\mathcal{H}^s_a(T)$ norm of $(I - J_h)U$, where $\mathcal{H}^s_a(T)$ is as defined in §2.2 (see Definition 2.4) and, in keeping with the notations of §3, $\{ J_h \,|\, h \in [0,1] \}$ is a Friedrichs' mollifier and $\Lambda^s := (I - \Delta)^{s/2}$.

In order to make our energy estimates applicable, the main difficulty is to verify that the commutator of $(I - J_h)\Lambda^s$ and the equations (2.1) can be seen as a source term. To do so we first note that for $h = O(\varepsilon)$, one can gain an extra factor $\varepsilon$ in the commutator estimates [see (5.2) below]. Yet, this costs a derivative. To compensate this loss of derivative, we use in an essential way the parabolic behavior of the equations. Consequently, we search $h$ under the form $c(\mu, \kappa)\varepsilon$. Since for our purposes the main smoothing effect concerns the penalized terms $\operatorname{div} v$ and $\nabla p$, we take $c(\mu, \kappa) = \sqrt{\mu + \kappa}$.

**Proposition 5.1.** *Given an integer $s > 1 + d/2$, there exists a continuous nondecreasing function $C$ such that for all $a = (\varepsilon, \mu, \kappa) \in A$, all $T \in [0,1]$ and all $U = (p, v, \theta) \in C^1([0,T]; H^\infty(\mathbb{D}))$ satisfying* (2.1),

$$\|(I - J_{\varepsilon\nu})U\|_{\mathcal{H}^s_a(T)} \leqslant C(\Omega_0) e^{\sqrt{T} C(\Omega)}, \tag{5.1}$$

*where $\nu := \sqrt{\mu + \kappa}$, $\Omega_0 := \|U(0)\|_{\mathcal{H}^s_{a,0}}$, $\Omega := \|U\|_{\mathcal{H}^s_a(T)}$ (see §2.2, Definition 2.4).*

**Remark 5.2.** Note that we establish estimates for the exact solutions of (2.1) and we do not estimate the solutions of the linearized system (4.1).

### 5.1. Preliminaries.

To avoid interruptions of the proofs later on, we now collect a few nonlinear estimates which will be used throughout this section.

**Lemma 5.3.** *Let $s > 1 + d/2$. There exists a constant $K$ such that for all $h \in [0,1]$, all $\varrho \geqslant 0$, all $f \in H^{s+1}(\mathbb{D})$ and all $u \in H^s(\mathbb{D})$, we have*

$$\big\| \big[ f, (I - J_h)\Lambda^s \big] u \big\|_{H^1_\varrho} \leqslant (\varrho + h) K \big\{ \|f\|_{W^{1,\infty}} \|u\|_{H^s} + \|f\|_{H^{s+1}} \|u\|_{L^\infty} \big\}, \tag{5.2}$$

*where $\|f\|_{W^{1,\infty}} := \|f\|_{L^\infty} + \|\nabla f\|_{L^\infty}$. Recall that $\|\cdot\|_{H^\sigma_\varrho} := \|\cdot\|_{H^{\sigma-1}} + \varrho \|\cdot\|_{H^\sigma}$.*

**Remark 5.4.** This inequality provides a way to gain an extra factor $h$ [with $\varrho = h$] or a derivative [with $\varrho = 1$]. In this respect it is like the estimate (3.10), to which we will return in a moment (see Lemma 6.6).

*Proof.* To prove this result we need a tame estimate version of (3.9). We use the following result, which is Proposition 3.6.A in [41] (see also [25]). Let

$m > 0$. For all Fourier multiplier $\mathcal{Q}$ with symbol $q \in S^m$ and for all $\sigma \geqslant 0$, we have

$$\big\| [f, \mathcal{Q}] u \big\|_{H^\sigma} \leqslant K \|f\|_{W^{1,\infty}} \|u\|_{H^{\sigma+m-1}} + K \|f\|_{H^{m+\sigma}} \|u\|_{L^\infty}, \tag{5.3}$$

where the constant $K$ depends only on $m$, $s$, $\sigma$, $d$ and a finite number of semi-norms of $q$ in $S^m$.

Since the support of the symbol $1 - \jmath_h$ is limited by $|\xi| \geqslant 1/h$, the most direct estimates show that $\{ h^{-1}(1 - \jmath(h\xi))\langle\xi\rangle^s \,|h \in [0,1] \}$ is uniformly bounded in the symbol class $S^{s+1}$. Thus, the commutator estimate (5.3), applied with $(m, \sigma) := (s+1, 0)$, implies that

$$\big\| [f, (I - J_h)\Lambda^s] u \big\|_{L^2} \lesssim h\big( \|f\|_{W^{1,\infty}} \|u\|_{H^s} + \|f\|_{H^{s+1}} \|u\|_{L^\infty} \big). \tag{5.4}$$

On the other hand, the family $\{ (1 - \jmath(h\xi))\langle\xi\rangle^s \,|h \in [0,1] \}$ is uniformly bounded in $S^s$, so that (5.3) applied with $(m, \sigma) = (s, 1)$ implies that

$$\big\| [f, (I - J_h)\Lambda^s] u \big\|_{H^1} \lesssim \|f\|_{W^{1,\infty}} \|u\|_{H^s} + \|f\|_{H^{s+1}} \|u\|_{L^\infty}. \tag{5.5}$$

Combining (5.4) with (5.5) multiplied by $\varrho$, yields (5.2). □

We next prove two Moser-type estimates for the norms $\|\cdot\|_{H^\sigma_\varrho}$.

**Lemma 5.5.** *Let $s > 1 + d/2$. There exists a constant $K$ such that for all $\varrho \geqslant 0$ and for all $u_i \in H^s_\varrho(\mathbb{D})$,*

$$\|u_1 u_2\|_{H^s_\varrho} \leqslant K \|u_1\|_{H^s_\varrho} \|u_2\|_{H^s_\varrho}. \tag{5.6}$$

*This result extends to vector valued functions.*

*Proof.* Using the standard tame estimate for products, we get

$$\|u_1 u_2\|_{H^{s-1}} \lesssim \|u_1\|_{L^\infty} \|u_2\|_{H^{s-1}} + \|u_1\|_{H^{s-1}} \|u_2\|_{L^\infty},$$
$$\|u_1 u_2\|_{H^s} \lesssim \|u_1\|_{L^\infty} \|u_2\|_{H^s} + \|u_1\|_{H^s} \|u_2\|_{L^\infty}.$$

We next put the parameter $\varrho$ in appropriate spots, to obtain

$$\|u_1 u_2\|_{H^s_\varrho} \lesssim \|u_1\|_{L^\infty} \|u_2\|_{H^s_\varrho} + \|u_1\|_{H^s_\varrho} \|u_2\|_{L^\infty},$$

which completes the proof since $\|u\|_{L^\infty} \lesssim \|u\|_{H^{s-1}} \leqslant \|u\|_{H^s_\varrho}$. □

**Lemma 5.6.** *Let $s > 1 + d/2$ and $F \colon \mathbb{R}^n \to \mathbb{C}$ be a $C^\infty$ function such that $F(0) = 0$. Then, for all $\varrho \geqslant 0$ and all $u \in H^s_\varrho(\mathbb{D})$ with values in $\mathbb{R}^n$,*

$$\|F(u)\|_{H^s_\varrho} \leqslant C\big(\|u\|_{H^s_\varrho}\big), \tag{5.7}$$

*where $C(\cdot)$ depends only on a finite number of semi-norms of $F$ in $C^\infty$.*

*Proof.* Recall that, for all $\sigma_0 > d/2$,

$$\|F(u)\|_{H^{\sigma_0}} \leqslant C(\|u\|_{L^\infty}) \|u\|_{H^{\sigma_0}}. \tag{5.8}$$

Using this estimate at orders $s-1$ and $s$, we have

$$\|F(u)\|_{H^s_\varrho} \leqslant C\big(\|u\|_{L^\infty}\big) \|u\|_{H^s_\varrho}. \tag{5.9}$$

This in turn implies the desired estimate. □

**Lemma 5.7.** *Let $s > 1 + d/2$, $F\colon \mathbb{R}^n \to \mathbb{C}$ be a $C^\infty$ function such that $F(0) = 0$ and $\mathcal{Q}$ be a Fourier multiplier with symbol $q \in S^s$. For all vector-valued function $u \in H^s(\mathbb{D})$ we have*

$$\big\|\mathcal{Q}\big(F(u)\big) - F'(u)\mathcal{Q}u\big\|_{H^1} \leqslant C\big(\|u\|_{H^s}\big), \tag{5.10}$$

*where $F'$ is the differential of $F$ and $C(\cdot)$ is a smooth nondecreasing function depending only on $s$, $d$, a finite number of semi-norms of $q$ in $S^s$ and a finite number of semi-norms of $F$ in $C^\infty$.*

*Proof.* To establish (5.10) we use the para-differential calculus of Bony [2]. Denote by $T_f$ the operator of para-multiplication by $f$. Starting from

$$F(u) = T_{F'(u)}u + R(u; x, D_x)u,$$

where $R(u; x, D_x)$ is a smoothing operator (see (5.11)), we obtain

$$\mathcal{Q}\big(F(u)\big) - F'(u)\mathcal{Q}u = \big(T_{F'(u)} - F'(u)\big)\mathcal{Q}u + \big[\mathcal{Q}, T_{F'(u)}\big]u + \mathcal{Q}R(u; x, D_x)u.$$

The claim then follows from the bounds (3.15)–(3.17) and the estimate

$$\|R(u; x, D_x)u\|_{H^{s+1}} \leqslant C\big(\|u\|_{H^s}\big). \tag{5.11}$$

See [17, Th. 10.3.1] and [17, Corollary 9.3.6] for the proof of (5.11). □

5.2. **Localization in the high frequency region.** To simplify the presentation, we fix a real number $s$ strictly greater than $1 + d/2$.

To proceed further, we need some more terminology.

**Notation 5.8.** For all $h \in [0, 2]$, define

$$\mathcal{Q}_h := (I - J_h)\Lambda^s.$$

Hereafter, the parameter $h$ is a product $\varepsilon\nu$ with $(\varepsilon, \nu) \in [0, 1] \times [0, 2]$.

Introduce next the commutator of the equations (2.1) and $\mathcal{Q}_{\varepsilon\nu}$.

**Notation 5.9.** Given $a = (\varepsilon, \mu, \kappa) \in A$, $\nu \in [0, 2]$ and $U = (p, v, \theta)$, set

$$\begin{aligned}
f^{a,\nu}_{1,\mathrm{HF}}(U) &:= \big[g_1(\phi), \mathcal{Q}_{\varepsilon\nu}\big]\partial_t p + \big[g_1(\phi)v, \mathcal{Q}_{\varepsilon\nu}\big]\cdot \nabla p \; - \frac{\kappa}{\varepsilon}\big[B_1(\phi), \mathcal{Q}_{\varepsilon\nu}\big]\theta,\\
f^{a,\nu}_{2,\mathrm{HF}}(U) &:= \big[g_2(\phi), \mathcal{Q}_{\varepsilon\nu}\big]\partial_t v + \big[g_2(\phi)v, \mathcal{Q}_{\varepsilon\nu}\big]\cdot \nabla v \; - \mu\,\big[B_2(\phi), \mathcal{Q}_{\varepsilon\nu}\big]v,\\
f^{a,\nu}_{3,\mathrm{HF}}(U) &:= \big[g_3(\phi), \mathcal{Q}_{\varepsilon\nu}\big]\partial_t \theta + \big\{g_3(\phi)v\cdot\nabla\theta; \mathcal{Q}_{\varepsilon\nu}\big\} - \kappa\,\big[B_3(\phi), \mathcal{Q}_{\varepsilon\nu}\big]\theta,
\end{aligned}$$

where $\phi$ is a shorthand notation for $(\theta, \varepsilon p)$ and

$$\big\{g_3(\phi)v\cdot\nabla\theta; \mathcal{Q}_{\varepsilon\nu}\big\} := g_3(\phi)v\cdot\nabla\mathcal{Q}_{\varepsilon\nu}\theta + g_3(\phi)\nabla\theta\cdot\mathcal{Q}_{\varepsilon\nu}v - \mathcal{Q}_{\varepsilon\nu}\big(g_3(\phi)v\cdot\nabla\theta\big).$$

**Remark 5.10.** For the purpose of proving estimates independent of $\kappa$, we do not consider the exact commutator of the third equation in (2.1) with $\mathcal{Q}_{\varepsilon\nu}$.

Let us recall that

$$\begin{aligned}
B_1(\phi) &:= \chi_1(\varepsilon p)\operatorname{div}(\beta(\theta)\nabla\,\cdot\,),\\
B_2(\phi) &:= \chi_2(\varepsilon p)\operatorname{div}(2\zeta(\theta)D\cdot) + \chi_2(\varepsilon p)\nabla(\eta(\theta)\operatorname{div}\cdot),\\
B_3(\phi) &:= \chi_3(\varepsilon p)\operatorname{div}(\beta(\theta)\nabla\cdot).
\end{aligned}$$

The following lemma shows that $f^{a,\nu}_{\rm HF}$ can be seen as a source term.

**Lemma 5.11.** *There exists a smooth nondecreasing function $C = C(\cdot)$ such that for all $a \in A$, all $T \geqslant 0$, all $\nu \in [0, 2]$ and all vector valued function $U := (p, v, \theta) \in C^1([0, T]; H^\infty(\mathbb{D}))$,*

$$\begin{aligned}\left\| f^{a,\nu}_{1,\rm HF}(U)\right\|_{H^1_{\varepsilon\nu}} &\leqslant C(R)\big\{1 + \|\varepsilon\partial_t p\|_{H^s_\nu} + \kappa\|\theta\|_{H^{s+2}}\big\},\\ \left\| f^{a,\nu}_{2,\rm HF}(U)\right\|_{H^1_{\varepsilon\nu}} &\leqslant C(R)\big\{1 + \|\varepsilon\partial_t v\|_{H^s_\nu} + \mu\|\varepsilon v\|_{H^{s+2}}\big\},\\ \left\| f^{a,\nu}_{3,\rm HF}(U)\right\|_{H^1_{\nu}} &\leqslant C(R)\big\{1 + \|\partial_t \theta\|_{H^s_\nu} + \kappa\|\theta\|_{H^{s+2}}\big\},\end{aligned}$$

*where $R := \|(p, v)\|_{H^{s+1}_{\varepsilon\nu}} + \|\theta\|_{H^{s+1}_\nu}$.*

**Remark 5.12.** We will apply this lemma with $\nu := \sqrt{\mu+\kappa}$. Yet, we prove estimates independent of $\nu \in [0, 2]$ to explain why the separation of the high and low frequency components occurs at frequencies of order of $1/(\varepsilon\sqrt{\mu+\kappa})$.

*Proof.* We make intensive use of the following obvious observations. Firstly, using the Sobolev embedding Theorem and the very definition of the norms $\|\cdot\|_{H^\sigma_\varrho}$, we have

$$\|u\|_{L^\infty} \lesssim \|u\|_{H^{s-1}} \leqslant \|u\|_{H^s_\nu} \quad \text{and} \quad \|u\|_{W^{1,\infty}} \lesssim \|u\|_{H^s} \leqslant \|u\|_{H^{s+1}_\nu}. \tag{5.12}$$

Since $\varepsilon \leqslant 1$ and $\nu \leqslant 2$, directly from the definition of $\|\cdot\|_{H^\sigma_\nu}$, we have

$$\nu\|u\|_{H^\sigma} \leqslant \|u\|_{H^\sigma_\nu} \leqslant 3\|u\|_{H^\sigma} \quad \text{and} \quad \|\varepsilon u\|_{H^{s+1}_\nu} \leqslant \|u\|_{H^{s+1}_{\varepsilon\nu}}. \tag{5.13}$$

Note the following corollary of the second inequality: $\|(\varepsilon p, \varepsilon v)\|_{H^{s+1}_\nu} \leqslant R$. This in turn implies $\|\phi\|_{H^{s+1}_\nu} := \|(\theta, \varepsilon p)\|_{H^{s+1}_\nu} \leqslant R$.

Step 1: Estimate for $f^{a,\nu}_{1,\rm HF}(\mathrm{U})$. **a)** We begin by proving that

$$\left\|\left[g_1(\phi), \mathcal{Q}_{\varepsilon\nu}\right]\partial_t p\right\|_{H^1_{\varepsilon\nu}} \leqslant C(R)\|\varepsilon\partial_t p\|_{H^s_\nu}. \tag{5.14}$$

Applying the commutator estimate (5.2) with $h = \varrho = \varepsilon\nu$, we have

$$\begin{aligned}\left\|\left[g_1(\phi), \mathcal{Q}_{\varepsilon\nu}\right]\partial_t p\right\|_{H^1_{\varepsilon\nu}} &\lesssim \varepsilon\nu\|\tilde g_1(\phi)\|_{W^{1,\infty}}\|\partial_t p\|_{H^s} + \varepsilon\nu\|\tilde g_1(\phi)\|_{H^{s+1}}\|\partial_t p\|_{L^\infty}\\ &\lesssim \|\tilde g_1(\phi)\|_{W^{1,\infty}}\|\varepsilon\partial_t p\|_{H^s_\nu} + \|\tilde g_1(\phi)\|_{H^{s+1}_\nu}\|\varepsilon\partial_t p\|_{L^\infty},\end{aligned}$$

where $\tilde g_1$ is defined by $\tilde g_1 = g_1 - g_1(0)$. The estimates (5.12) imply that the right side is less than $K\|\tilde g_1(\phi)\|_{H^{s+1}_\nu}\|\varepsilon\partial_t p\|_{H^s_\nu}$. Using Lemma 5.6, we obtain $\|\tilde g_1(\phi)\|_{H^{s+1}_\nu} \leqslant C(\|\phi\|_{H^{s+1}_\nu}) \leqslant C(R)$. This proves (5.14).

**b)** Next, we prove that

$$\left\|\left[g_1(\phi)v, \mathcal{Q}_{\varepsilon\nu}\right]\cdot\nabla p\right\|_{H^1_{\varepsilon\nu}} \leqslant C(R). \tag{5.15}$$

Again, this follows from the commutator estimate (5.2) applied with $h = \varrho = \varepsilon\nu$. Indeed, it yields

$$\begin{aligned}&\left\|\left[g_1(\phi)v, \mathcal{Q}_{\varepsilon\nu}\right]\cdot\nabla p\right\|_{H^1_{\varepsilon\nu}}\\ &\qquad\lesssim \|\tilde g_1(\phi)v\|_{W^{1,\infty}}\|\varepsilon\nabla p\|_{H^s_\nu} + \|\varepsilon\tilde g_1(\phi)v\|_{H^{s+1}_\nu}\|\nabla p\|_{L^\infty}.\end{aligned}$$

The first term in the right-hand side is estimated by $C(R)$ since

$$\|\tilde{g}_1(\phi)v\|_{W^{1,\infty}} \leqslant C\big(\|\phi\|_{W^{1,\infty}}, \|v\|_{W^{1,\infty}}\big) \leqslant C(\|\phi\|_{H^s}, \|v\|_{H^s}) \leqslant C(R),$$

and $\|\varepsilon\nabla p\|_{H^s_\nu} \leqslant \|p\|_{H^{s+1}_{\varepsilon\nu}} \leqslant R$.

It thus remains to estimate $\|\varepsilon\tilde{g}_1(\phi)v\|_{H^{s+1}_\nu}$. To do so use Lemma 5.5 and 5.6, to obtain

$$\|\varepsilon\tilde{g}_1(\phi)v\|_{H^{s+1}_\nu} \lesssim \|\tilde{g}_1(\phi)\|_{H^{s+1}_\nu} \|\varepsilon v\|_{H^{s+1}_\nu} \leqslant C(R)R.$$

**c)** Let us prove that

$$\frac{\kappa}{\varepsilon}\left\|\big[B_1(\phi), \mathcal{Q}_{\varepsilon\nu}\big]\theta\right\|_{H^1_{\varepsilon\nu}} \leqslant C(R)\|\kappa\theta\|_{H^{s+2}}. \tag{5.16}$$

By definition $B_1(\phi) = \chi_1(\varepsilon p)\operatorname{div}(\beta(\theta)\nabla\,\cdot\,)$, so we can decompose the commutator $\big[B_1(\phi), \mathcal{Q}_{\varepsilon\nu}\big]\theta$ as

$$\big[\mathcal{E}_1(\phi), \mathcal{Q}_{\varepsilon\nu}\big]\Delta\theta + \big[\mathcal{E}_2(\phi, \nabla\phi), \mathcal{Q}_{\varepsilon\nu}\big]\cdot\nabla\theta, \tag{5.17}$$

where $\mathcal{E}_1(\phi) = \chi_1(\varepsilon p)\beta(\theta)$ and $\mathcal{E}_2(\phi,\nabla\phi) = \chi_1(\varepsilon p)\nabla\beta(\theta) = \chi_1(\varepsilon p)\beta'(\theta)\nabla\theta$.

The commutator estimate (5.2) [applied with $h = \varrho = \varepsilon\nu$] implies that

$$\left\|\big[\mathcal{E}_1, \mathcal{Q}_{\varepsilon\nu}\big]\Delta\theta\right\|_{H^1_{\varepsilon\nu}} \lesssim \varepsilon\|\tilde{\mathcal{E}}_1\|_{W^{1,\infty}}\|\Delta\theta\|_{H^s_\nu} + \varepsilon\|\tilde{\mathcal{E}}_1\|_{H^{s+1}_\nu}\|\Delta\theta\|_{L^\infty},$$
$$\left\|\big[\mathcal{E}_2, \mathcal{Q}_{\varepsilon\nu}\big]\cdot\nabla\theta\right\|_{H^1_{\varepsilon\nu}} \lesssim \varepsilon\|\nu\mathcal{E}_2\|_{W^{1,\infty}}\|\nabla\theta\|_{H^s} + \varepsilon\|\mathcal{E}_2\|_{H^{s+1}_\nu}\|\nabla\theta\|_{L^\infty},$$

where $\tilde{\mathcal{E}}_1 := \mathcal{E}_1 - \mathcal{E}_1(0)$. Claim (5.16) then easily follows from the estimates:

$$\begin{aligned}
&\|\tilde{\mathcal{E}}_1(\phi)\|_{W^{1,\infty}} \leqslant C(\|\phi\|_{H^s}) \leqslant C(R), && \|\Delta\theta\|_{H^s_\nu} \lesssim \|\theta\|_{H^{s+2}},\\
&\|\tilde{\mathcal{E}}_1(\phi)\|_{H^{s+1}_\nu} \leqslant C(\|\phi\|_{H^{s+1}_\nu}) \leqslant C(R), && \|\Delta\theta\|_{L^\infty} \lesssim \|\theta\|_{H^{s+2}},\\
&\|\nu\mathcal{E}_2(\phi,\nabla\phi)\|_{W^{1,\infty}} \leqslant C(\|\phi\|_{H^{s+1}_\nu}) \leqslant C(R), && \|\nabla\theta\|_{H^s} \leqslant \|\theta\|_{H^{s+2}},\\
&\|\mathcal{E}_2(\phi,\nabla\phi)\|_{H^{s+1}_\nu} \leqslant C(R)\|\theta\|_{H^{s+2}}, && \|\nabla\theta\|_{L^\infty} \lesssim \|\theta\|_{H^s} \leqslant R.
\end{aligned}$$

The first six inequalities follows from Lemma 5.6, the Sobolev Theorem and the estimates (5.12)–(5.13). To estimate $\|\tilde{\mathcal{E}}_2(\phi,\nabla\phi)\|_{H^{s+1}_\nu}$, we first use Lemma 5.5, to obtain

$$\|\mathcal{E}_2(\phi,\nabla\phi)\|_{H^{s+1}_\nu} \lesssim (1 + \|\tilde{\chi_1}(\varepsilon p)\|_{H^{s+1}_\nu})(1 + \|\tilde{\beta}'(\theta)\|_{H^{s+1}_\nu})\|\nabla\theta\|_{H^{s+1}_\nu}, \tag{5.18}$$

and next use Lemma 5.6 to bound the $H^{s+1}_\nu$-norms of $\tilde{\chi_1}(\varepsilon\theta)$ and $\tilde{\beta}'(\theta)$ by $C(R)$. This yields the desired bound since $\|\nabla\theta\|_{H^{s+1}_\nu} \lesssim \|\theta\|_{H^{s+2}}$.

Step 2: Estimate for $f^{a,\nu}_{2,\mathrm{HF}}(U)$. Note that it is possible to obtain $f^{a,\nu}_{2,\mathrm{HF}}(U)$ from $f^{a,\nu}_{1,\mathrm{HF}}(U)$ by replacing $p$ by $v$, $\theta$ by $\varepsilon v$ and $\kappa$ by $\mu$. Therefore, we are back in the situation of the previous step, and hence conclude that

$$\begin{aligned}
\left\|\big[g_2(\phi), \mathcal{Q}_{\varepsilon\nu}\big]\partial_t v\right\|_{H^1_{\varepsilon\nu}} &\leqslant C(R)\|\varepsilon\partial_t v\|_{H^s_\nu},\\
\left\|\big[g_2(\phi)v, \mathcal{Q}_{\varepsilon\nu}\big]\cdot\nabla v\right\|_{H^1_{\varepsilon\nu}} &\leqslant C(R),\\
\mu\left\|\big[B_2(\phi), \mathcal{Q}_{\varepsilon\nu}\big]v\right\|_{H^1_{\varepsilon\nu}} &\leqslant C(R)\|\mu\varepsilon v\|_{H^{s+2}}.
\end{aligned}$$

Step 3: Estimate for $f^{a,\nu}_{3,\mathrm{HF}}(U)$. The estimates for the first and the last terms in $f^{a,\nu}_{3,\mathrm{HF}}(U)$ can be deduced by following the previous analysis. We only indicate the point at which the argument differs: we use the commutator estimate (5.5) with $h = \varepsilon\nu$ and $\varrho = \nu$ (instead of $\varrho = \varepsilon\nu$).

Let us concentrate on the second term. We claim that

$$\big\| \big\{ g_3(\phi) v \cdot \nabla\theta ; \mathcal{Q}_{\varepsilon\nu} \big\} \big\|_{H^1_\nu} \leqslant C(R).$$

We first decompose $\big\{ g_3(\phi) v \cdot \nabla\theta ; \mathcal{Q}_{\varepsilon\nu} \big\}$ as

$$\big[ g_3(\phi), \mathcal{Q}_{\varepsilon\nu} \big] v \cdot \nabla\theta + g_3(\phi) Z \tag{5.19}$$

where $Z := v \cdot \nabla \mathcal{Q}_{\varepsilon\nu}\theta + \nabla\theta \cdot \mathcal{Q}_{\varepsilon\nu} v - \mathcal{Q}_{\varepsilon\nu}(v \cdot \nabla\theta)$.

Applying the commutator estimate (5.2) with $h = \varepsilon\nu$ and $\varrho = \nu$, and using the estimates (5.12), we get

$$\big\| \big[ g_3(\phi), \mathcal{Q}_{\varepsilon\nu} \big] v \cdot \nabla\theta \big\|_{H^1_\nu} \lesssim \| \tilde{g}_3(\phi) \|_{H^{s+1}_\nu} \| v \cdot \nabla\theta \|_{H^s_\nu} .$$

Using Lemma 5.5 and 5.6, we find that the right-hand side of the previous estimate is dominated by $C(\|\phi\|_{H^{s+1}_\nu}) \|v\|_{H^s_\nu} \|\nabla\theta\|_{H^s_\nu} \leqslant C(R)$.

Since $\|g_3(\phi) Z\|_{H^1_\nu} \leqslant \|g_3(\phi)\|_{W^{1,\infty}} \|Z\|_{H^1_\nu} \leqslant C(R) \|Z\|_{H^1_\nu}$, to control the $H^1_\nu$-norm of the second term in (5.19), only the estimate of $\|Z\|_{H^1_\nu}$ is missing. We split $\|Z\|_{H^1_\nu}$ as $\|Z\|_{L^2} + \nu \|Z\|_{H^1}$, and we prove that $\|Z\|_{L^2}$ and $\nu \|Z\|_{H^1}$ are both estimated by $C(R)$. To estimate $\|Z\|_{L^2}$, note that

$$Z = \nabla\theta \cdot \mathcal{Q}_{\varepsilon\nu} v + [v, \mathcal{Q}_{\varepsilon\nu}] \cdot \nabla\theta. \tag{5.20}$$

The second term in the right-hand side is estimated by way of the commutator estimate (3.9). Indeed, applying Lemma 3.3 with $\sigma_0 = m = s$ and $\sigma = s - 1$, we get

$$\| [v, \mathcal{Q}_{\varepsilon\nu}] \cdot \nabla\theta \|_{L^2} \lesssim \|v\|_{H^s} \|\nabla\theta\|_{H^{s-1}} \leqslant \|v\|_{H^s} \|\theta\|_{H^s} \leqslant R^2.$$

With regards to the first term in the right-hand side of (5.20), we write

$$\|\nabla\theta \cdot \mathcal{Q}_{\varepsilon\nu} v\|_{L^2} \leqslant \|\nabla\theta\|_{L^\infty} \|\mathcal{Q}_{\varepsilon\nu} v\|_{L^2} \lesssim \|\theta\|_{H^s} \|v\|_{H^s} \leqslant R^2.$$

Moving to the estimate of $\nu \|Z\|_{H^1}$, we remark that

$$\nu Z = F'(u) \mathcal{Q}_{\varepsilon\nu} u - \mathcal{Q}_{\varepsilon\nu}(F(u)),$$

with $u = (u_1, u_2) := (v, \nu\nabla\theta)$ and $F(u) = u_1 u_2$. Hence, (5.10) yields

$$\nu \|Z\|_{H^1} \leqslant C\big( \|(v, \nu\nabla\theta)\|_{H^s} \big) \leqslant C\big( \|v\|_{H^s} + \|\theta\|_{H^{s+1}_\nu} \big) \leqslant C(R).$$

The previous estimates imply that $\|Z\|_{H^1_\nu}$ is controlled by $C(R)$. This completes the proof of Lemma 5.11. □

**Remark 5.13.** Let us explain the reason why we assume that $\beta$ depends only on $\theta$. Had we worked instead with general coefficient $\beta$ depending also on $\varepsilon p$, the corresponding inequality (5.18) would have involved $\|\varepsilon\nabla p\|_{H^{s+1}_\nu}$. The problem presents itself: $\|\varepsilon p\|_{L^1(0,T;H^{s+2}_\nu)}$ is not controlled by the norm $\|(p, v, \theta)\|_{\mathcal{H}^s_a(T)}$. It is possible to get around the previous problem, yet we do not address this question.

In view of Lemma 5.11 we are led to estimate $\partial_t \psi$ where $\psi := (\theta, \varepsilon p, \varepsilon v)$.

**Lemma 5.14.** *There exists a continuous nondecreasing function $C(\cdot)$ such that for all $a \in A$, all $T \geqslant 0$, all $\nu \in [0,2]$ and all $(p,v,\theta) \in \mathcal{H}^s_a(T)$ solving* (2.1)*, the function $\psi := (\theta, \varepsilon p, \varepsilon v)$ satisfies*

$$\|\partial_t \psi\|_{H^s_\nu} \leqslant C(R)\big\{1 + R'\big\}, \tag{5.21}$$

*where*

$$\begin{aligned} R &:= \|(p,v)\|_{H^{s+1}_{\varepsilon\nu}} + \|\theta\|_{H^{s+1}_{\nu}}\,, \\ R' &:= \nu\,\|(\operatorname{div} v, \nabla p)\|_{H^s} + \sqrt{\mu}\,\|\nabla v\|_{H^{s+1}_{\varepsilon\nu}} + \sqrt{\kappa}\,\|\nabla\theta\|_{H^{s+1}_{\nu}}\,. \end{aligned} \tag{5.22}$$

**Remark 5.15.** This estimate plays a key role in proving estimates independent of $\mu$ and $\kappa$. Indeed, this is the only step in which we use the additional smoothing effect for $\operatorname{div} v$ and $\nabla p$. More precisely, the fact that the estimate (5.21) is tame [linear in $R'$] allows us to control $\|\partial_t\psi\|_{L^2(0,T;H^s_\nu)}$ by $C\big(\|(p,v,\theta)\|_{\mathcal{H}^s_a(T)}\big)$ for all $\nu \leqslant \sqrt{\mu+\kappa}$.

*Proof.* The proof is straightforward. We write

$$G(\phi)\big(\partial_t\psi + v\cdot\nabla\psi\big) + LU + B_{\mu,\kappa}(\phi)\psi = 0, \tag{5.23}$$

where $\psi$ and $U$ are identified with $(\varepsilon p, \varepsilon v, \theta)^t$ and $(p,v,\theta)^t$, respectively, and

$$G := \begin{pmatrix} g_1 & 0 & 0 \\ 0 & g_2 I_d & 0 \\ 0 & 0 & g_3 \end{pmatrix},\ L := \begin{pmatrix} 0 & \operatorname{div} & 0 \\ \nabla & 0 & 0 \\ 0 & \operatorname{div} & 0 \end{pmatrix},\ B_{\mu,\kappa} := \begin{pmatrix} 0 & 0 & \kappa B_1 \\ 0 & \mu B_2 & 0 \\ 0 & 0 & \kappa B_3 \end{pmatrix}.$$

Since $\max\{\mu,\kappa\} \leqslant 1$, we can easily verify that there is a family $\{F_a \,|\, a \in A\}$ uniformly bounded in $C^\infty$, such that $F_a(0) = 0$ and $\partial_t\psi = F_a(\Xi)$, with $\Xi := \big(v, \psi, \nabla\psi, \operatorname{div} v, \nabla p, \varepsilon\mu\nabla^2 v, \kappa\nabla^2\theta\big)$. Using the Moser-type estimate (5.9), we find that $\|\partial_t\psi\|_{H^s_\nu} \leqslant C(\|\Xi\|_{H^{s-1}})\{1 + \nu\,\|\Xi\|_{H^s}\}$. To complete the proof, note that, directly from the definitions (5.22) and the assumption $(\mu,\kappa) \in [0,1]^2$, we have $\|\Xi\|_{H^{s-1}} \leqslant R$ and $\nu\,\|\Xi\|_{H^s} \leqslant R + R'$. □

We are now prepared to prove Proposition 5.1.

5.3. **Proof of Proposition 5.1.** Set $\nu := \sqrt{\mu+\kappa} \leqslant \sqrt{2}$. We first show that

$$\widetilde{U} := (\widetilde{p}, \widetilde{v}, \widetilde{\theta}) := \big(\mathcal{Q}_{\varepsilon\nu} p, \mathcal{Q}_{\varepsilon\nu} v, \mathcal{Q}_{\varepsilon\nu}\theta\big),$$

satisfies (4.4) for suitable source terms $f_1$, $f_2$, $f_3$ and coefficients $\beta_1$, $\beta_2$, $\beta_2^\sharp$, $\beta_3$, $G$. It readily follows from Notation 5.9 that $(\widetilde{p}, \widetilde{v}, \widetilde{\theta})$ solves

$$\begin{cases} g_1(\phi)\big(\partial_t\widetilde{p} + v\cdot\nabla\widetilde{p}\big) + \dfrac{1}{\varepsilon}\operatorname{div}\widetilde{v} - \dfrac{\kappa}{\varepsilon}B_1(\phi)\widetilde{\theta} = f^{a,\nu}_{1,\mathrm{HF}}(U), \\ g_2(\phi)\big(\partial_t\widetilde{v} + v\cdot\nabla\widetilde{v}\big) + \dfrac{1}{\varepsilon}\nabla\widetilde{p} - \mu B_2(\phi)\widetilde{v} = f^{a,\nu}_{2,\mathrm{HF}}(U), \\ g_3(\phi)\big(\partial_t\widetilde{\theta} + v\cdot\nabla\widetilde{\theta} + \widetilde{v}\cdot\nabla\theta\big) + \operatorname{div}\widetilde{v} - \kappa B_3(\phi)\widetilde{\theta} = f^{a,\nu}_{3,\mathrm{HF}}(U), \end{cases}$$

where $\phi = (\theta, \varepsilon p)$. Set $\beta_1 := \chi_1 \beta$, $\beta_2 := 2\chi_2\zeta$, $\beta_2^\sharp := \chi_2\eta$, $\beta_3 := \chi_3\beta$ and $G(\phi, \nabla\phi) = g_3(\phi)\nabla\theta$, to obtain that $(\widetilde{p}, \widetilde{v}, \widetilde{\theta})$ satisfies (4.4) where

$$
\begin{aligned}
f_1 &:= f_{1,\mathrm{HF}}^{a,\nu}(U) + f_{1,\mathrm{HF}}' + \mathcal{Q}_{\varepsilon\nu}\Upsilon_1 \quad \text{with } f_{1,\mathrm{HF}}' := -\frac{\kappa}{\varepsilon}\nabla\chi_1(\varepsilon p)\cdot\big(\beta(\theta)\nabla\widetilde{\theta}\big),\\
f_2 &:= f_{2,\mathrm{HF}}^{a,\nu}(U) + f_{2,\mathrm{HF}}' \quad \text{with } f_{2,\mathrm{HF}}' := \mu\chi_2(\varepsilon p)\big\{2D\widetilde{v}\nabla\zeta(\theta) + \operatorname{div}\widetilde{v}\nabla\eta(\theta)\big\},\\
f_3 &:= f_{3,\mathrm{HF}}^{a,\nu}(U) + f_{3,\mathrm{HF}}' + \varepsilon\mathcal{Q}_{\varepsilon\nu}\Upsilon_3 \quad \text{with } f_{3,\mathrm{HF}}' := \kappa\chi_3(\varepsilon p)\nabla\beta(\theta)\cdot\nabla\widetilde{\theta},
\end{aligned}
$$

where $\Upsilon_i$ are as in system (2.1).

By definition of the norm $\|\cdot\|_{\mathcal{H}_a^s(T)}$ (see Definition 2.4), the Sobolev embedding Theorem and Lemma 5.14 imply that

$$\|(\phi, \partial_t\phi, \nabla\phi, \nu\nabla^2\phi, v, \nabla v)\|_{L^\infty([0,T]\times\mathbb{D})} \leqslant C(\Omega),$$

where $\Omega := \|U\|_{\mathcal{H}_a^s(T)}$. Similarly, $\|\phi(0)\|_{L^\infty(\mathbb{D})} \leqslant \Omega_0 := \|U(0)\|_{\mathcal{H}_{a,0}^s}$.

Observe that Assumption 2.1 implies that the conditions in Assumption 4.1 are satisfied.

With these preliminary remarks in hands, Theorem 4.5 yields

$$\|\widetilde{U}\|_{\mathcal{H}_a^0(T)} \leqslant C(\Omega_0)e^{TC(\Omega)}\|\widetilde{U}(0)\|_{\mathcal{H}_{a,0}^0} + C(\Omega)\big\{\mathfrak{F}(T) + \mathfrak{F}'(T) + \mathfrak{F}''(T)\big\},$$

with

$$
\begin{aligned}
\mathfrak{F}(T) &:= \big\|(f_{1,\mathrm{HF}}^{a,\nu}(U), f_{2,\mathrm{HF}}^{a,\nu}(U))\big\|_{L^1(0,T;H^1_{\varepsilon\nu})} + \big\|f_{3,\mathrm{HF}}^{a,\nu}(U)\big\|_{L^1(0,T;H^1_\nu)},\\
\mathfrak{F}'(T) &:= \big\|(f_{1,\mathrm{HF}}', f_{2,\mathrm{HF}}')\big\|_{L^1(0,T;H^1_{\varepsilon\nu})} + \big\|f_{3,\mathrm{HF}}'\big\|_{L^1(0,T;H^1_\nu)},\\
\mathfrak{F}''(T) &:= \big\|\mathcal{Q}_{\varepsilon\nu}\Upsilon_1\big\|_{L^1(0,T;H^1_{\varepsilon\nu})} + \big\|\varepsilon\mathcal{Q}_{\varepsilon\nu}\Upsilon_3\|_{L^1(0,T;H^1_\nu)}.
\end{aligned}
$$

By definition, $\|\widetilde{U}\|_{\mathcal{H}_a^0(T)} = \|(I - J_{\varepsilon\nu})U\|_{\mathcal{H}_a^s(T)}$. Similarly, we have the bound $\|\widetilde{U}(0)\|_{\mathcal{H}_{a,0}^0} = \|(I - J_{\varepsilon\nu})U(0)\|_{\mathcal{H}_{a,0}^s} \leqslant \Omega_0$. Therefore, in view of the elementary inequality $x + y \leqslant 2xe^y$ (for all $x \geqslant 1$ and $y \geqslant 0$), the proof of Proposition 5.1 reduces to establishing that $\mathfrak{F}(T)+\mathfrak{F}'(T)+\mathfrak{F}''(T) \leqslant \sqrt{T}C(\Omega)$.

The estimate $\mathfrak{F}(T) \leqslant \sqrt{T}C(\Omega)$ immediately follows from the preliminaries. Indeed, define $R$ and $R'$ as in (5.22). By the definition of $\nu := \sqrt{\mu+\kappa}$, we have $\mu \leqslant \sqrt{\mu}\nu$ and $\kappa \leqslant \sqrt{\kappa}\nu$. Hence, it can be easily verified that Lemma 5.11 and Lemma 5.14 imply

$$\big\|f_{1,\mathrm{HF}}^{a,\nu}(U)\big\|_{H^1_{\varepsilon\nu}} + \big\|f_{2,\mathrm{HF}}^{a,\nu}(U)\big\|_{H^1_{\varepsilon\nu}} + \big\|f_{3,\mathrm{HF}}^{a,\nu}(U)\big\|_{H^1_\nu} \leqslant C(R)\big\{1 + R'\big\}. \tag{5.24}$$

Since $C(\cdot)$ is nondecreasing, by integrating and using the Cauchy–Schwarz estimate, we obtain

$$\mathfrak{F}(T) \leqslant \sqrt{T}C\big(\|R\|_{L^\infty(0,T)}\big)\big\{1 + \|R'\|_{L^2(0,T)}\big\}.$$

This in turn implies the desired estimate $\mathfrak{F}(T) \leqslant \sqrt{T}C(\Omega)$ since, by definition, we have $\Omega \approx \|R\|_{L^\infty(0,T)} + \|R'\|_{L^2(0,T)}$.

Let us prove that, similarly, $\mathfrak{F}'(T) \leqslant \sqrt{T}C(\Omega)$. To do that it is sufficient to prove that (5.24) holds true with $f_{i,\mathrm{HF}}^{a,\nu}(U)$ replaced by $f_{i,\mathrm{HF}}'$. This in turn follows from direct estimates. Indeed, observe that

$$f'_{1,\mathrm{HF}} = \kappa \mathcal{F}_1(\phi, \nabla p)\nabla\widetilde{\theta},\ f'_{2,\mathrm{HF}} = \mu\mathcal{F}_2(\phi, \nabla\theta)\nabla\widetilde{v},\ f'_{3,\mathrm{HF}} = \kappa\mathcal{F}_3(\phi, \nabla\theta)\nabla\widetilde{\theta},$$

for some $C^\infty$ functions $\mathcal{F}_i$ vanishing at the origin. As already seen, one can give estimates for the coefficients $\mathcal{F}_i$ by combining the Sobolev embedding Theorem with the Moser estimate (5.7). It is found that

$$\left\|(f'_{1,\mathrm{HF}}, f'_{2,\mathrm{HF}})\right\|_{H^1_{\varepsilon\nu}} + \left\|f'_{3,\mathrm{HF}}\right\|_{H^1_\nu} \leqslant C(R)\{\|\sqrt{\kappa}\nabla\widetilde{\theta}\|_{H^1_\nu} + \|\sqrt{\mu}\nabla\widetilde{v}\|_{H^1_{\varepsilon\nu}}\}. \tag{5.25}$$

We next use $\mathcal{Q}_{\varepsilon\nu} \leqslant \Lambda^s$, to obtain

$$\|\sqrt{\kappa}\nabla\widetilde{\theta}\|_{H^1_\nu} + \|\sqrt{\mu}\nabla\widetilde{v}\|_{H^1_{\varepsilon\nu}} \leqslant \sqrt{\kappa}\,\|\theta\|_{H^{s+2}_\nu} + \sqrt{\mu}\,\|v\|_{H^{s+2}_{\varepsilon\nu}} \leqslant R'.$$

Consequently, the left-hand side of (5.25) is controlled by $C(R)R'$.

To conclude the proof it remains to show that $\mathfrak{F}''(T) \leqslant \sqrt{T}C(\Omega)$. This is nothing new in that it follows from the estimates:

$$\begin{aligned}
\mathfrak{F}''(T) &\lesssim \sqrt{T}\,\|(\Upsilon_1, \Upsilon_3)\|_{L^2(0,T;H^{s+1}_{\varepsilon\nu})} \\
&\leqslant \sqrt{T}C\big(\|(\phi, \sqrt{\mu}\nabla v)\|_{L^\infty([0,T]\times\mathbb{D})}\big)\big\{1 + \|(\phi, \sqrt{\mu}\nabla v)\|_{L^2(0,T;H^{s+1}_{\varepsilon\nu})}\big\} \\
&\leqslant \sqrt{T}C(\|R\|_{L^\infty(0,T)})\{1 + \|R'\|_{L^2(0,T)}\} \\
&\leqslant \sqrt{T}C(\Omega).
\end{aligned}$$

We have proved Proposition 5.1. This completes the analysis of the high frequency regime.

## 6. Low frequency regime

This section is devoted to the proof of *a priori* estimates in the low frequency region, which is the most delicate part.

We prove the following estimates.

**Proposition 6.1.** *Given an integer $s > 1 + d/2$, there exists a continuous nondecreasing function $C$ such that for all $a = (\varepsilon, \mu, \kappa) \in A$, all $T \in [0, 1]$ and all $U = (p, v, \theta) \in C^1([0, T]; H^\infty(\mathbb{D}))$ satisfying* (2.1),

$$\|J_{\varepsilon\nu}U\|_{\mathcal{H}^s_a(T)} \leqslant C(\Omega_0)e^{(\sqrt{T}+\varepsilon)C(\Omega)}, \tag{6.1}$$

*where $\nu := \sqrt{\mu + \kappa}$, $\Omega_0 := \|U(0)\|_{\mathcal{H}^s_{a,0}}$, $\Omega := \|U\|_{\mathcal{H}^s_a(T)}$ (see Definition 2.4).*

As alluded to previously, the nonlinear energy estimates cannot be obtained from the $L^2$ estimates by an elementary argument using differentiation of the equations with respect to spatial derivatives. For such problems a general strategy can be used. First we apply to the equations some operators based on $(\varepsilon\partial_t)$. Next, one uses the special structure of the equations to estimate the spatial derivatives.

This basic strategy has many roots, at least for hyperbolic problems (see, e.g., [1, 18, 39, 40]). For our purposes, the key point is that the hyperbolic behavior prevails in the low frequency regime. Yet, in sharp contrast with the Euler equations ($\mu = \kappa = 0$), the form of the equations (2.1) shows that the time derivative and the spatial derivatives do not have the same weight. In particular, our analysis requires some preparation.

We begin our discussion in §6.1 by establishing some estimates which allows us to commute $J_{\varepsilon\nu}(\varepsilon\partial_t)^m$ with the equations. The latter task is achieved in §6.3. With these preliminaries established, we can proceed to give an estimate for $J_{\varepsilon\nu}(\varepsilon\partial_t)^s U$. The fast components $J_{\varepsilon\nu}(\operatorname{div} v, \nabla p)$ are estimated next by using an induction argument. To conclude, we give the estimates for the slow components $\theta$ and $\operatorname{curl} v$.

For the sake of notational clarity, in this section we deliberately omit the terms $\Upsilon_1$ and $\varepsilon\Upsilon_3$ in the system (2.1). Nothing is changed in the statements of the results, nor in their proofs.

6.1. **Non-isotropic estimates.** The fact that the time derivative and the spatial derivatives do not have the same weight is made precise by the following lemma (whose easy proof is left to the reader).

**Lemma 6.2.** *There is a family* $\{ B_{a,\alpha} \,|\, a \in A,\ \alpha \in \mathbb{N}^d,\ 1 \leqslant |\alpha| \leqslant 2 \}$ *uniformly bounded in* $C^\infty(\mathbb{R}^N; \mathbb{R}^{N\times N})$ *(where* $N = (d+2)^2$*) such that for all* $a \in A$ *and all smooth solution* $(p, v, \theta)$ *of* (2.1)*, the function*

$$\Psi := \big(\psi, \partial_t\psi, \nabla\psi\big) \quad \text{where} \quad \psi := (\theta, \varepsilon p, \varepsilon v), \tag{6.2}$$

*solves*

$$\varepsilon\partial_t\Psi = \sum_{1\leqslant j\leqslant d} B_{a,j}(\Psi)\partial_j\Psi + \varepsilon(\mu+\kappa) \sum_{1\leqslant j,k\leqslant d} \partial_j\big(B_{a,jk}(\Psi)\partial_k\Psi\big). \tag{6.3}$$

We want to introduce an operator based on $(\varepsilon\partial_t)$ which has the weight of a spatial derivative. The previous result suggests introducing the following family of operators.

**Definition 6.3.** *For all* $\varepsilon \geqslant 0$, $\nu \geqslant 0$ *and* $\ell \in \mathbb{N}$*, define the operators*

$$Z^\ell_{\varepsilon,\nu} := \Lambda^{-\ell}_{\varepsilon\nu}(\varepsilon\partial_t)^\ell,$$

*where, we recall that* $\Lambda^{-\ell}_{\varepsilon\nu} := \big(I - (\varepsilon\nu)^2\Delta\big)^{-\ell/2}$.

We will need the following technical ingredient.

**Lemma 6.4.** *Given* $F \in C^\infty(\mathbb{R}^n)$ *satisfying* $F(0) = 0$ *and* $\sigma_0 > d/2$*, there exists a function* $C(\cdot)$ *such that for all* $\varepsilon \in [0,1]$*, all* $\nu \in [0,2]$*, all* $T > 0$*, all vector-valued function* $U \in C^\infty([0,T]; H^\infty(\mathbb{D}))$ *and all* $\mathbb{N} \ni m \leqslant \sigma_0$*,*

$$\big\| Z^m_{\varepsilon,\nu} F(U) \big\|_{H^{\sigma_0 - m}} \leqslant C\Big( \sum_{0\leqslant\ell\leqslant m} \big\| Z^\ell_{\varepsilon,\nu} U \big\|_{H^{\sigma_0-\ell}} \Big). \tag{6.4}$$

*Proof.* To prove this claim, observe that $(\varepsilon\partial_t)^m F(U)$ is a sum of terms of the form

$$f(U)(\varepsilon\partial_t)^{\ell_1}u_1 \cdots (\varepsilon\partial_t)^{\ell_p}u_p,$$

with $p \leqslant m$ and $\sum_{1\leqslant i\leqslant p} \ell_i = m$. In this formula, $f$ is a $C^\infty$ function and $u_1, \ldots, u_p$ denote coefficients of $U$.

In order to estimate these terms, we use the following result (whose proof follows from Proposition 3.2 by induction): let $\alpha = (\alpha_0, , \dots, \alpha_p) \in \mathbb{N}^{p+1}$ be such that $\sum_{i=0}^p \alpha_i =: |\alpha| \leqslant \sigma_0$, then

$$\Big\| \Lambda_{\varepsilon\nu}^{-|\alpha|} \prod_{i=0}^p V_i \Big\|_{H^{\sigma_0 - |\alpha|}} \lesssim \prod_{i=0}^p \big\| \Lambda_{\varepsilon\nu}^{-\alpha_i} V_i \big\|_{H^{\sigma_0-\alpha_i}} ,$$

where the implicit constant depends only on $d$,$\sigma_0$ and $\alpha$.

Set $\widetilde{f} := f - f(0)$. We apply the previous result with $\alpha_0 = 0$, $\alpha_i = \ell_i$ ($i \geqslant 1$), $V_0 = \widetilde{f}(U)$ and $V_i = (\varepsilon\partial_t)^{\ell_i} u_i$ ($i \geqslant 1$). This yields

$$\begin{aligned} &\big\| \Lambda_{\varepsilon\nu}^{-m} \big( f(U)(\varepsilon\partial_t)^{\ell_1} u_1 \cdots (\varepsilon\partial_t)^{\ell_p} u_p \big) \big\|_{H^{\sigma_0-m}} \\ &\qquad \lesssim \big( 1 + \big\| \widetilde{f}(U) \big\|_{H^{\sigma_0}} \big) \big\| Z_{\varepsilon,\nu}^{\ell_1} u_1 \big\|_{H^{\sigma_0-\ell_1}} \cdots \big\| Z_{\varepsilon,\nu}^{\ell_p} u_p \big\|_{H^{\sigma_0-\ell_p}}. \end{aligned}$$

Since $Z_{\varepsilon,\nu}^0 = I$, the estimate (5.8) implies $\|\widetilde{f}(U)\|_{H^{\sigma_0}} \leqslant C\big( \big\| Z_{\varepsilon,\nu}^0 U \big\|_{H^{\sigma_0}} \big)$. Which completes the proof. □

Now we are in position to prove that $Z_{\varepsilon,\nu}^1$ has the weight of a spatial derivative. The following result states that, for $m \in \mathbb{N}$, $Z_{\varepsilon,\nu}^m \Psi$ satisfies the same estimates as $\Lambda^m F(\Psi)$ does (where $F$ is a given function).

**Proposition 6.5.** *Let $s > 1 + d/2$ be an integer. There exists a function $C(\cdot)$ such that for all $a = (\varepsilon, \mu, \kappa) \in A$, all $T > 0$ and all smooth solution $(p, v, \theta) \in C^\infty([0, T]; H^\infty(\mathbb{D}))$ of* (2.1) *, if $\nu \in [(\mu+\kappa)/2, 2]$ then the function $\Psi$ defined by* (6.2) *satisfies*

$$\sum_{0 \leqslant \ell \leqslant s} \big\| Z_{\varepsilon,\nu}^\ell \Psi \big\|_{H^{s-\ell-1}} \leqslant C\big( \|\Psi\|_{H^{s-1}} \big), \tag{6.5}$$

$$\sum_{0 \leqslant \ell \leqslant s} \big\| Z_{\varepsilon,\nu}^\ell \Psi \big\|_{H_\nu^{s-\ell}} \leqslant C\big( \|\Psi\|_{H^{s-1}} \big) \|\Psi\|_{H_\nu^s} . \tag{6.6}$$

*Proof.* We prove by induction on $m \in \{0, \dots, s\}$ that

$$\sum_{0 \leqslant \ell \leqslant m} \big\| Z_{\varepsilon,\nu}^\ell \Psi \big\|_{H^{s-\ell-1}} \leqslant C\big( \|\Psi\|_{H^{s-1}} \big), \tag{6.7}$$

$$\sum_{0 \leqslant \ell \leqslant m} \big\| Z_{\varepsilon,\nu}^\ell \Psi \big\|_{H^{s-\ell}} \leqslant C\big( \|\Psi\|_{H^{s-1}} \big) \|\Psi\|_{H^s} . \tag{6.8}$$

Assume the results (6.7)–(6.8) at order $m < s$ (note that they are obvious with $m = 0$).

By definition, $Z_{\varepsilon,\nu}^{m+1} = \Lambda_{\varepsilon\nu}^{-1} Z_{\varepsilon,\nu}^m (\varepsilon\partial_t)$. It thus follows from (6.3) that

$$Z_{\varepsilon,\nu}^{m+1} \Psi = \sum_{j,k} \Lambda_{\varepsilon\nu}^{-1} Z_{\varepsilon,\nu}^m \big( B_{a,j}(\Psi)\partial_j \Psi \big) + \big( \varepsilon(\mu+\kappa)\partial_j \Lambda_{\varepsilon\nu}^{-1} \big) Z_{\varepsilon,\nu}^m \big( B_{a,jk}(\Psi)\partial_k \Psi \big).$$

Since $\nu \gtrsim \mu + \kappa$ we have $\varepsilon(\mu+\kappa)\partial_j \Lambda_{\varepsilon\nu}^{-1} \lesssim I$ (see (3.3)). Similarly, $\Lambda_{\varepsilon\nu}^{-1} \lesssim I$. The proof thus reduces to estimating terms having the form $Z_{\varepsilon,\nu}^m \big( B(\Psi)\partial_j \Psi \big)$.

More precisely, it is sufficient to prove that, for all integers $m < s$ and for all smooth functions $B$, we have

$$\left\| Z^m_{\varepsilon,\nu} \left( B(\Psi)\partial_j \Psi \right) \right\|_{H^{s-m-2}} \leqslant C\left( \|\Psi\|_{H^{s-1}} \right), \tag{6.9}$$

$$\left\| Z^m_{\varepsilon,\nu} \left( B(\Psi)\partial_j \Psi \right) \right\|_{H^{s-m-1}} \leqslant C\left( \|\Psi\|_{H^{s-1}} \right) \|\Psi\|_{H^s} \,, \tag{6.10}$$

where $C(\cdot)$ depends only on a finite number of semi-norms of $B$ in $C^\infty$.

Firstly, writing $B(\Psi)\partial_j\Psi$ as $\partial_j F(\Psi)$ for some smooth function $F$ such that $F(0) = 0$, and using Lemma 6.4 with $\sigma_0 = s-1$, we find that

$$\begin{aligned} \left\| Z^m_{\varepsilon,\nu} \left( B(\Psi)\partial_j \Psi \right) \right\|_{H^{s-m-2}} &\leqslant \left\| Z^m_{\varepsilon,\nu} F(\Psi) \right\|_{H^{s-1-m}} \\ &\leqslant C\Big( \sum_{\ell=0}^{m} \left\| Z^\ell_{\varepsilon,\nu} \Psi \right\|_{H^{s-1-\ell}} \Big). \end{aligned}$$

As a consequence, the estimate (6.9) follows from the induction hypothesis (6.7). Moving to the proof of (6.10), we begin with the Leibniz rule

$$(\varepsilon\partial_t)^m \big( B(\Psi)\partial_j\Psi \big) = \sum_{0\leqslant \ell\leqslant m} \binom{m}{\ell} (\varepsilon\partial_t)^{m-\ell} B(\Psi) (\varepsilon\partial_t)^\ell \partial_j \Psi.$$

Let $\mathbb{N} \ni \ell \leqslant m$. Since $s > 1 + d/2$ and $m \leqslant s-1$, Proposition 3.2 applies with $\sigma_0 = s-1$, $\sigma_1 = m_1 = m-\ell$ and $\sigma_2 = m_2 = \ell$. It yields

$$\begin{aligned} &\left\| Z^m_{\varepsilon,\nu} \left( B(\Psi)\partial_j \Psi \right) \right\|_{H^{s-m-1}} \\ &\qquad \lesssim \sum_{\ell=0}^{m} \left\| Z^{m-\ell}_{\varepsilon,\nu} B(\Psi) \right\|_{H^{s-1-(m-\ell)}} \left\| Z^\ell_{\varepsilon,\nu} \partial_j \Psi \right\|_{H^{s-1-\ell}} . \end{aligned}$$

Using Lemma 6.4 to estimate the first term in the summand, we get

$$\left\| Z^m_{\varepsilon,\nu} \left( B(\Psi)\partial_j \Psi \right) \right\|_{H^{s-m-1}} \leqslant C\Big( \sum_{\ell=0}^{m} \left\| Z^\ell_{\varepsilon,\nu} \Psi \right\|_{H^{s-\ell-1}} \Big) \sum_{\ell=0}^{m} \left\| Z^\ell_{\varepsilon,\nu} \Psi \right\|_{H^{s-\ell}} .$$

Using the induction hypotheses (6.7)–(6.7), we have proved (6.10). □

We next prove a commutator estimate with gain of a factor $\varepsilon$.

**Lemma 6.6.** *Given $s > 1 + d/2$, there exists a constant $K$ such that for all $\varepsilon \in [0,1]$, all $\nu \in [0,2]$, all $T > 0$, all $m \in \mathbb{N}$ such that $1 \leqslant m \leqslant s$ and all $f, u \in C^\infty([0,T]; H^\infty(\mathbb{D}))$,*

$$\begin{aligned} \left\| [f, J_{\varepsilon\nu}(\varepsilon\partial_t)^m] u \right\|_{H^{s-m+1}_{\varepsilon\nu}} \leqslant K\varepsilon \Big\{ \|f\|_{H^s} + \sum_{\ell=0}^{m-1} \left\| Z^\ell_{\varepsilon,\nu} \partial_t f \right\|_{H^{s-1-\ell}} \Big\} \\ \times \Big\{ \left\| Z^m_{\varepsilon,\nu} u \right\|_{H^{s-m}_\nu} + \sum_{\ell=0}^{m-1} \left\| Z^\ell_{\varepsilon,\nu} u \right\|_{H^{s-1-\ell}} \Big\}. \end{aligned}$$

**Remark 6.7.** This technical ingredient is an analogue of the commutator estimate (5.2) (with $\varrho = h$) which we used in the analysis of the high frequency regime.

*Proof.* The proof is based on the tools we developed in Section 3.

The commutator $\big[f, J_{\varepsilon\nu}(\varepsilon\partial_t)^m\big]u$ is expanded to

$$\big[f, J_{\varepsilon\nu}\big](\varepsilon\partial_t)^m u + J_{\varepsilon\nu}\big[f, (\varepsilon\partial_t)^m\big]u. \tag{6.11}$$

**a)** We first claim that

$$\big\| \big[f, J_{\varepsilon\nu}\big](\varepsilon\partial_t)^m u\big\|_{H^{s-m+1}_{\varepsilon\nu}} \lesssim \varepsilon\nu \, \|f\|_{H^s} \, \big\| Z^m_{\varepsilon,\nu}\, u\big\|_{H^{s-m}} . \tag{6.12}$$

Since $s > 1+d/2$ and $1 \leqslant m \leqslant s$, Proposition 3.10 applies with $(h, m, \sigma_0, \sigma)$[2] replaced by $(\varepsilon\nu, 1, s, s-m)$. It yields

$$\big\| \big[f, J_{\varepsilon\nu}\big](\varepsilon\partial_t)^m u\big\|_{H^{s-m}} \lesssim \varepsilon\nu \, \|f\|_{H^s} \, \|\Lambda^{-(2s-m)}_{\varepsilon\nu}(\varepsilon\partial_t)^m u\|_{H^{s-m}} .$$

Since $m \leqslant s$, there holds $\Lambda^{-(2s-m)}_{\varepsilon\nu} \leqslant \Lambda^{-m}_{\varepsilon\nu}$. Hence, we have the first half of (6.12), namely:

$$\big\| \big[f, J_{\varepsilon\nu}\big](\varepsilon\partial_t)^m u\big\|_{H^{s-m}} \lesssim \varepsilon\nu \, \|f\|_{H^s} \, \big\| Z^m_{\varepsilon,\nu}\, u\big\|_{H^{s-m}} . \tag{6.13}$$

The technique for obtaining the second half is similar. We first apply Proposition 3.10 with $m = 0$, to obtain

$$\big\| \big[f, J_{\varepsilon\nu}\big](\varepsilon\partial_t)^m u\big\|_{H^{s-m+1}} \lesssim \|f\|_{H^s} \, \big\| Z^m_{\varepsilon,\nu}\, u\big\|_{H^{s-m}} , \tag{6.14}$$

and we next multiply (6.14) by $\varepsilon\nu$.

**b)** Moving to the second term in (6.11), we claim that

$$\big\| J_{\varepsilon\nu}\big[f, (\varepsilon\partial_t)^m\big]u\big\|_{H^{s-m+1}_{\varepsilon\nu}} \lesssim \varepsilon \sum_{\ell=0}^{m-1} \big\| Z^\ell_{\varepsilon,\nu}\, \partial_t f\big\|_{H^{s-1-\ell}} \sum_{\ell=0}^{m-1} \big\| Z^\ell_{\varepsilon,\nu}\, u\big\|_{H^{s-1-\ell}} . \tag{6.15}$$

Starting from the Leibniz rule, we get

$$\big[f, (\varepsilon\partial_t)^m\big]u = \varepsilon \sum_{\ell=0}^{m-1} \binom{m}{\ell+1} \big((\varepsilon\partial_t)^\ell \partial_t f\big)\, (\varepsilon\partial_t)^{m-1-\ell} u. \tag{6.16}$$

Let $0 \leqslant \ell \leqslant m-1$. Since $s > 1+d/2$ and $1 \leqslant m \leqslant s$, Proposition 3.2 applies with $\sigma_0 = s-1$, $\sigma_1 = m_1 = \ell$ and $\sigma_2 = m_2 = m-1-\ell$. It yields

$$\begin{aligned} &\big\|\Lambda^{-m+1}_{\varepsilon\nu}\big((\varepsilon\partial_t)^\ell \partial_t f\big)\, (\varepsilon\partial_t)^{m-1-\ell} u\big)\big\|_{H^{s-m}} \\ &\qquad \lesssim \|\Lambda^{-\ell}_{\varepsilon\nu}(\varepsilon\partial_t)^\ell \partial_t f\|_{H^{s-1-\ell}} \, \|\Lambda^{-(m-1-\ell)}_{\varepsilon\nu}(\varepsilon\partial_t)^{m-1-\ell} u\|_{H^{s-1-(m-1-\ell)}} . \end{aligned}$$

By summing over all $0 \leqslant \ell \leqslant m-1$, we obtain from the definition of the operators $Z^{\cdot}_{\varepsilon,\nu}$ that $\big\|\Lambda^{-m+1}_{\varepsilon\nu}\big[f, (\varepsilon\partial_t)^m\big]u\big\|_{H^{s-m}}$ is estimated by the right-hand side of (6.15). To complete the proof of (6.15), use the elementary estimate

$$\|J_{\varepsilon\nu} u\|_{H^{s-m+1}_{\varepsilon\nu}} \lesssim \|J_{\varepsilon\nu} u\|_{H^{s-m}} \lesssim \|\Lambda^{-m+1}_{\varepsilon\nu} u\|_{H^{s-m}} . \tag{6.17}$$

□

[2]Here, $m$ refers to the index used in the statement of Lemma 3.3.

6.2. **Notations.** Let us pause here to fix a few notations and to make some running conventions.

From now on, we consider a time $0 < T \leqslant 1$, a fixed triple of parameter $a = (\varepsilon, \mu, \kappa) \in A$ and a smooth solution $U = (p, v, \theta) \in C^\infty([0,T]; H^\infty(\mathbb{D}))$ of the system (2.1). The notations $\phi$, $\psi$ and $\Psi$ are shorthand notations for

$$\phi := (\theta, \varepsilon p), \quad \psi := (\theta, \varepsilon p, \varepsilon v) \quad \text{and} \quad \Psi := (\psi, \partial_t \psi, \nabla \psi). \tag{6.18}$$

Hereafter, $s$ always denotes a fixed integer $s > 1 + d/2$. We set

$$\nu := \sqrt{\mu + \kappa}, \quad \Omega_0 := \| U(0) \|_{\mathcal{H}^s_{a,0}} \quad \text{and} \quad \Omega := \| U \|_{\mathcal{H}^s_a(T)},$$

where the norms are defined in Definition 2.4.

As in (5.22), we set

$$\begin{aligned} R &:= \| (p,v) \|_{H^{s+1}_{\varepsilon\nu}} + \| \theta \|_{H^{s+1}_\nu}, \\ R' &:= \sqrt{\mu + \kappa} \, \| (\operatorname{div} v, \nabla p) \|_{H^s} + \sqrt{\mu} \, \| \nabla v \|_{H^{s+1}_{\varepsilon\nu}} + \sqrt{\kappa} \, \| \nabla \theta \|_{H^{s+1}_\nu}. \end{aligned} \tag{6.19}$$

With these notations, we have $\Omega \approx \| R \|_{L^\infty(0,T)} + \| R' \|_{L^2(0,T)}$.

To say that a smooth nondecreasing function $C\colon [0,+\infty) \to [1,+\infty)$ is *generic* means that $C(\cdot)$ is independent of $T$, $a$ and $(p,v,\theta)$. Given a generic function $C(\cdot)$, we denote by $\widetilde{C}$ the positive constant

$$\widetilde{C} := C(\Omega_0) e^{(\sqrt{T} + \varepsilon) C(\Omega)}. \tag{6.20}$$

The factor $\varepsilon$ is of no consequence but makes some arguments work more smoothly. We will often use the simple observation that for all generic function $C(\cdot)$, we have $C(\Omega_0) + (\sqrt{T} + \varepsilon) C(\Omega) \leqslant \widetilde{C}$.

To clarify matters, with these conventions, Proposition 6.1 is formulated in the following way: there exists a generic function $C(\cdot)$ such that $\| J_{\varepsilon\nu} U \|_{\mathcal{H}^s_a(T)} \leqslant \widetilde{C}$.

6.3. **Localization in the low frequency region.**

**Definition 6.8.** *Given $m \in \mathbb{N}$, set*

$$\mathcal{X}^m := J_{\varepsilon\nu} (\varepsilon \partial_t)^m. \tag{6.21}$$

**Notation 6.9.** Denote by $f^m_{\mathrm{LF}} = (f^m_{1,\mathrm{LF}}, f^m_{2,\mathrm{LF}}, f^m_{3,\mathrm{LF}})$ the commutator of the equations (2.1) and the operator $\mathcal{X}^m$:

$$\begin{aligned} f^m_{1,\mathrm{LF}} &:= \big[ g_1(\phi), \mathcal{X}^m \big] \partial_t p + \big[ g_1(\phi) v, \mathcal{X}^m \big] \cdot \nabla p - \frac{\kappa}{\varepsilon} \big[ B_1(\phi), \mathcal{X}^m \big] \theta, \\ f^m_{2,\mathrm{LF}} &:= \big[ g_2(\phi), \mathcal{X}^m \big] \partial_t v + \big[ g_2(\phi) v, \mathcal{X}^m \big] \cdot \nabla v - \mu \big[ B_2(\phi), \mathcal{X}^m \big] v, \\ f^m_{3,\mathrm{LF}} &:= \big[ g_3(\phi), \mathcal{X}^m \big] \partial_t \theta + \big[ g_3(\phi) v, \mathcal{X}^m \big] \cdot \nabla \theta - \kappa \big[ B_3(\phi), \mathcal{X}^m \big] \theta. \end{aligned}$$

The aim of this paragraph is to estimate $f^m_{\mathrm{LF}}$. To begin with, consider the case when $m \geqslant 1$.

**Lemma 6.10.** *There exists a generic function $C(\cdot)$ such that for all $m \in \mathbb{N}$ such that $1 \leqslant m \leqslant s$,*

$$\|f^m_{\mathrm{LF}}\|_{\mathcal{H}^{s-m}_{a,0}} := \big\|\big(f^m_{1,\mathrm{LF}}, f^m_{2,\mathrm{LF}}\big)\big\|_{H^{s-m+1}_{\varepsilon\nu}} + \big\|f^m_{3,\mathrm{LF}}\big\|_{H^{s-m+1}_{\nu}} \leqslant C(R)\{1 + R'\},$$

*where $R$ and $R'$ are as defined in* (6.19)*.*

*Proof.* To emphasize the role of the slow component $\Psi$ (as defined in (6.18)), many bounds are given in terms of:

$$\gamma := \|\Psi\|_{H^{s-1}} \quad \text{and} \quad \Gamma := \|\Psi\|_{H^s_\nu}.$$

Directly from Lemma 5.14, $\|\partial_t\psi\|_{H^{s-1}} \leqslant C(R)$ and $\|\partial_t\psi\|_{H^s_\nu} \leqslant C(R)\{1+R'\}$. Furthermore, directly from the definition of $\psi$ (see (6.18)), we have the estimate $\|(\psi, \nabla\psi)\|_{H^s_\nu} \leqslant \|(p, v)\|_{H^{s+1}_{\varepsilon\nu}} + \|\theta\|_{H^{s+1}_\nu} \leqslant R$. Hence,

$$\gamma = \|\Psi\|_{H^{s-1}} \leqslant C(R) \quad \text{and} \quad \Gamma = \|\Psi\|_{H^s_\nu} \leqslant C(R)\{1 + R'\}. \tag{6.22}$$

Note that Proposition 6.5 applies since the condition $\nu \in [(\mu + \kappa)/2, 2]$ is fulfilled.

Step 1: Estimate for $f^m_{1,\mathrm{LF}}$.
**a)** We begin by proving

$$\big\|\big[g_1(\phi), \mathcal{X}^m\big]\partial_t p\big\|_{H^{s-m+1}_{\varepsilon\nu}} \leqslant C(\gamma)\Gamma. \tag{6.23}$$

Starting from Lemma 6.6, we find that the left-hand side is bounded by

$$\Big\{\|\tilde{g}_1(\phi)\|_{H^s} + \sum_{\ell=0}^{m-1} \big\|Z^\ell_{\varepsilon,\nu}\, \partial_t g_1(\phi)\big\|_{H^{s-1-\ell}}\Big\}\Big\{\sum_{\ell=0}^{m} \big\|Z^\ell_{\varepsilon,\nu}\, (\varepsilon\partial_t p)\big\|_{H^{s-\ell}_\nu}\Big\}. \tag{6.24}$$

Since $\Psi = (\ldots, (\varepsilon\partial_t)p, \ldots)$, the estimate (6.6) implies that the second factor in (6.24) is estimated by $C(\gamma)\Gamma$. Moving to the first factor, note that $\partial_t g_1(\phi) = F(\phi, \partial_t\phi)$ for some $C^\infty$ function $F$ such that $F(0) = 0$. As a consequence, Lemma 6.4 implies that

$$\sum_{\ell=0}^{m-1} \big\|Z^\ell_{\varepsilon,\nu}\, \partial_t g_1(\phi)\big\|_{H^{s-1-\ell}} \leqslant C\Big(\sum_{\ell=0}^{m-1} \big\|Z^\ell_{\varepsilon,\nu}\, (\phi, \partial_t\phi)\big\|_{H^{s-1-\ell}}\Big). \tag{6.25}$$

Again, since $\Psi = (\phi, \ldots, \partial_t\phi, \ldots)$, the estimate (6.5) implies that the right-hand side in the previous estimate is bounded by $C(\gamma)$. It remains to estimate $\|\tilde{g}_1(\phi)\|_{H^s}$. To do so we write

$$\|\tilde{g}_1(\phi)\|_{H^s} \leqslant C\big(\|\phi\|_{H^s}\big) \leqslant C\big(\|(\psi, \nabla\psi)\|_{H^{s-1}}\big) \leqslant C(\gamma).$$

**b)** We next prove that

$$\big\|\big[g_1(\phi)v, \mathcal{X}^m\big] \cdot \nabla p\big\|_{H^{s-m+1}_{\varepsilon\nu}} \leqslant C(\gamma)\big\{\|\nabla p\|_{H^{s-1}} + \Gamma\big\}. \tag{6.26}$$

Lemma 6.6 implies that the left-hand side is bounded by

$$
\begin{aligned}
&\Big\{\|\varepsilon g_1(\phi)v\|_{H^s} + \sum_{\ell=0}^{m-1} \big\|Z^\ell_{\varepsilon,\nu}\,\partial_t(\varepsilon g_1(\phi)v)\big\|_{H^{s-1-\ell}}\Big\} \\
&\qquad\qquad \times\Big\{\,\|\nabla p\|_{H^{s-1}} + \sum_{\ell=1}^{m} \big\|Z^\ell_{\varepsilon,\nu}\,\nabla p\big\|_{H^{s-\ell}_\nu}\Big\}.
\end{aligned} \tag{6.27}
$$

By definition $\psi = (\theta, \varepsilon p, \varepsilon v)$. Hence, one can rewrite $\varepsilon g_1(\phi)v$ as $G(\psi)$ for some $C^\infty$ function $G$. Consequently, the first term in the above written product is the exact analogue of the first term in (6.24) with $g_1(\phi)$ replaced by $G(\psi)$. We thus obtain that this term is estimated by $C(\gamma)$.

Next, using the very definitions of $Z^\ell_{\varepsilon,\nu}$, we write

$$
\sum_{\ell=1}^{m} \big\|Z^\ell_{\varepsilon,\nu}\,\nabla p\big\|_{H^{s-\ell}_\nu} = \sum_{\ell=1}^{m} \big\|Z^{\ell-1}_{\varepsilon,\nu}\,\Lambda^{-1}_{\varepsilon\nu}\nabla(\varepsilon\partial_t p)\big\|_{H^{s-\ell}_\nu}\,. \tag{6.28}
$$

Since $\Lambda^{-1}_{\varepsilon\nu} \lesssim I$, this in turn implies

$$
\sum_{\ell=1}^{m} \big\|Z^\ell_{\varepsilon,\nu}\,\nabla p\big\|_{H^{s-\ell}_\nu} \leqslant \sum_{\ell=0}^{m-1} \big\|Z^\ell_{\varepsilon,\nu}\,(\varepsilon\partial_t p)\big\|_{H^{s-\ell}_\nu}\,,
$$

which, as in the previous step **a)**, is estimated by $C(\gamma)\Gamma$.

**c)** To complete the estimate of $f^m_{1,\mathrm{LF}}$, we establish

$$
\frac{\kappa}{\varepsilon}\,\big\|\big[B_1(\phi),\mathcal{X}^m\big]\theta\big\|_{H^{s-m+1}_{\varepsilon\nu}} \leqslant C(\gamma,R)\big\{\|\kappa\theta\|_{H^{s+2}_\nu} + \Gamma\big\}. \tag{6.29}
$$

Parallel to (5.17), we decompose the commutator $\kappa\varepsilon^{-1}\big[B_1(\phi),\mathcal{X}^m\big]\theta$ as

$$
\frac{\kappa}{\varepsilon}\big[\mathcal{E}_1(\phi),\mathcal{X}^m\big]\Delta\theta + \frac{\kappa}{\varepsilon}\big[\mathcal{E}_2(\phi,\nabla\phi),\mathcal{X}^m\big]\cdot\nabla\theta, \tag{6.30}
$$

where $\mathcal{E}_1(\phi) := \chi_1(\varepsilon p)\beta(\theta)$ and $\mathcal{E}_2(\phi,\nabla\phi) := \chi_1(\varepsilon p)\beta'(\theta)\nabla\theta$.

Replacing $\partial_t p$ with $\kappa\varepsilon^{-1}\Delta\theta$, the arguments given in **a)** yield

$$
\frac{\kappa}{\varepsilon}\,\big\|\big[\mathcal{E}_1(\phi),\mathcal{X}^m\big]\Delta\theta\big\|_{H^{s-m+1}_{\varepsilon\nu}} \leqslant C(\gamma)\sum_{\ell=0}^{m} \big\|Z^\ell_{\varepsilon,\nu}\,(\kappa\Delta\theta)\big\|_{H^{s-\ell}_\nu}\,.
$$

We split the sum in the previous right-hand side as

$$
\|\kappa\Delta\theta\|_{H^s_\nu} + \sum_{\ell=1}^{m} \big\|Z^\ell_{\varepsilon,\nu}\,(\kappa\Delta\theta)\big\|_{H^{s-\ell}_\nu}\,.
$$

The first term is bounded by $\|\kappa\theta\|_{H^{s+2}_\nu}$. With regards to the second term, we proceed as in **b)**. Namely, we write

$$\begin{aligned}\sum_{\ell=1}^{m} \kappa \left\| Z^{\ell}_{\varepsilon,\nu}\, \Delta\theta \right\|_{H^{s-\ell}_{\nu}} &= \sum_{\ell=1}^{m} \left\| Z^{\ell-1}_{\varepsilon,\nu} \left(\varepsilon\kappa\Lambda^{-1}_{\varepsilon\nu}\Delta\right)\partial_t\theta \right\|_{H^{s-\ell}_{\nu}} \\ &\leqslant \sum_{\ell=0}^{m-1} \left\| Z^{\ell}_{\varepsilon,\nu}\, \partial_t\theta \right\|_{H^{s-\ell}_{\nu}},\end{aligned}$$

where we have used $\varepsilon\kappa\Lambda^{-1}_{\varepsilon\nu}\Delta \lesssim \Lambda^1$ (which stems from $\kappa \leqslant \sqrt{\kappa} \leqslant \nu$). Note that $\Psi = (\ldots, \partial_t\theta, \ldots)$. Therefore, estimate (6.6) implies that the right-hand side of the previous inequality is controlled by $C(\gamma)\Gamma$.

We now have to estimate the $H^{s-m+1}_{\varepsilon\nu}$-norm of the second term in (6.30). Since $\kappa \leqslant \nu$ and $\kappa \leqslant 1$, Lemma 6.6 implies that this term is bounded by

$$\Big\{ \|\nu\mathcal{E}_2\|_{H^s} + \sum_{\ell=0}^{m-1} \left\| \nu Z^{\ell}_{\varepsilon,\nu}\, \partial_t\mathcal{E}_2 \right\|_{H^{s-1-\ell}} \Big\} \Big\{ \|\nabla\theta\|_{H^{s-1}} + \sum_{\ell=1}^{m} \left\| \nu Z^{\ell}_{\varepsilon,\nu}\, \nabla\theta \right\|_{H^{s-\ell}} \Big\}.$$

This in turn is bounded by $C(\gamma)\{C(R) + \Gamma\}$ since

$$\|\nu\mathcal{E}_2(\phi,\nabla\phi)\|_{H^s} \leqslant C(R), \qquad \nu \left\| Z^{\ell}_{\varepsilon,\nu}\, \partial_t\mathcal{E}_2(\phi,\nabla\phi) \right\|_{H^{s-1-\ell}} \leqslant C(\gamma)\Gamma, \tag{6.31}$$

$$\|\nabla\theta\|_{H^{s-1}} \leqslant \gamma, \qquad \forall \ell \in [1,m], \quad \left\| \nu Z^{\ell}_{\varepsilon,\nu}\, \nabla\theta \right\|_{H^{s-\ell}} \leqslant C(\gamma). \tag{6.32}$$

The first inequality in (6.31) follows from (5.7). The second inequality follows from (6.10) since the term $\partial_t\mathcal{E}_2(\phi,\nabla\phi)$ can be written as $F(\Psi)\nabla\Psi$ for some $C^\infty$ function $F$. The first inequality in (6.32) is obvious. In order to prove the last one, as already seen, we can write

$$\begin{aligned}\left\| \nu Z^{\ell}_{\varepsilon,\nu}\, \nabla\theta \right\|_{H^{s-\ell}} &\lesssim \left\| Z^{\ell-1}_{\varepsilon,\nu} \left(\varepsilon\nu\Lambda^{-1}_{\varepsilon\nu}\nabla\right)\partial_t\theta \right\|_{H^{s-\ell}} \\ &\lesssim \left\| Z^{\ell-1}_{\varepsilon,\nu}\, \Psi \right\|_{H^{s-1-(\ell-1)}} \leqslant C(\gamma).\end{aligned}$$

**d)** By combining (6.23) with (6.26) and (6.29), we obtain

$$\left\| f^m_{1,\mathrm{LF}} \right\|_{H^{s-m+1}_{\varepsilon\nu}} \leqslant C(\gamma,R)\{\|\kappa\theta\|_{H^{s+2}_\nu} + \Gamma\},$$

so that the desired estimate $\|f^m_{1,\mathrm{LF}}\|_{H^{s-m+1}_{\varepsilon\nu}} \leqslant C(R)\{1+R'\}$ follows from observation (6.22).

Step 2: Estimate for $f^m_{2,\mathrm{LF}}$.

Note that one can obtain $f^m_{2,\mathrm{LF}}$ from $f^m_{1,\mathrm{LF}}$ by replacing $p$ by $v$, $\theta$ by $\varepsilon v$ and $\kappa$ by $\mu$. Therefore, we are in the situation of the previous step and hence we can conclude that

$$\begin{aligned}\left\| \left[g_2(\phi),\mathcal{X}^m\right]\partial_t v \right\|_{H^{s-m+1}_{\varepsilon\nu}} &\leqslant C(\gamma)\Gamma, \\ \left\| \left[g_2(\phi)v,\mathcal{X}^m\right]\cdot\nabla v \right\|_{H^{s-m+1}_{\varepsilon\nu}} &\leqslant C(\gamma)\big\{\|\nabla v\|_{H^{s-1}} + \Gamma\big\}, \\ \mu \left\| \left[B_2(\phi),\mathcal{X}^m\right] v \right\|_{H^{s-m+1}_{\varepsilon\nu}} &\leqslant C(\gamma,R)\big\{\|\varepsilon\mu v\|_{H^{s+2}_\nu} + \Gamma\big\}.\end{aligned}$$

Step 3: Estimate for $f^m_{3,\mathrm{LF}}$.

Note that, for all $u \in H^\sigma_\nu$,

(6.33) $$\|u\|_{H^\sigma_\nu} \leqslant \varepsilon^{-1} \|u\|_{H^\sigma_{\varepsilon\nu}}.$$

Hence,

$$\begin{aligned}\big\| \big[ g_3(\phi), \mathcal{X}^m \big] \partial_t \theta \big\|_{H^{s-m+1}_\nu} + \kappa \big\| \big[ B_3(\phi), \mathcal{X}^m \big] \theta \big\|_{H^{s-m+1}_\nu} \\ \leqslant \frac{1}{\varepsilon} \big\| \big[ g_3(\phi), \mathcal{X}^m \big] \partial_t \theta \big\|_{H^{s-m+1}_{\varepsilon\nu}} + \frac{\kappa}{\varepsilon} \big\| \big[ B_3(\phi), \mathcal{X}^m \big] \theta \big\|_{H^{s-m+1}_{\varepsilon\nu}}.\end{aligned}$$

The first term in the right-hand side is estimated as in **a)** above (by replacing $\partial_t p$ by $\varepsilon^{-1}\partial_t\theta$) and the second term has been estimated in **c)**.

For technical reasons, the estimate for $[g_3(\phi)v, \mathcal{X}^m]\cdot\nabla\theta$ is somewhat more complicated. We argue as in the proof of Lemma 6.6. Set $f := g_3(\phi)v$ and $u = \nabla\theta$. We begin by splitting the commutator $[f, \mathcal{X}^m]\cdot u$ as

$$P + Q := J_{\varepsilon\nu}[f, (\varepsilon\partial_t)^m]\cdot u + [f, J_{\varepsilon\nu}]\cdot(\varepsilon\partial_t)^m u.$$

By combining (6.13) with (6.14) multiplied by $\nu$, we find that

$$\|Q\|_{H^{s-m+1}_\nu} \lesssim \|f\|_{H^s} \big\| Z^m_{\varepsilon,\nu} u \big\|_{H^{s-m}_\nu} \leqslant C(\gamma, R)\Gamma,$$

where we have used (6.32).

Our next task is to show a similar estimate for $P$. To do so we decompose $P$ into two parts:

(6.34) $$P_1 + P_2 := \Big\{ J_{\varepsilon\nu}[f, (\varepsilon\partial_t)^m]\cdot u - J_{\varepsilon\nu}\big(u\cdot(\varepsilon\partial_t)^m f\big) \Big\} + J_{\varepsilon\nu}\big(u\cdot(\varepsilon\partial_t)^m f\big).$$

Let us prove that $\|P_1\|_{H^{s-m+1}_\nu} \leqslant C(\gamma)$. In light of (6.33), all we need to prove is that

(6.35) $$\|P_1\|_{H^{s-m+1}_{\varepsilon\nu}} \leqslant \varepsilon C(\gamma).$$

We repeat the proof of Lemma 6.6, to obtain

$$\|P_1\|_{H^{s-m+1}_{\varepsilon\nu}} \lesssim \varepsilon \sum_{\ell=1}^{m-1} \big\| Z^\ell_{\varepsilon,\nu}\, \partial_t f \big\|_{H^{s-1-\ell}} \big\| Z^\ell_{\varepsilon,\nu}\, u \big\|_{H^{s-1-\ell}}.$$

The sum differs from the one that appears in Lemma 6.6 in that it is indexed by $\ell \geqslant 1$ instead of $\ell \geqslant 0$. This fact allows us to write

$$\|P_1\|_{H^{s-m+1}_{\varepsilon\nu}} \lesssim \varepsilon \sum_{\ell=1}^{m-1} \big\| Z^\ell_{\varepsilon,\nu}\, (\varepsilon\partial_t) f \big\|_{H^{s-1-\ell}} \big\| Z^{\ell-1}_{\varepsilon,\nu}\, \partial_t u \big\|_{H^{s-1-\ell}}.$$

Let $1 \leqslant \ell \leqslant m-1$. We write $(\varepsilon\partial_t)f$ as $F(\Psi)$ for some $C^\infty$ function $F$ such that $F(0) = 0$. By combining Lemma 6.4 and the estimate (6.5), we obtain

$$\big\| Z^\ell_{\varepsilon,\nu}\, (\varepsilon\partial_t) f \big\|_{H^{s-1-\ell}} \leqslant C(\gamma).$$

Moving to the estimate of $Z^{\ell-1}_{\varepsilon,\nu}\, \partial_t u$, we use the very definitions of $u = \nabla\theta$ and $\Psi = (\dots, \partial_t\theta, \dots)$, to obtain thanks to (6.5):

$$\big\| Z^{\ell-1}_{\varepsilon,\nu}\, \partial_t u \big\|_{H^{s-1-\ell}} \leqslant \big\| Z^{\ell-1}_{\varepsilon,\nu}\, \Psi \big\|_{H^{s-\ell}} = \big\| Z^{\ell-1}_{\varepsilon,\nu}\, \Psi \big\|_{H^{s-(\ell-1)-1}} \leqslant C(\gamma).$$

We have proved (6.35). So to conclude it remains only to estimate the second term $P_2$ in (6.34). This is accomplished using

$$\|P_2\|_{H_\nu^{s-m+1}} \leqslant \|u\|_{H_\nu^s} \left\|Z_{\varepsilon,\nu}^{m-1}\,(\varepsilon\partial_t)f\right\|_{H_\nu^{s-m+1}},$$

as the reader can verify, yielding the bound $\|P_2\|_{H_\nu^{s-m+1}} \leqslant C(\gamma,R)\Gamma$.

This completes the proof of Lemma 6.10. □

Note that in the case when $m=0$, the previous method does not work as it stands, since the estimate (6.12) is no longer correct.

Now we give an estimate valid for all $0 \leqslant m \leqslant s-1$.

**Lemma 6.11.** *There exists a generic function $C(\cdot)$ such that for all $m\in\mathbb{N}$ such that $0\leqslant m\leqslant s-1$,*

$$\|f_{\mathrm{LF}}^m\|_{\mathcal{H}_{a,0}^{s-1-m}} := \left\|\big(f_{1,\mathrm{LF}}^m, f_{2,\mathrm{LF}}^m\big)\right\|_{H_{\varepsilon\nu}^{s-m}} + \left\|f_{3,\mathrm{LF}}^m\right\|_{H_\nu^{s-m}} \leqslant C(R). \tag{6.36}$$

The estimate (6.36) is nothing new in that it can be deduced by following the proof of Lemma 6.10. One only has to use the following analogue of the calculus inequality in Lemma 6.6:

$$\begin{aligned}\left\|[f, J_{\varepsilon\nu}(\varepsilon\partial_t)^m]u\right\|_{H_{\varepsilon\nu}^{s-m}} \lesssim \varepsilon\Big\{\|f\|_{H^s} + \sum_{\ell=0}^{m-1}\left\|Z_{\varepsilon,\nu}^{\ell}\,\partial_t f\right\|_{H^{s-2-\ell}}\Big\}\\ \times\Big\{\left\|Z_{\varepsilon,\nu}^{m}\,u\right\|_{H^{s-1-m}} + \sum_{\ell=0}^{m-1}\left\|Z_{\varepsilon,\nu}^{\ell}\,u\right\|_{H^{s-1-\ell}}\Big\}.\end{aligned}$$

In the case when $m=0$ the sum $\sum_0^{-1}$ is interpreted as 0. The index $s-2-\ell$ in the second term of the first set of parentheses is not a typographical error. It is of use to us for the estimate of $\partial_t\mathcal{E}_2(\phi,\nabla\phi)$ where $\mathcal{E}_2$ is as in (6.31).

6.4. **The fast components.** We give here the estimates for the fast components $\operatorname{div} v$ and $\nabla p$. We use of the notations introduced in §6.2.

**Notation 6.12.** For all integer $m\leqslant s$, set $U_m := \mathcal{X}^m U = J_{\varepsilon\nu}(\varepsilon\partial_t)^m U$.

As a preliminary step towards the estimate of $(\operatorname{div} v, \nabla p)$, we estimate the $\mathcal{H}_a^0(T)$-norm of $U_m$.

**Lemma 6.13.** *For all integers $m\leqslant s$, there exists a generic function $C$ such that $\|U_m\|_{\mathcal{H}_a^0(T)} \leqslant \widetilde{C}$, where $\widetilde{C}$ is as defined in* (6.20).

*Proof.* Having estimated the commutators $f_{\mathrm{LF}}^m$, this result can be deduced by following the end of proof of Proposition 5.1 (see §5.3). We therefore only indicate the points at which the argument is slightly different.

It readily follows from Notation 6.9 that $(\widetilde{p},\widetilde{v},\widetilde{\theta}) := (p_m, v_m, \theta_m)$ satisfies the linearized system (4.1) where

$$\begin{aligned}
&f_1 := f_{1,\mathrm{LF}}^m + f_{1,\mathrm{LF}}' \quad \text{with} \quad f_{1,\mathrm{LF}}' := -\frac{\kappa}{\varepsilon}\nabla\chi_1(\varepsilon p)\cdot\big(\beta(\theta)\nabla\theta_m\big),\\
&f_2 := f_{2,\mathrm{LF}}^m + f_{2,\mathrm{LF}}' \quad \text{with} \quad f_{2,\mathrm{LF}}' := \mu\chi_2(\varepsilon p)\big\{2Dv_m\nabla\zeta(\theta) + \operatorname{div} v_m\nabla\eta(\theta)\big\},\\
&f_3 := f_{3,\mathrm{LF}}^m + f_{3,\mathrm{LF}}' \quad \text{with} \quad f_{3,\mathrm{LF}}' := \kappa\chi_3(\varepsilon p)\nabla\beta(\theta)\cdot\nabla\theta_m,
\end{aligned}$$

recalling that we deliberately omit the terms $\Upsilon_1$ and $\varepsilon\Upsilon_3$ in the system (2.1).

Set

$$\mathfrak{F}(T) := \left\|(f^m_{1,\mathrm{LF}}, f^m_{2,\mathrm{LF}})\right\|_{L^1(0,T;H^1_{\varepsilon\nu})} + \left\|f^m_{3,\mathrm{LF}}\right\|_{L^1(0,T;H^1_\nu)},$$
$$\mathfrak{F}'(T) := \left\|(f'_{1,\mathrm{LF}}, f'_{2,\mathrm{LF}})\right\|_{L^1(0,T;H^1_{\varepsilon\nu})} + \left\|f'_{3,\mathrm{LF}}\right\|_{L^1(0,T;H^1_\nu)}.$$

Applying Theorem 4.3, we get

$$\|U_m\|_{\mathcal{H}^0_a(T)} \leqslant C(\Omega_0)e^{TC(\Omega)}\,\|U_m(0)\|_{\mathcal{H}^0_{a,0}} + C(\Omega)\mathfrak{F}(T) + C(\Omega)\mathfrak{F}'(T).$$

The proof thus reduces to establishing that

$$\|U_m(0)\|_{\mathcal{H}^0_{a,0}} \leqslant C(\Omega_0), \qquad \mathfrak{F}(T) + \mathfrak{F}'(T) \leqslant \sqrt{T}C(\Omega).$$

To fix matters, we concentrate on the hardest case when $m = s$. Note that the conclusion of Lemma 6.10 is the exact analogue of (5.24). Hence, we have $\mathfrak{F}(T) \leqslant \sqrt{T}C(\Omega)$. As in §5.3, all that has to be done in order to prove $\mathfrak{F}'(T) \leqslant \sqrt{T}C(\Omega)$ is to check $\sqrt{\kappa}\|\nabla\theta_s\|_{H^1_\nu} + \sqrt{\mu}\|\nabla v_s\|_{H^1_{\varepsilon\nu}} \leqslant C(R)\{1 + R'\}$. To do so, using the definition of $\theta_s = J_{\varepsilon\nu}(\varepsilon\partial_t)^s\theta$, we first rewrite $\nabla\theta_s$ as $\varepsilon J_{\varepsilon\nu}(\varepsilon\partial_t)^{s-1}\nabla\partial_t\theta$. Next, by combining the estimate $\|\varepsilon J_{\varepsilon\nu}u\|_{H^1_\nu} \leqslant \|J_{\varepsilon\nu}u\|_{H^1_{\varepsilon\nu}} \lesssim \|J_{\varepsilon\nu}u\|_{L^2}$ with the inequality $\sqrt{\kappa} \leqslant \nu$ and the definition of $\Psi = (\ldots, \partial_t\theta, \ldots)$, we obtain

$$\sqrt{\kappa}\,\|\nabla\theta_s\|_{H^1_\nu} \lesssim \nu\left\|Z^{s-1}_{\varepsilon,\nu}\,\nabla\Psi\right\|_{L^2} \leqslant \left\|Z^{s-1}_{\varepsilon,\nu}\,\Psi\right\|_{H^1_\nu}.$$

Analogous computations lead to

$$\sqrt{\mu}\|\nabla v_s\|_{H^1_{\varepsilon\nu}} \lesssim \|v_s\|_{H^1_\nu} \lesssim \left\|Z^{s-1}_{\varepsilon,\nu}\,(\varepsilon\partial_t)v\right\|_{H^1_\nu} \leqslant \left\|Z^{s-1}_{\varepsilon,\nu}\,\Psi\right\|_{H^1_\nu}.$$

Hence, the desired bound follows from (6.6) and (6.22).

The technique for estimating the initial data is similar. Indeed, we have

$$\|U_s\|_{\mathcal{H}^0_{a,0}} \lesssim \left\|Z^{s-1}_{\varepsilon,\nu}\,\Psi\right\|_{L^2},$$

as the reader can verify, yielding $\|U_s(0)\|_{\mathcal{H}^0_{a,0}} \leqslant C(R(0)) = C(\Omega_0)$. □

We now come to the induction argument.

**Notation 6.14.** Define $\|u\|_{\mathcal{K}^\sigma_\nu(T)} := \|u\|_{L^\infty(0,T;H^{\sigma-1})} + \nu\,\|u\|_{L^2(0,T;H^\sigma)}$.

**Lemma 6.15.** *Let $\widetilde{U} := (\widetilde{p}, \widetilde{v}, \widetilde{\theta})$ be a solution of the system:*

$$\begin{cases} g_1(\phi)\partial_t\widetilde{p} + \dfrac{1}{\varepsilon}\operatorname{div}\widetilde{v} - \dfrac{\kappa}{\varepsilon}B_1(\phi)\widetilde{\theta} = f_1, \\ g_2(\phi)\partial_t\widetilde{v} + \dfrac{1}{\varepsilon}\nabla\widetilde{p} - \mu B_2(\phi)\widetilde{v} = f_2, \\ g_3(\phi)\partial_t\widetilde{\theta} + \operatorname{div}\widetilde{v} - \kappa B_3(\phi)\widetilde{\theta} = f_3. \end{cases} \tag{6.37}$$

*If the Fourier transform of $\widetilde{U}$ is supported in the ball $\{|\xi| \leqslant 2/\varepsilon\nu\}$, then there exists a generic function $C(\cdot)$ such that for all $\sigma \in [1, s]$,*

$$\begin{aligned}
&\|\widetilde{p}\|_{\mathcal{K}_\nu^{\sigma+1}(T)} + \|\operatorname{div}\widetilde{v}\|_{\mathcal{K}_\nu^{\sigma}(T)} \\
&\quad \leqslant \widetilde{C}\,\|(\varepsilon\partial_t)\widetilde{p}\|_{\mathcal{K}_\nu^{\sigma}(T)} + \widetilde{C}\,\|(\varepsilon\partial_t)\operatorname{div}\widetilde{v}\|_{\mathcal{K}_\nu^{\sigma-1}(T)} \\
&\qquad + \widetilde{C}\,\|\widetilde{p}\|_{L^\infty(0,T;L^2)} + \widetilde{C}\,\|\widetilde{\theta}(0)\|_{H_\nu^{\sigma+1}} + \varepsilon C(\Omega)\,\|\mu\widetilde{v}\|_{\mathcal{K}_\nu^{\sigma+1}(T)} \\
&\qquad + \varepsilon C(\Omega)\,\|(f_1, f_2)\|_{\mathcal{K}_\nu^{\sigma}(T)} + \nu\widetilde{C}\,\|f_3\|_{L^2(0,T;H^\sigma)}\,.
\end{aligned} \tag{6.38}$$

*Proof.* **a)** For further reference, we first prove three estimates.

**1)** Let $\sigma_0 > d/2$ and $\sigma \in [0, \sigma_0]$. There exists a constant $K$ such that for all $\varrho \geqslant 0$ and for all $(u_1, u_2) \in H^{\sigma_0} \times H^\sigma_\varrho(\mathbb{D})$,

$$\|u_1 u_2\|_{H^\sigma_\varrho} \leqslant K\,\|u_1\|_{H^{\sigma_0}}\,\|u_2\|_{H^\sigma_\varrho}\,. \tag{6.39}$$

To prove this result, we use a standard Moser's estimate. Since $\sigma_0 > d/2$, the product maps continuously $H^{\sigma_0} \times H^r$ to $H^r$ for all $r \in [-\sigma_0, \sigma_0]$. Hence, we have

$$\begin{aligned}
\|u_1 u_2\|_{H^\sigma_\varrho} &:= \|u_1 u_2\|_{H^{\sigma-1}} + \varrho\,\|u_1 u_2\|_{H^\sigma} \\
&\lesssim \|u_1\|_{H^{\sigma_0}}\,(\|u_2\|_{H^{\sigma-1}} + \varrho\,\|u_2\|_{H^\sigma}) = \|u_1\|_{H^{\sigma_0}}\,\|u_2\|_{H^\sigma_\varrho}\,.
\end{aligned}$$

**2)** Let $\sigma \in [0, s]$ and $q \in [1, +\infty]$. Given a $C^\infty$ function $F$, there exists a generic function $C$ such that for all $\varrho \geqslant 0$, and for all $u \in L^q(0, T; H^\sigma_\varrho)$,

$$\|F(\psi)u\|_{L^q(0,T;H^\sigma_\varrho)} \lesssim \widetilde{C}\,\|u\|_{L^q(0,T;H^\sigma_\varrho)}\,, \tag{6.40}$$

where $\psi$ is as defined in (6.18).

In light of (6.39), to prove this estimate we need only show that

$$\|F(\psi)\|_{L^\infty(0,T;H^s)} \leqslant \widetilde{C}. \tag{6.41}$$

This will be established (independently) in (6.63) below.

**2')** Let us infer from (6.40) that, for all $\sigma \in [0, s]$,

$$\|F(\psi)u\|_{\mathcal{K}_\nu^\sigma(T)} \leqslant \widetilde{C}\,\|u\|_{\mathcal{K}_\nu^\sigma(T)}\,, \tag{6.42}$$

To see this, we write

$$\begin{aligned}
\|F(\psi)u\|_{\mathcal{K}_\nu^\sigma(T)} &:= \|F(\psi)u\|_{L^\infty(0,T;H^{\sigma-1})} + \nu\,\|F(\psi)u\|_{L^2(0,T;H^\sigma)} \\
&\leqslant \|F(\psi)u\|_{L^\infty(0,T;H^{\sigma-1})} + \|F(\psi)u\|_{L^2(0,T;H^\sigma_\nu)} \\
&\leqslant \widetilde{C}\big\{\|u\|_{L^\infty(0,T;H^{\sigma-1})} + \|u\|_{L^2(0,T;H^\sigma_\nu)}\big\} \leqslant \widetilde{C}\,\|u\|_{\mathcal{K}_\nu^\sigma(T)}\,.
\end{aligned}$$

**3)** Let $\sigma \in [0, s]$, $f \in H^{\sigma-1}(\mathbb{D})$ and $\gamma \colon \mathbb{D} \to (0, +\infty)$ be a function bounded from below by a positive constant. Furthermore, suppose $\tilde{\gamma} := \gamma - \underline{\gamma} \in H^s(\mathbb{D})$ for some constant $\underline{\gamma}$. We claim that if $u \in H^\sigma(\mathbb{D})$ satisfies $\operatorname{div}(\gamma\nabla u) = f$, then there exists a constant $K = K(d, s)$ such that

$$\|u\|_{H^{\sigma+1}} \leqslant K\,\|\gamma^{-1}\|_{L^\infty}\,\big\{\|f\|_{H^{\sigma-1}} + \|\tilde{\gamma}\|_{H^s}\,\|u\|_{L^2}\big\} + \|u\|_{L^2}\,. \tag{6.43}$$

Commuting the equation $\operatorname{div}(\gamma\nabla u) = f$ with $\Lambda^\sigma$ and using the commutator estimate (3.9) to bound $\operatorname{div}([\gamma, \Lambda^\sigma]\nabla u)$, we see that the proof of (6.43) can be reduced to the special case $\sigma = 0$. On the other hand, the case $\sigma = 0$ is immediate by the usual integration by parts and duality arguments.

**b)** Hereafter, RHS denotes the right-hand side of (6.38). To prove that

$$\|\operatorname{div}\widetilde{v}\|_{\mathcal{K}^\sigma_\nu(T)} \leqslant \text{RHS}, \tag{6.44}$$

we begin by showing that

$$\|\operatorname{div}\widetilde{v}\|_{\mathcal{K}^\sigma_\nu(T)} \leqslant \widetilde{C}\big\{\|(\varepsilon\partial_t)\widetilde{p}\|_{\mathcal{K}^\sigma_\nu(T)} + \|\kappa\widetilde{\theta}\|_{\mathcal{K}^{\sigma+2}_\nu(T)} + \varepsilon\|f_1\|_{\mathcal{K}^\sigma_\nu(T)}\big\}. \tag{6.45}$$

We rewrite $\operatorname{div}\widetilde{v}$ as

$$-g_1(\phi)(\varepsilon\partial_t)\widetilde{p} + \kappa\mathcal{E}_1(\phi)\Delta\widetilde{\theta} + \kappa\mathcal{E}_2(\phi,\nabla\phi)\nabla\widetilde{\theta} + \varepsilon f_1,$$

where $\mathcal{E}_1$ and $\mathcal{E}_2$ are as in (6.30). Using (6.42), we have

$$\|g_1(\phi)(\varepsilon\partial_t)\widetilde{p}\|_{\mathcal{K}^\sigma_\nu(T)} \leqslant \widetilde{C}\,\|(\varepsilon\partial_t)\widetilde{p}\|_{\mathcal{K}^\sigma_\nu(T)}\,,$$
$$\|\kappa\mathcal{E}_1(\phi)\Delta\widetilde{\theta}\|_{\mathcal{K}^\sigma_\nu(T)} \leqslant \widetilde{C}\,\|\kappa\widetilde{\theta}\|_{\mathcal{K}^{\sigma+2}_\nu(T)}\,.$$

To infer (6.45) it remains only to estimate the $\mathcal{K}^\sigma_\nu(T)$-norm of $\kappa\mathcal{E}_2(\phi,\nabla\phi)\nabla\widetilde{\theta}$. Using the very definition of $\mathcal{E}_2(\phi,\nabla\phi)$ and the estimate (6.41), we obtain

$$\|\mathcal{E}_2(\phi,\nabla\phi)\|_{L^\infty(0,T;H^{s-1})} \lesssim \widetilde{C}.$$

Since $s-1 > d/2$, the estimate (6.39) applies with $\sigma_0 := s-1$. It yields

$$\|\mathcal{E}_2(\phi,\nabla\phi)\nabla\widetilde{\theta}\|_{L^\infty(0,T;H^{\sigma-1})} \leqslant \widetilde{C}\,\|\nabla\widetilde{\theta}\|_{L^\infty(0,T;H^{\sigma-1})} \leqslant \widetilde{C}\,\|\widetilde{\theta}\|_{L^\infty(0,T;H^\sigma)}\,. \tag{6.46}$$

Moreover, one can easily verify that

$$\begin{aligned}
&\nu\,\|\mathcal{E}_2(\phi,\nabla\phi)\nabla\widetilde{\theta}\|_{L^2(0,T;H^\sigma)}\\
&\quad\leqslant \|\nu\mathcal{E}_2(\phi,\nabla\phi)\|_{L^2(0,T;H^s)}\,\|\nabla\widetilde{\theta}\|_{L^\infty(0,T;H^\sigma)}\\
&\quad\leqslant \sqrt{T}\,\|\mathcal{E}_2(\phi,\nabla\phi)\|_{L^\infty(0,T;H^s_\nu)}\,\|\widetilde{\theta}\|_{L^\infty(0,T;H^{\sigma+1})}\\
&\quad\leqslant \sqrt{T}C(\|\phi\|_{L^\infty(0,T;H^{s+1}_\nu)})\,\|\widetilde{\theta}\|_{L^\infty(0,T;H^{\sigma+1})}\\
&\quad\leqslant \sqrt{T}C(\Omega)\,\|\widetilde{\theta}\|_{L^\infty(0,T;H^{\sigma+1})} \leqslant \widetilde{C}\,\|\widetilde{\theta}\|_{L^\infty(0,T;H^{\sigma+1})}\,.
\end{aligned}$$

From this together with (6.46), we conclude that

$$\|\mathcal{E}_2(\phi,\nabla\phi)\nabla\widetilde{\theta}\|_{\mathcal{K}^\sigma_\nu(T)} \leqslant \widetilde{C}\,\|\widetilde{\theta}\|_{L^\infty(0,T;H^{\sigma+1})}\,.$$

This completes the proof of (6.45) since $\|\widetilde{\theta}\|_{L^\infty(0,T;H^{\sigma+1})} \leqslant \|\widetilde{\theta}\|_{\mathcal{K}^{\sigma+2}_\nu(T)}$. So to prove (6.44) it remains only to show that

$$\|\kappa\widetilde{\theta}\|_{\mathcal{K}^\sigma_\nu(T)} \leqslant \text{RHS}. \tag{6.47}$$

To see this, solve the first equation in (6.37) for $\operatorname{div} v$ and substitute the result in the third equation, to obtain

$$g_3(\phi)\partial_t\widetilde{\theta} - \kappa(B_3(\phi) - B_1(\phi))\widetilde{\theta} = f_3 + f_3', \tag{6.48}$$

with $f_3' := g_1(\phi)(\varepsilon\partial_t)\widetilde{p} - \varepsilon f_1$. Since

$$B_3(\phi) - B_1(\phi) := (\chi_3(\varepsilon p) - \chi_1(\varepsilon p)) \operatorname{div}(\beta(\theta)\nabla\cdot),$$

the assumption $\chi_1(\wp) < \chi_3(\wp)$ (which is (A3) in Assumption 2.1) implies that the equation (6.48) is parabolic. Hence, one can use the following estimate (see (9.2) below):

$$\begin{aligned} \|\kappa\widetilde{\theta}\|_{\mathcal{K}_\nu^{\sigma+2}(T)} &\leqslant \|\sqrt{\kappa}\nu\widetilde{\theta}\|_{L^\infty(0,T;H^{\sigma+1})} + \sqrt{\kappa}\,\|\sqrt{\kappa}\nu\widetilde{\theta}\|_{L^2(0,T;H^{\sigma+2})} \\ &\leqslant \widetilde{C}\big\{\|\sqrt{\kappa}\nu\widetilde{\theta}(0)\|_{H^{\sigma+1}} + \nu\,\|f_3 + f_3'\|_{L^2(0,T;H^\sigma)}\big\}. \end{aligned} \tag{6.49}$$

From which we easily infer the desired bound (6.47).

**c)** *Estimate for* $\nabla\widetilde{p}$. Note that, by combining (6.42) with (6.43), we have

$$\|\nabla\widetilde{p}\|_{\mathcal{K}_\nu^\sigma(T)} \leqslant \widetilde{C}\,\|\operatorname{div}(g_2(\phi)^{-1}\nabla\widetilde{p})\|_{\mathcal{K}_\nu^{\sigma-1}(T)} + \widetilde{C}\,\|\widetilde{p}\|_{L^\infty(0,T;L^2)}\,. \tag{6.50}$$

Starting from $\nabla\widetilde{p} = -g_2(\phi)(\varepsilon\partial_t)\widetilde{v} + \varepsilon\mu B_2(\phi)\widetilde{v} + \varepsilon f_2$, it is found that

$$\begin{aligned} \operatorname{div}(g_2(\phi)^{-1}\nabla\widetilde{p}) &= -(\varepsilon\partial_t)\operatorname{div}\widetilde{v} + \varepsilon\mu\mathcal{F}_1(\phi)\nabla^2\operatorname{div}\widetilde{v} \\ &\quad + \varepsilon\mu\mathcal{F}_2(\phi,\nabla\phi)\nabla^2\widetilde{v} + \varepsilon\mu\mathcal{F}_3(\phi,\nabla\phi,\nabla^2\phi)\nabla\widetilde{v} + \varepsilon\operatorname{div}(g_2(\phi)^{-1}f_2), \end{aligned} \tag{6.51}$$

for some $C^\infty$ functions $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$ vanishing at the origin.

The most direct estimates show that the $\mathcal{K}_\nu^{\sigma-1}(T)$-norms of the first and the last three terms in the right-hand side of (6.51) are bounded by RHS. So we only need to concentrate on the second term. Since $0 \leqslant \sigma - 1 \leqslant s$, the estimate (6.42) implies that

$$\|\varepsilon\mu\mathcal{F}_1(\phi)\nabla^2\operatorname{div}\widetilde{v}\|_{\mathcal{K}_\nu^{\sigma-1}(T)} \leqslant \widetilde{C}\,\|\varepsilon\mu\nabla^2\operatorname{div}\widetilde{v}\|_{\mathcal{K}_\nu^{\sigma-1}(T)}\,.$$

The following observation is important since this where the spectral localization of $\widetilde{U}$ enters. Since $\widetilde{v} = J_{\varepsilon\nu/3}\widetilde{v}$ and since $\varepsilon\mu\nabla^2 J_{\varepsilon\nu/3} \lesssim \Lambda^1$, we have

$$\|\varepsilon\mu\nabla^2\operatorname{div}\widetilde{v}\|_{\mathcal{K}_\nu^{\sigma-1}(T)} \lesssim \|\operatorname{div}\widetilde{v}\|_{\mathcal{K}_\nu^\sigma(T)}\,.$$

The piece of information already determined in (6.44) implies that this in turn is $\leqslant$ RHS. Hence, the right-hand side of (6.50) is $\leqslant$ RHS. The proof of (6.38) is complete. □

Recall that the purpose of the two previous lemmas is to estimate the fast components $\operatorname{div} J_{\varepsilon\nu}v$ and $\nabla J_{\varepsilon v}p$.

**Lemma 6.16.** *There exists a generic function $C(\cdot)$ such that*

$$\|J_{\varepsilon\nu}p\|_{L^\infty(0,T;H^s)} + \nu\,\|J_{\varepsilon\nu}p\|_{L^2(0,T;H^{s+1})} \leqslant \widetilde{C}, \tag{6.52}$$

$$\|\operatorname{div} J_{\varepsilon\nu}v\|_{L^\infty(0,T;H^{s-1})} + \nu\,\|\operatorname{div} J_{\varepsilon\nu}v\|_{L^2(0,T;H^s)} \leqslant \widetilde{C}. \tag{6.53}$$

**Remark 6.17.** Since we have already estimated the high frequency components (see Proposition 5.1), the previous lemma implies that

$$\begin{aligned} \|p\|_{L^\infty(0,T;H^s)} + \nu\,\|p\|_{L^2(0,T;H^{s+1})} &\leqslant \widetilde{C}, \\ \|\operatorname{div} v\|_{L^\infty(0,T;H^{s-1})} + \nu\,\|\operatorname{div} v\|_{L^2(0,T;H^s)} &\leqslant \widetilde{C}. \end{aligned} \tag{6.54}$$

*Proof.* To avoid repetition, we just give the scheme of the analysis. Set

$$X_m := \|p_m\|_{\mathcal{K}_\nu^{s-m+1}(T)} + \|\operatorname{div} v_m\|_{\mathcal{K}_\nu^{s-m}(T)} .$$

Setting $\widetilde{U} := U_m$ and $f := f_{\mathrm{LF}}^m - G(\phi)v \cdot \nabla U_m$, where $G(\phi)$ is as defined in (5.23), we are in the situation of the previous lemma and hence the estimate (6.38) easily implies that $X_m \leqslant \widetilde{C} X_{m+1} + Y_m$, where

$$\begin{aligned} Y_m \leqslant \widetilde{C}\big\{ \|U_m\|_{\mathcal{H}_a^0(T)} + \|U_m(0)\|_{\mathcal{H}_{a,0}^{s-m}} \big\} \\ + (\varepsilon + \sqrt{T})C(\Omega)\big\{ \|f_{\mathrm{LF}}^m\|_{L^\infty(0,T;\mathcal{H}_{a,0}^{s-1-m})} + \|U_m\|_{\mathcal{H}_a^{s-m}(T)} \big\}. \end{aligned}$$

Gathering the results of the previous lemma, one obtains

$$Y_m \leqslant \widetilde{C}\big\{ \widetilde{C} + C(\Omega_0)\big\} + (\varepsilon + \sqrt{T})C(\Omega)\big\{ C(\Omega) + C(\Omega)\big\} \leqslant \widetilde{C}.$$

Hence, we end up with $X_m \leqslant \widetilde{C} X_{m+1} + \widetilde{C}$. By an elementary induction, we therefore obtain $X_0 \leqslant \widetilde{C} X_s + \widetilde{C}$. Finally, noting that $X_s \lesssim \|U_s\|_{\mathcal{H}_a^0(T)}$ and using Lemma 6.13 leads to the desired bound $X_0 \leqslant \widetilde{C}$. □

6.5. **The slow components.** Having proved the estimates for the fast components, we now prove the estimates for the slow components $\theta$ and $\operatorname{curl} v$. To prove both estimates we use Assumption 2.2. This furnishes us with a $C^\infty$ function $S = S(\vartheta, \wp)$ such that $S(0,0) = 0$ and

$$\mathsf{d}S(\vartheta,\wp) = g_3(\vartheta,\wp)\,\mathsf{d}\vartheta - g_1(\vartheta,\wp)\,\mathsf{d}\wp, \tag{6.55}$$

$$(\vartheta,\wp) \mapsto (S(\vartheta,\wp),\wp) \text{ is a } C^\infty \text{ change of variables.} \tag{6.56}$$

Recall that $(\vartheta,\wp) \in \mathbb{R}^2$ is the place holder of $(\theta, \varepsilon p)$.

**Notation 6.18.** Set $\sigma := S(\theta, \varepsilon p)$.

One reason it is interesting to introduce the coordinate $S = S(\vartheta,\wp)$ is that $\sigma$ is well transported by the flow (see also Remarks 6.21 and 6.23).

**Lemma 6.19.** *Given $F \in C^\infty(\mathbb{R})$ satisfying $F(0) = 0$, there exists a generic function $C(\cdot)$ such that*

$$\|(\partial_t + v\cdot\nabla)F(\sigma)\|_{L^2(0,T;H^s)} \leqslant C(\Omega), \tag{6.57}$$

$$\|F(\sigma)\|_{L^\infty(0,T;H^s)} \leqslant \widetilde{C}. \tag{6.58}$$

*Proof.* Firstly, we form an evolution equation for $\sigma$. Directly from the identity (6.55), we have $\partial_{t,x}\sigma = g_3(\phi)\partial_{t,x}\theta - \varepsilon g_1(\phi)\partial_{t,x} p$. By combining the first and the last equations in (2.1) with this identity, we compute

$$\partial_t\sigma + v\cdot\nabla\sigma - \kappa\big(\chi_3(\varepsilon p) - \chi_1(\varepsilon p)\big)\operatorname{div}(\beta(\theta)\nabla\theta) = 0. \tag{6.59}$$

Therefore,

$$\partial_t\sigma + v\cdot\nabla\sigma = \kappa G_1(\phi,\nabla\theta) + \kappa G_2(\phi)\Delta\theta$$

for some $C^\infty$ functions $G_1$ and $G_2$, with $G_1(0) = 0$. Hence, the Moser estimate (5.8) implies at once that

$$\|\partial_t\sigma + v\cdot\nabla\sigma\|_{H^s} \leqslant C(\|\phi\|_{H^s})\big\{1 + \kappa\,\|\nabla\theta\|_{H^{s+1}}\big\}.$$

Since $\sqrt{\kappa} \leqslant \nu$, we have $\kappa \left\| \nabla\theta \right\|_{H^{s+2}} \leqslant \sqrt{\kappa} \left\| \nabla\theta \right\|_{H^{s+1}_\nu}$, so that

$$\left\| \partial_t \sigma + v \cdot \nabla \sigma \right\|_{H^s} \leqslant C(R)\{1 + R'\},$$

where $R$ and $R'$ are as defined in (6.19). Since $\Omega \approx \left\| R \right\|_{L^\infty(0,T)} + \left\| R' \right\|_{L^2(0,T)}$,

$$\left\| (\partial_t + v \cdot \nabla)\sigma \right\|_{L^2(0,T;H^s)} \leqslant C(\Omega). \tag{6.60}$$

The chain rule and the rule of product in Sobolev spaces imply that the left-hand side of (6.57) is estimated by

$$K(1 + \| \widetilde{F'}(\sigma) \|_{L^\infty(0,T;H^s)}) \left\| (\partial_t + v \cdot \nabla)\sigma \right\|_{L^2(0,T;H^s)}, \tag{6.61}$$

where $\widetilde{F'}(z) = F'(z) - F'(0)$ and $F'$ denotes the differential of $F$.

The first term in (6.61) is bounded by $C(\left\| \sigma \right\|_{L^\infty(0,T;H^s)}) \leqslant C(\Omega)$ (here is where we use the hypothesis that $S(0,0) = 0$). Hence, (6.57) follows from (6.60).

Moving to the proof of (6.58), we first recall an usual estimate in Sobolev spaces for hyperbolic equations (see the estimate (9.2) below with $\eta = 0$). For all functions $u, \chi \in C^0([0,T]; H^s(\mathbb{D}))$, with $s > 1 + d/2$,

$$\left\| u \right\|_{L^\infty(0,T;H^s)} \lesssim e^{TX} \left\| u(0) \right\|_{H^s} + \int_0^T e^{(T-t)X} \left\| \partial_t u + \chi \cdot \nabla u \right\|_{H^s} \, dt, \tag{6.62}$$

where $X := K \int_0^T \left\| \chi \right\|_{H^s} \, dt$ for some constant $K$ depending only on $(s, d)$. The Cauchy-Schwarz inequality readily implies that the right-hand side is estimated by

$$e^{KT\left\| \chi \right\|_{L^\infty(0,T;H^s)}} \big\{ \left\| u(0) \right\|_{H^s} + \sqrt{T} \left\| \partial_t u + \chi \cdot \nabla u \right\|_{L^2(0,T;H^s)} \big\}.$$

Applying this bound with $(u, \chi) = (F(\sigma), v)$, we find that the $L^\infty(0,T;H^s)$ norm of $F(\sigma)$ is estimated by

$$e^{KT\Omega} \big\{ C(\Omega_0) + \sqrt{T} \left\| \partial_t F(\sigma) + v \cdot \nabla F(\sigma) \right\|_{L^2(0,T;H^s)} \big\}.$$

We complete the proof of (6.58) by using (6.57). □

**Corollary 6.20.** *Given $F \in C^\infty(\mathbb{R}^{2+d})$ satisfying $F(0) = 0$, there exists a generic function $C(\cdot)$ such that*

$$\left\| F(\psi) \right\|_{L^\infty(0,T;H^s)} \leqslant \widetilde{C}. \tag{6.63}$$

*Proof.* The property (6.56) implies that there exists a $C^\infty$ function $F^*$ such that $F(\vartheta, \wp, \mathsf{v}) = F^*(S(\vartheta, \wp), \wp, \mathsf{v})$ for all $(\vartheta, \wp, \mathsf{v}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d$. Moreover, the hypothesis $S(0,0) = 0$ implies $F^*(0,0,0) = 0$.

We decompose $F(\psi) := F(\theta, \varepsilon p, \varepsilon v) = F^*(\sigma, \varepsilon p, \varepsilon v)$ as

$$F^*(\sigma, 0, 0) + \big\{ F^*(\sigma, \varepsilon p, \varepsilon v) - F^*(\sigma, 0, 0) \big\}.$$

The first term is estimated by way of (6.58). To bound the second term, we observe that one can factor out $(\varepsilon p, \varepsilon v)$. Consequently, there exists a function $C(\cdot)$ depending only on $F^*$ such that

$$\|F^*(\sigma,\varepsilon p,\varepsilon v) - F^*(\sigma,0,0)\|_{H^s} \leqslant \varepsilon C(\|(\theta,p,v)\|_{H^s}) \leqslant \varepsilon C(R) \leqslant \widetilde{C}.$$

This completes the proof. □

**Remark 6.21.** The estimate (5.21) implies $\|\partial_t F(\psi)\|_{L^2(0,T;H^s_\nu)} \leqslant C(\Omega)$. Just as in (4.22), it yields $\|F(\psi)\|_{L^\infty(0,T;H^s_\nu)} \leqslant C(\Omega_0) + \sqrt{T}C(\Omega) \leqslant \widetilde{C}$. Yet, this is weaker than (6.63) [indeed $H^s_0 = H^{s-1}$].

Now we really use the special feature of the low frequency analysis. The following lemma states that $\Lambda_{\varepsilon\nu}^{-1}\sigma$ satisfies parabolic-type estimates.

**Lemma 6.22.** *There exists a generic function $C(\cdot)$ such that*

$$\|\Lambda_{\varepsilon\nu}^{-1}\sigma\|_{L^\infty(0,T;H^{s+1}_\nu)} + \sqrt{\kappa}\,\|\Lambda_{\varepsilon\nu}^{-1}\sigma\|_{L^2(0,T;H^{s+2}_\nu)} \leqslant \widetilde{C}.$$

*Proof.* In light of the estimate $\|\sigma\|_{L^\infty(0,T;H^s)} \leqslant \widetilde{C}$ (see (6.58)), it suffices to prove that $\dot\sigma := \nu\Lambda_{\varepsilon\nu}^{-1}\nabla\sigma$ satisfies

$$\|\dot\sigma\|_{L^\infty(0,T;H^s)} + \sqrt{\kappa}\,\|\dot\sigma\|_{L^2(0,T;H^{s+1})} \leqslant \widetilde{C}. \tag{6.64}$$

Let us form a parabolic evolution equation for $\dot\sigma := \nu\Lambda_{\varepsilon\nu}^{-1}\nabla\sigma$. Writing the identity (6.55) in the form $\mathsf{d}\vartheta = c_1(\vartheta,\wp)\,\mathsf{d}S + c_2(\vartheta,\wp)\,\mathsf{d}\wp$ with $c_1 := 1/g_3^{-1}$ and $c_2 := g_1/g_3$, yields

$$\nabla\theta = c_1(\phi)\nabla\sigma + \varepsilon c_2(\phi)\nabla p. \tag{6.65}$$

Inserting this expression for $\nabla\theta$ into the equation (6.59), yields

$$\partial_t\sigma + v\cdot\nabla\sigma - \kappa k(\phi)\Delta\sigma = \kappa\mathsf{G}_3 + \kappa\mathsf{G}_4, \tag{6.66}$$

where $k := (\chi_3 - \chi_1)\beta c_1$ is a smooth positive function[3] and

$$\mathsf{G}_3 + \mathsf{G}_4 := G_3(\phi,\nabla\phi) + \varepsilon G_4(\phi,\nabla\phi)\Delta p,$$

for some $C^\infty$ functions $G_3$ and $G_4$, with $G_3(0) = 0$.

We compute

$$\partial_t\dot\sigma + v\cdot\nabla\dot\sigma - \kappa k(\phi)\Delta\dot\sigma = \mathcal{G}, \tag{6.67}$$

where the source term is given by

$$\begin{aligned}\mathcal{G} := &-\nu\Lambda_{\varepsilon\nu}^{-1}(\nabla v\cdot\nabla\sigma) + \nu\big[v,\Lambda_{\varepsilon\nu}^{-1}\big]\cdot\nabla\nabla\sigma\\ &+\kappa\Lambda_{\varepsilon\nu}^{-1}(\nabla k(\phi)\Delta\sigma) + \kappa\nu[k(\phi),\Lambda_{\varepsilon\nu}^{-1}]\Delta\nabla\sigma\\ &+\nu\kappa\Lambda_{\varepsilon\nu}^{-1}\nabla\mathsf{G}_3 + \nu\kappa\Lambda_{\varepsilon\nu}^{-1}\nabla\mathsf{G}_4.\end{aligned}$$

Let us give the scheme of the analysis. In light of the standard estimate (9.2) (given in the appendix below), to prove (6.64) we need only show that the source term $\mathcal{G}$ can be split as $\mathcal{G}_1 + \sqrt{\kappa}\mathcal{G}_2$ with

$$\|\mathcal{G}_1\|_{L^1(0,T;H^s)} \leqslant \widetilde{C} \quad\text{and}\quad \|\mathcal{G}_2\|_{L^2(0,T;H^{s-1})} \leqslant \widetilde{C}.$$

[3]By the hypotheses, $\chi_1 < \chi_3$, $0 < c_1 := 1/g_3$ and $0 < \beta$ (see Assumption 2.1).

To see this, there are only two nontrivial points. Firstly, note that

$$\begin{aligned}
&\|\nu\sqrt{\kappa}\Lambda_{\varepsilon\nu}^{-1}\nabla \mathsf{G}_4\|_{L^2(0,T;H^{s-1})}\\
&\quad := \|\nu\sqrt{\kappa}\Lambda_{\varepsilon\nu}^{-1}\nabla\{\varepsilon G_4(\phi,\nabla\phi)\Delta p\}\|_{L^2(0,T;H^{s-1})}\\
&\quad \lesssim \|\sqrt{\kappa}G_4(\phi,\nabla\phi)\Delta p\|_{L^2(0,T;H^{s-1})} && (\varepsilon\nu\nabla\Lambda_{\varepsilon\nu}^{-1}\lesssim I)\\
&\quad \lesssim (1+\|\tilde{G}_4(\phi,\nabla\phi)\|_{L^\infty(0,T;H^{s-1})})\,\|\sqrt{\kappa}p\|_{L^2(0,T;H^{s+1})} && \text{(straightforward)}\\
&\quad \lesssim \widetilde{C}\widetilde{C}. && \text{(by (6.63), (6.54))}
\end{aligned}$$

Secondly, set $I_{\mu,\kappa} := \nu/(\sqrt{\mu}+\sqrt{\kappa}) \leqslant 1$ and decompose $\nu\nabla v\cdot\nabla\sigma$ into two parts:

$$\nu\nabla v\cdot\nabla\sigma = I_{\mu,\kappa}\sqrt{\mu}\nabla v\cdot\nabla\sigma + I_{\mu,\kappa}\sqrt{\kappa}\nabla v\cdot\nabla\sigma.$$

Using the usual Moser estimates as well as the weighted versions (5.6)-(5.7), it is easily found that

$$\begin{aligned}
\sqrt{\mu}\,\|\nabla v\cdot\nabla\sigma\|_{H^s} &\lesssim \|\nabla v\|_{H^s_{\sqrt{\mu}}}\|\nabla\sigma\|_{H^s_{\sqrt{\mu}}} \lesssim (R+R')C(R),\\
\|\nabla v\cdot\nabla\sigma\|_{H^{s-1}} &\lesssim \|v\|_{H^s}\|\sigma\|_{H^s} \lesssim RC(R),
\end{aligned}$$

so that

$$\sqrt{\mu}\,\|\nabla v\cdot\nabla\sigma\|_{L^1(0,T;H^s)} + \|\nabla v\cdot\nabla\sigma\|_{L^2(0,T;H^{s-1})} \leqslant \sqrt{T}C(\Omega) \leqslant \widetilde{C}.$$

With these estimates established, the proof easily follows. □

**Remark 6.23.** One interesting feature of the equation (6.66) for $\sigma$ is that it is coupled to the momentum equation only through the convective term. Indeed, for the purpose of proving estimates independent of $\kappa$, we cannot see the term $\operatorname{div} v$ [in the equation for $\theta$] as a source term.

We denote by $\operatorname{curl} v$ the matrix with coefficients $(\operatorname{curl} v)_{ij} := \partial_j v_i - \partial_i v_j$. The basic idea of the forthcoming computations is to apply the curl operator to the equation for $v$ so as to cancel the large term $\varepsilon^{-1}\nabla p$. Yet, this requires some preparation because the factor $g_2(\theta,\varepsilon p)$ multiplying the time derivative of $v$ admits large oscillations in $O(1)$. To get around this, we follow the analysis of [30]. Namely, as in the proof of Corollary 6.20, we decompose $g_2(\theta,\varepsilon p)$ into two parts: the first which is well transported by the flow, the second which admits small oscillations of typical size $O(\varepsilon)$.

In particular, we do not estimate $\operatorname{curl} v$ directly. Instead we estimate $\operatorname{curl}(\gamma_0 v)$ where the coefficient $\gamma_0$ is defined as follows.

**Notation 6.24.** By (6.56), one can write $g_2(\vartheta,\wp) = \Gamma\big(S(\vartheta,\wp),\wp\big)$ for some $C^\infty$ positive function $\Gamma$. Set $\Gamma_0(\vartheta,\wp) := \Gamma\big(S(\vartheta,\wp),0\big)$ and

$$\gamma_0 := \Gamma_0(\theta,\varepsilon p). \tag{6.68}$$

**Lemma 6.25.** *There exists a generic function $C(\cdot)$ such that*

$$\|\operatorname{curl}(\gamma_0 v)\|_{L^\infty(0,T;H^{s-1})} + \sqrt{\mu}\,\|\operatorname{curl}(\gamma_0 v)\|_{L^2(0,T;H^s)} \leqslant \widetilde{C}. \tag{6.69}$$

*Proof.* We begin by computing the equation satisfied by $\omega := \operatorname{curl}(\gamma_0 v)$. It follows from elementary calculus that there exists a $C^\infty$ function $\Gamma_1$ such that for all $(\wp,\vartheta)\in\mathbb{R}^2$

$$\Gamma_0(\vartheta,\wp)/g_2(\vartheta,\wp) = 1+\wp\Gamma_1(\vartheta,\wp). \tag{6.70}$$

We first insert the expression for $g_2(\theta,\varepsilon p)$ given by (6.70) into the equation for $v$, thereby obtaining

$$\gamma_0(\partial_t v + v\cdot\nabla v) + \varepsilon^{-1}\nabla p - \mu B_2 v = F^\dagger,$$

with $F^\dagger := -p\gamma_1\nabla p + \varepsilon\mu p\gamma_1 B_2 v$ where $\gamma_1 := \Gamma_1(\theta,\varepsilon p)$. Consequently, the equation for $v$ is equivalent to

$$(\partial_t + v\cdot\nabla)(\gamma_0 v) + \varepsilon^{-1}\nabla p - \mu B_2 v = F^\dagger + (\partial_t\gamma_0 + v\cdot\nabla\gamma_0)v.$$

Using the elementary identity $\operatorname{curl}\nabla = 0$, we find

$$\big(\partial_t + v\cdot\nabla\big)\operatorname{curl}(\gamma_0 v) - \mu B_2\operatorname{curl} v = F^{\dagger\dagger},$$

where

$$F^{\dagger\dagger} := \operatorname{curl} F^\dagger - \big[\operatorname{curl}, v\big]\cdot\nabla(\gamma_0 v) + \mu\big[\operatorname{curl}, B_2\big]v + \operatorname{curl}\big(v(\partial_t\gamma_0 + v\cdot\nabla\gamma_0)\big).$$

Next, write $B_2\operatorname{curl} = \gamma_0^{-1}\gamma_0 B_2\operatorname{curl} = \gamma_0^{-1}B_2\operatorname{curl}(\gamma_0\cdot) + \gamma_0^{-1}\big[\gamma_0, B_2\operatorname{curl}\big]$ to obtain that $\omega = \operatorname{curl}(\gamma_0 v)$ satisfies

$$\partial_t\omega + v\cdot\nabla\omega - \mu\gamma_0^{-1}B_2\omega = F, \tag{6.71}$$

with $F := F^{\dagger\dagger} + \mu\gamma_0^{-1}[\gamma_0, B_2\operatorname{curl}]v$.

To sum up, $F = \sum_{1\leqslant i\leqslant 6} F_i$ with

$$\begin{aligned}
&F_1 := -\operatorname{curl}(p\gamma_1\nabla p), && F_4 := \mu[\operatorname{curl}, B_2]v,\\
&F_2 := \varepsilon\mu\operatorname{curl}(p\gamma_1 B_2 v), && F_5 := -[\operatorname{curl}, v]\cdot\nabla(\gamma_0 v),\\
&F_3 := \operatorname{curl}\big(v(\partial_t\gamma_0 + v\cdot\nabla\gamma_0)\big), && F_6 := \mu\gamma_0^{-1}[\gamma_0, B_2\operatorname{curl}]v.
\end{aligned}$$

*Estimate for $F$.* As in the proof of Lemma 6.22, to prove (6.69) it suffices to show that one can decompose $F$ as $f_1 + \sqrt{\mu}f_2$ with

$$\left\| f_1\right\|_{L^1(0,T;H^{s-1})} \leqslant \widetilde{C} \quad\text{and}\quad \left\| f_2\right\|_{L^2(0,T;H^{s-2})} \leqslant \widetilde{C}. \tag{6.72}$$

To do so we decompose $F$ as $f_1 + \sqrt{\mu}f_2$ with

$$\left\| f_1\right\|_{L^2(0,T;H^{s-1})} \leqslant C(\Omega) \quad\text{and}\quad \left\| f_2\right\|_{L^\infty(0,T;H^{s-2})} \leqslant C(\Omega). \tag{6.73}$$

[Note that (6.73) implies (6.72) since $\sqrt{T}C(\Omega)\leqslant\widetilde{C}$.]

Set $f_2 := \varepsilon\sqrt{\mu}p\gamma_1\operatorname{curl}(B_2 v) - \sqrt{\mu}\gamma_0^{-1}v\wedge B_2\nabla\gamma_0$ and $f_1 := F - f_2$. It follows from the very definition of $\Omega$ that $f_2$ satisfies the second estimate in (6.73). With regards to $f_1$, the key point is the following: Starting from

$$\left\| F_3\right\|_{H^{s-1}} \lesssim \left\| v\right\|_{H^s}\left\|\partial_t\gamma_0 + v\cdot\nabla\gamma_0\right\|_{H^s},$$

we obtain $\left\| F_3\right\|_{L^2(0,T;H^{s-1})} \lesssim \left\| v\right\|_{L^\infty(0,T;H^s)}\left\|\partial_t\gamma_0 + v\cdot\nabla\gamma_0\right\|_{L^2(0,T;H^s)}$. As a consequence, using the very definition of $\gamma_0$ (see (6.68)), the estimate (6.57) implies that $\left\| F_3\right\|_{L^2(0,T;H^{s-1})} \leqslant C(\Omega)$.

Let us estimate the other terms without repeated uses of the rules of product. To do so set $P := (p, \nabla p)$, $V := (v, \nabla v)$ and $\Theta := (\theta, \nabla\theta)$. Direct computations show that one can write $F_1 + F_2 + F_4 + F_5 + F_6 - f_2$ as

$$\sum (\sqrt{\mu}\partial_x)^\alpha G_{\alpha,a}(P, V, \Theta) \qquad (\alpha \in \mathbb{N}^d,\ |\alpha| \leqslant 1), \tag{6.74}$$

for some $C^\infty$ functions $G_{\alpha,a}$ such that the family $\{G_{\alpha,a} \,|\, a \in A, |\alpha| \leqslant 1\}$ is bounded in $C^\infty$, with in addition $G_{0,a}(0) = 0$. The Moser estimate (5.9) implies

$$\begin{aligned} \|(\sqrt{\mu}\partial_x)^\alpha G_{\alpha,a}(P, V, \Theta)\|_{H^{s-1}} &\leqslant \|G_{\alpha,a}(P, V, \Theta)\|_{H^s_{\sqrt{\mu}}} \\ &\leqslant C(\|(P, V, \Theta)\|_{L^\infty}) \,\|(P, V, \Theta)\|_{H^s_{\sqrt{\mu}}}\,. \end{aligned}$$

Hence, directly from the definition of $\Omega$, we arise at the desired bound

$$\|(\sqrt{\mu}\partial_x)^\alpha G_{\alpha,a}(P, V, \Theta)\|_{L^2(0,T;H^{s-1})} \leqslant C(\Omega),$$

which completes the proof of the claim (6.73). □

6.6. **The end of the proof of Proposition 6.1.** In order to prove Proposition 6.1 it remains to estimate $J_{\varepsilon\nu}\theta$ and $J_{\varepsilon\nu}v$.

**Lemma 6.26.** *Given $F \in C^\infty(\mathbb{R}^2)$ such that $F(0) = 0$, there exists a generic function $C(\cdot)$ such that*

$$\|J_{\varepsilon\nu}F(\phi)\|_{L^\infty(0,T;H^{s+1}_\nu)} + \sqrt{\kappa}\,\|J_{\varepsilon\nu}\nabla F(\phi)\|_{L^2(0,T;H^{s+1}_\nu)} \leqslant \widetilde{C}.$$

**Remark 6.27.** With $F(\phi) = \theta$ this is the expected bound for $J_{\varepsilon\nu}\theta$.

*Proof.* In light of (6.63), it suffices to prove that

$$\nu\,\|J_{\varepsilon\nu}\nabla F(\phi)\|_{L^\infty(0,T;H^s)} + \nu\sqrt{\kappa}\,\|J_{\varepsilon\nu}\nabla^2 F(\phi)\|_{L^2(0,T;H^s)} \leqslant \widetilde{C}. \tag{6.75}$$

By (6.56) we have

$$\nabla F(\phi) = F_1(\phi)\nabla\sigma + \varepsilon F_2(\phi)\nabla p, \tag{6.76}$$

for some $C^\infty$ functions $F_1$ and $F_2$. Applying Proposition 3.2 with $\sigma_0 = s$, $\sigma_1 = \sigma_2 = m_1 = 0$ and $m_2 = 1$ leads to

$$\begin{aligned} \|\nu\nabla\Lambda_{\varepsilon\nu}^{-1}F(\phi)\|_{H^s} \lesssim{}& \big(1 + \|\tilde{F}_1(\phi)\|_{H^s}\big)\,\|\nabla\nu\Lambda_{\varepsilon\nu}^{-1}\sigma\|_{H^s} \\ &+ \big(1 + \|\tilde{F}_2(\phi)\|_{H^s}\big)\,\|\nabla\varepsilon\nu\Lambda_{\varepsilon\nu}^{-1}p\|_{H^s}\,, \end{aligned}$$

where $\tilde{F} := F - F(0)$. Using $\nu\,\|\cdot\|_{H^s} \leqslant \|\cdot\|_{H^s_\nu}$, $\varepsilon\nu\nabla\Lambda_{\varepsilon\nu}^{-1} \lesssim I$ and (6.63), we obtain

$$\|\nu\nabla\Lambda_{\varepsilon\nu}^{-1}F(\phi)\|_{L^\infty(0,T;H^s)} \leqslant \widetilde{C}\big\{\|\Lambda_{\varepsilon\nu}^{-1}\sigma\|_{L^\infty(0,T;H^{s+1}_\nu)} + \|p\|_{L^\infty(0,T;H^s)}\big\}.$$

As a consequence, from Lemma 6.22 and (6.54), we get the first half of (6.75), namely: $\|\nu J_{\varepsilon\nu}\nabla F(\phi)\|_{L^\infty(0,T;H^s)} \leqslant \widetilde{C}$.

The technique for obtaining the second half is similar. We differentiate (6.76) to obtain: $\partial_i\partial_j F(\phi) := \mathcal{F}_I + \mathcal{F}_{II}$ with

$$\mathcal{F}_I := F_1(\phi)\partial_i\partial_j\sigma + \varepsilon F_2(\phi)\partial_i\partial_j p,$$
$$\mathcal{F}_{II} := \partial_i F_1(\phi)\partial_j\sigma + \varepsilon\partial_i F_2(\phi)\partial_j p.$$

Exactly as above, we have

$$\nu\sqrt{\kappa}\left\|J_{\varepsilon\nu}\mathcal{F}_I\right\|_{L^2(0,T;H^{s+2})}$$
$$\lesssim \widetilde{C}\big\{\nu\sqrt{\kappa}\left\|\Lambda_{\varepsilon\nu}^{-1}\sigma\right\|_{L^2(0,T;H^{s+2})} + \sqrt{\kappa}\left\|p\right\|_{L^2(0,T;H^{s+1})}\big\} \leqslant \widetilde{C}\widetilde{C}.$$

In order to estimate $\mathcal{F}_{II}$, note that $\mathcal{F}_{II} := G(\phi,\nabla\phi)$ for some $C^\infty$ function $G$ such that $G(0)=0$. Hence,

$$\begin{aligned}\nu\sqrt{\kappa}\left\|\mathcal{F}_{II}\right\|_{L^2(0,T;H^s)} &\lesssim \sqrt{T}\left\|G(\phi,\nabla\phi)\right\|_{L^\infty(0,T,H^s_\nu)} && \text{(straightforward)}\\ &\lesssim \sqrt{T}C\big(\left\|\phi\right\|_{L^\infty(0,T;H^{s+1}_\nu)}\big) && \text{(by (5.7))}\\ &\lesssim \sqrt{T}C(\Omega) \leqslant \widetilde{C} && \text{(by definition).}\end{aligned}$$

This completes the proof. □

The next estimate finishes the proof of Proposition 6.1.

**Lemma 6.28.** *There exists a generic function $C(\cdot)$ such that*

$$\left\|J_{\varepsilon\nu}v\right\|_{L^\infty(0,T;H^{s+1}_{\varepsilon\nu})} + \sqrt{\mu}\left\|J_{\varepsilon\nu}\nabla v\right\|_{L^2(0,T;H^{s+1}_{\varepsilon\nu})} \leqslant \widetilde{C}.$$

*Proof.* We need only explain how to combine all the previous Lemma.

**a)** To shorten notations, given a function $f$ such that $f - \underline{f} \in H^\sigma$ for some constant $\underline{f}$, we denote by $\left\|f\right\|_{\dot{H}^\sigma}$ the norm $\left\|f-\underline{f}\right\|_{H^\sigma}$. The semi-norm $\left\|\cdot\right\|_{\dot{H}^\sigma_\nu}$ is defined in the same way.

We claim that for all smooth positive functions $\gamma$, there exists a constant $K$ such that for all $(\varepsilon,\nu)$ and for all vector fields $u \in H^s(\mathbb{D})$, we have

$$\begin{aligned}\left\|J_{\varepsilon\nu}u\right\|_{H^s} \leqslant KM^{2s+1}\big\{&\left\|\operatorname{div} J_{\varepsilon\nu}u\right\|_{H^{s-1}} + \left\|\operatorname{curl}(\gamma u)\right\|_{H^{s-1}}\\ &+ \left\|J_{\varepsilon\nu}u\right\|_{L^2} + \varepsilon\left\|u\right\|_{H^s}\big\},\end{aligned} \tag{6.77}$$

where $M := 1 + \left\|\gamma^{-1}\right\|_{\dot{H}^s} + \left\|\gamma\right\|_{\dot{H}^{s+1}_\nu}$.

To prove this result, we start from the following estimate: for all $\sigma \geqslant 0$,

$$\left\|u\right\|_{H^{\sigma+1}} \lesssim \left\|\operatorname{div} u\right\|_{H^\sigma} + \left\|\operatorname{curl} u\right\|_{H^\sigma} + \left\|u\right\|_{L^2}. \tag{6.78}$$

As $\operatorname{curl} u = \gamma^{-1}\operatorname{curl}(\gamma u) - \gamma^{-1}\nabla\gamma\wedge u$, by the usual rule of product in Sobolev spaces, we infer that for all $\sigma\in[0,s-1]$,

$$\begin{aligned}\left\|\operatorname{curl} u\right\|_{H^\sigma} &\lesssim (1+\left\|\gamma^{-1}\right\|_{\dot{H}^s})\{\left\|\operatorname{curl}(\gamma u)\right\|_{H^\sigma} + \left\|\gamma\right\|_{\dot{H}^s}\left\|u\right\|_{H^\sigma}\}\\ &\lesssim M\left\|\operatorname{curl}(\gamma u)\right\|_{H^\sigma} + M^2\left\|u\right\|_{H^\sigma}.\end{aligned}$$

Thus, the estimate (6.78) turns into

$$\left\|u\right\|_{H^{\sigma+1}} \lesssim M^2\big\{\left\|\operatorname{div} u\right\|_{H^\sigma} + \left\|\operatorname{curl}(\gamma u)\right\|_{H^\sigma} + \left\|u\right\|_{H^\sigma}\big\}.$$

By induction on $\mathbb{N} \ni \sigma \leqslant s-1$, this yields

$$\|u\|_{H^s} \lesssim M^{2s}\big\{\|\operatorname{div} u\|_{H^{s-1}} + \|\operatorname{curl}(\gamma u)\|_{H^{s-1}} + \|u\|_{L^2}\big\}. \tag{6.79}$$

Furthermore, by (3.10) applied with $(m,\sigma_0,\sigma)=(1,s+1,s)$, we have

$$\begin{aligned}\|\operatorname{curl}(\gamma J_{\varepsilon\nu} u)\|_{H^{s-1}} &\leqslant \|J_{\varepsilon\nu}\operatorname{curl}(\gamma u)\|_{H^{s-1}} + \|[\gamma, J_{\varepsilon\nu}]u\|_{H^s}\\ &\lesssim \|\operatorname{curl}(\gamma u)\|_{H^\sigma} + \varepsilon\nu\,\|\gamma\|_{\dot H^{s+1}}\,\|u\|_{H^s}\,.\end{aligned}$$

From the estimate $\nu\,\|\gamma\|_{\dot H^{s+1}} \leqslant \|\gamma\|_{\dot H^{s+1}_\nu} \leqslant M$ and (6.79) applied with $u$ replaced by $J_{\varepsilon\nu}u$, the previous fact implies (6.77).

**b)** Let $\gamma := \gamma_0$ (as defined in (6.68)) and let $M := 1 + \|\gamma^{-1}\|_{\dot H^s} + \|\gamma\|_{\dot H^{s+1}_\nu}$. We list our bounds:

| | |
|---|---|
| by Lemma 6.26 : | $\|M\|_{L^\infty(0,T)} \leqslant \widetilde{C}$, |
| by (6.53) : | $\|\operatorname{div} J_{\varepsilon\nu} v\|_{L^\infty(0,T;H^{s-1})} \leqslant \widetilde{C}$, |
| by Lemma 6.25 : | $\|\operatorname{curl}(\gamma_0 v)\|_{L^\infty(0,T;H^{s-1})} \leqslant \widetilde{C}$, |
| by Lemma 6.13(with $m=0$) : | $\|J_{\varepsilon\nu} v\|_{L^\infty(0,T;L^2)} \leqslant \widetilde{C}$, |
| directly from the definitions : | $\|\varepsilon v\|_{L^\infty(0,T;H^s)} \leqslant \varepsilon\Omega \leqslant \widetilde{C}$. |

Hence, we deduce from (6.77) that $\|J_{\varepsilon\nu} v\|_{L^\infty(0,T;H^s)} \leqslant \widetilde{C}$. This in turn implies $\|J_{\varepsilon\nu} v\|_{L^\infty(0,T;H^{s+1}_{\varepsilon\nu})} \leqslant \widetilde{C}$, indeed recall that $\|J_{\varepsilon\nu} v\|_{H^{\sigma+1}_{\varepsilon\nu}} \lesssim \|J_{\varepsilon\nu} v\|_{H^\sigma}$.

**c)** By (6.78), (6.53) and Lemma 6.13, to prove $\sqrt{\mu}\,\|J_{\varepsilon\nu}\nabla v\|_{L^2(0,T;H^{s+1}_{\varepsilon\nu})} \leqslant \widetilde{C}$ we need only check that $\sqrt{\mu}\,\|\operatorname{curl} J_{\varepsilon\nu} v\|_{L^2(0,T;H^s)} \leqslant \widetilde{C}$. For this purpose, write

$$\|\operatorname{curl} J_{\varepsilon\nu} v\|_{H^s} \lesssim (1 + \|\gamma_0^{-1}\|_{\dot H^s})\{\|\operatorname{curl}(\gamma_0 J_{\varepsilon\nu} v)\|_{H^s} + \|\gamma_0\|_{\dot H^{s+1}}\,\|v\|_{H^s}\},$$

and

$$\|\operatorname{curl}(\gamma_0 J_{\varepsilon\nu} v)\|_{H^s} \lesssim \|\operatorname{curl}(\gamma_0 v)\|_{H^s} + \|\gamma_0\|_{\dot H^{s+1}}\,\|v\|_{H^s}\,,$$

where we used the commutator estimate (3.10) with $(m,\sigma_0,\sigma)=(0,s+1,s)$. These two facts and the estimate $\sqrt{\mu}\,\|\gamma_0\|_{L^\infty(0,T;\dot H^{s+1})} \leqslant \|\gamma_0\|_{L^\infty(0,T;\dot H^{s+1}_\nu)} \leqslant \widetilde{C}$ (see Lemma 6.26) imply that

$$\sqrt{\mu}\,\|\operatorname{curl} J_{\varepsilon\nu} v\|_{L^2(0,T;H^s)} \leqslant \widetilde{C}\big\{\sqrt{\mu}\,\|\operatorname{curl}(\gamma_0 v)\|_{L^2(0,T;H^s)} + \|v\|_{L^2(0,T;H^s)}\big\}.$$

Hence, the expected bound directly follows from Lemma 6.25 and the estimate $\|v\|_{L^2(0,T;H^s)} \leqslant \sqrt{T}\,\|v\|_{L^\infty(0,T;H^s)} \leqslant \widetilde{C}$. □

## 7. Uniform stability

In this section we complete the proof of Theorem 2.7.

7.1. **Local existence result.** Granted the uniform *a priori* bounds proved in the previous sections, the end of the proof of Theorem 2.7 essentially reduces to establishing a local existence result for fixed $a \in A$.

For fixed $a \in A$, the fact that the Cauchy problem for (1.1) is well-posed is immediate provided that one chooses to work with the unknown $(\rho, v, \mathcal{T})$. With this choice of dependent variables the system is a coupled hyperbolic/parabolic system in symmetric form, so that the general theory applies. We shall show that this remains valid for the system (2.1).

Introduce first a definition clarifying the structure of the systems we produce. Consider a system of nonlinear equations

$$A_0(u)\partial_t u + \sum_{\alpha \in \mathbb{N}^d,\, 1 \leqslant |\alpha| \leqslant 2} A_\alpha(u)\partial_x^\alpha u = F(u, \nabla u), \tag{7.1}$$

where each $A_\beta$ $(0 \leqslant |\beta| \leqslant 2)$ is a $n \times n$ matrix smooth in its arguments and furthermore symmetric: $A_\beta = A_\beta^t$. We suppose $F$ is smooth in its arguments, with values in $\mathbb{R}^n$.

**Definition 7.1.** *The system is said to be of coupled hyperbolic/parabolic type provided that there exists* $(n_1, n_2) \in \mathbb{N}^2$*, and a splitting of the unknowns* $u = (u^1, u^2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ *so that* $F(u, \nabla u) = (0, F^2(u, \nabla u))^t$ *and*

$$A_0(u) = \begin{pmatrix} A_0^{11}(u) & 0 \\ 0 & A_0^{22}(u) \end{pmatrix}, \quad A_\alpha(u) = \begin{pmatrix} 0 & 0 \\ 0 & A_\alpha^{22}(u) \end{pmatrix} \quad \text{for all } |\alpha| = 2,$$

*where in addition*

$$\forall (u, \xi) \in \mathbb{R}^n \times \mathbb{S}^{d-1}, \qquad A_0^{11}(u) > 0, \quad A_0^{22}(u) > 0, \quad -\sum_{|\alpha|=2} A_\alpha^{22}(u)\xi^\alpha > 0.$$

*(here $M^{jk}$ denote the sub-blocks of the matrix $M$ which correspond to the splitting $\mathbb{R}^n \ni u = (u^1, u^2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$).*

The general theory [28, 41] for hyperbolic or parabolic problems applies for coupled hyperbolic/parabolic ones. One can prove the following results.

**Proposition 7.2.** *Let* $s > 2 + d/2$*. Assume that* (7.1) *is of coupled hyperbolic/parabolic type. Then for all initial data* $u_0$ *in* $H^s(\mathbb{D})$ *there exists a positive* $T = T(\|u_0\|_{H^s})$ *such that the Cauchy problem for* (7.1) *has a unique classical solution* $\Phi(u_0) \in C^0([0,T]; H^s(\mathbb{D}))$ *such that* $\Phi^2(u_0) \in L^2(0,T; H^{s+1}(\mathbb{D}))$ *where* $\Phi^2(u_0)$ *are the last* $n_2$ *components of* $\Phi(u_0)$ *where* $n_2$ *is as in Definition 7.1.*

**Proposition 7.3.** *With notations as in Proposition 7.2, the interval* $[0, T^\star)$ *with* $T^\star < +\infty$*, is a maximal interval of* $H^s$ *existence if and only if*

$$\limsup_{t \to T^\star} \|\Phi(u_0)(t)\|_{W^{1,\infty}(\mathbb{D})} = +\infty.$$

**Corollary 7.4.** *Let $u$ be a classical solution of* (7.1)*. If* $u(0) \in H^\infty(\mathbb{D})$ *then* $u \in C^\infty([0,T]; H^\infty(\mathbb{D}))$.

Fix $a = (\varepsilon, \mu, \kappa) \in A$. We next show that an appropriate change of variables transforms System (2.1) into a system of coupled hyperbolic/parabolic type. Let $\varrho$ denote the function given by Assumption 2.2. Since the mapping $\mathbb{R}^2 \ni (\vartheta, \wp) \mapsto (\vartheta, \varrho(\vartheta, \wp)) \in \mathbb{R}^2$ is a $C^\infty$ diffeomorphism, it is equivalent to work with $(p, v, \theta)$ or with $(\rho, v, \theta)$ where $\rho := \varrho(\theta, \varepsilon p)$.

Firstly, we form an evolution equation for $\rho := \varrho(\theta, \varepsilon p)$. By combining the first and the last equations in (2.1) with the second identity in (2.2) written in the form $\chi_3 \,\mathsf{d}\varrho = \chi_3 g_1 \,\mathsf{d}\wp - \chi_1 g_3 \,\mathsf{d}\vartheta$, we get

$$\chi_3(\partial_t \rho + v \cdot \nabla \rho) + (\chi_3 - \chi_1) \operatorname{div} v = 0. \tag{7.2}$$

The second identity in (2.2) written in the form

$$\mathsf{d}\wp = \gamma_1 \,\mathsf{d}\vartheta + \gamma_2 \,\mathsf{d}\varrho := \frac{\chi_1 g_3}{\chi_3 g_1} \,\mathsf{d}\vartheta + \frac{1}{g_1} \,\mathsf{d}\varrho,$$

yields $\varepsilon \nabla p = \gamma_1 \nabla \theta + \gamma_2 \nabla \rho$. Summing up, we have

$$\begin{cases} \chi_3(\partial_t \rho + v \cdot \nabla \rho) + (\chi_3 - \chi_1) \operatorname{div} v = 0, \\ g_2(\partial_t v + v \cdot \nabla v) + \varepsilon^{-2} \gamma_1 \nabla \theta + \varepsilon^{-2} \gamma_2 \nabla \rho - \mu B_2 v = 0, \\ g_3(\partial_t \theta + v \cdot \nabla \theta) + \operatorname{div} v - \kappa B_3 \theta = 0, \end{cases} \tag{7.3}$$

where all the coefficients are positive by Assumption 2.1. We also mention that all the coefficients are evaluated at $(\theta, \varepsilon p) = (\theta, \mathsf{P}(\theta, \rho))$ for some $C^\infty$ function $\mathsf{P}$. We next symmetrize the system (7.3). To do so, we multiply the first equation by $(\chi_3 - \chi_1)^{-1} \gamma_1$, the second by $\varepsilon^2$ and the third one by $\gamma_2$, to obtain

$$\begin{cases} \delta_1(\partial_t \rho + v \cdot \nabla \rho) + \gamma_1 \operatorname{div} v = 0, \\ \delta_2(\partial_t v + v \cdot \nabla v) + \gamma_1 \nabla \rho + \gamma_2 \nabla \theta - \varepsilon^2 \mu B_2 v = 0, \\ \delta_3(\partial_t \theta + v \cdot \nabla \theta) + \gamma_2 \operatorname{div} v - \kappa \gamma_2 B_3 \theta = 0, \end{cases} \tag{7.4}$$

with $\delta_1 := (\chi_3 - \chi_1)^{-1} \gamma_1 \chi_3$, $\delta_2 := \varepsilon^2 g_2$ and $\delta_3 := \gamma_2 g_3$ (compare with the system (4.15)). By Assumption 2.1, the coefficients $\delta_i$'s are positive, so that:

**Lemma 7.5.** *For all $(\mu, \kappa) \in [0,1]^2$, the symmetric system* (7.4) *is of coupled hyperbolic/parabolic type.*

Therefore, we obtain that the Cauchy problem for (7.4) is well-posed for all fixed $a \in A$, so is the Cauchy problem for (2.1) since (7.4) has been deduced from (2.1) by using a $C^\infty$ diffeomorphism.

7.2. **The end of the proof of Theorem 2.7.** Let $s > 1 + d/2$ and $M_0 > 0$. The previous analysis implies that, for all $a \in A$, there exists $T_a = T_a(M_0)$ such that for all initial data $U_0$ in the ball $B(\mathcal{H}^s_{a,0}, M_0)$, the Cauchy problem for (2.1) has a unique classical solution $U_a$ in $\mathcal{H}^s_a(T_a)$. We denote by $T_a^\star$ the maximal time of existence of such a classical solution. For all $t < T_a^\star$, set $\Omega_a(t) := \|U_a\|_{\mathcal{H}^s_a(t)}$. Our task is to show that there exists $T > 0$ and $M < +\infty$ such that

$$\forall a \in A, \qquad T_a^\star \geqslant T \quad \text{and} \quad \Omega_a(T) \leqslant M. \tag{7.5}$$

Up to replacing the initial data $U_0$ by $J_\delta U_0$, letting $\delta$ goes to zero and using a continuity argument for solutions of coupled hyperbolic/parabolic systems, we can assume that $U_0 \in H^\infty(\mathbb{D})$. In light of Corollary 7.4, it means that we can assume $U_a \in C^\infty([0, T_a^\star); H^\infty(\mathbb{D}))$. Hence, we can apply Propositions 5.1 and 6.1. We thus know that there exists a smooth nondecreasing function $C(\cdot)$ such that for all $a \in A$ and all $t < \min\{T_a^\star, 1\}$,

$$\Omega_a(t) \leqslant C(\Omega_a(0)) e^{(\sqrt{t}+\varepsilon)C(\Omega_a(t))}. \tag{7.6}$$

Choose $M_1 > C(M_0)$ and next $T_1 \leqslant 1$ and $\varepsilon_1 \leqslant 1$ such that

$$C(M_0) e^{(\sqrt{T_1}+\varepsilon_1)C(M_1)} < M_1. \tag{7.7}$$

Let $t < \min\{T_a^\star, T_1\}$ and $a = (\varepsilon, \mu, \kappa) \in A$ be such that $\varepsilon \leqslant \varepsilon_1$. By combining the inequalities (7.6) and (7.7) with the hypothesis $\Omega_a(0) < M_0$, we infer that $\Omega_a(t) \neq M_1$. Besides, we can assume without restriction that $M_0 < M_1$, so that $\Omega_a(0) < M_1$. Since the function $\Omega_a$ is continuous, we infer $\Omega_a(t) < M_1$. Consequently, the continuation principle (which is Proposition 7.3) shows that $T_a^\star > T_1$.

On the other hand the (omitted) proof of Proposition 7.2 implies that there exists $T_2 > 0$ and $M_2 < +\infty$ such that for all $a = (\varepsilon, \mu, \kappa) \in A$ with $\varepsilon \geqslant \varepsilon_1$, we have $T_a^\star \geqslant T_2$ and $\Omega_a(T_2) \leqslant M_2$.

We have proved (7.5) with $T := \min\{T_1/2, T_2\}$ and $M := \max\{M_1, M_2\}$, which completes the proof of Theorem 2.7.

## 8. Decay of the local energy

We will consider systems which include (2.1) as a special case. The motivation is to consider a structure general enough to include, say, the combustion equations. To do so we allow the limit constraint on the divergence of the velocity field to read $\operatorname{div} v = F(D^m\psi)$ where $\psi$ is the slow variable (namely $\partial_t \psi = O(1)$), $m$ is a given integer and

$$D^m\psi := \left\{ \partial_x^\alpha \psi \mid \alpha \in \mathbb{N}^d,\ |\alpha| \leqslant m \right\}.$$

More precisely, we consider systems of the form:

$$\begin{cases} g_1(\psi)\partial_t p + \dfrac{1}{\varepsilon} \operatorname{div} v = \dfrac{1}{\varepsilon} F_1(D^m\psi) + F_2(D^m\psi, D^m p, D^m v), \\ g_2(\psi)\partial_t v + \dfrac{1}{\varepsilon}\nabla p = F_3(D^m\psi, D^m p, D^m v), \\ \partial_t \psi = F_4(D^m\psi, D^m p, D^m v), \end{cases} \tag{8.1}$$

where $(p, v, \psi)$ is defined on $[0, T] \times \mathbb{R}^d$ with values in $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^n$, $n \geqslant 0$ and $m \geqslant 0$ are given integers.

**Assumption 8.1.** The functions $g_i$ $(i = 1, 2)$ and $F_j$ $(j = 1, 2, 3, 4)$ are smooth in their arguments, satisfying

$$g_1(y), g_2(y) > 0, \quad F_1(Y), F_2(Y, Z) \in \mathbb{R}, \quad F_3(Y, Z) \in \mathbb{R}^d, \quad F_4(Y, Z) \in \mathbb{R}^n,$$

where $y$, $Y$ and $Z$ are the place holders of $\psi$, $D^m\psi$ and $(D^m p, D^m v)$. Moreover $F_j(0) = 0$ for $j = 1, 2, 3, 4$.

Assuming that the slow variable decays sufficiently rapidly at spatial infinity we prove that the penalized terms converge to 0.

**Proposition 8.2.** *Let $T > 0$, $m \geqslant 1$ and $s > 3m + d/2$. Assume that the functions $(p^\varepsilon, v^\varepsilon, \psi^\varepsilon)$ satisfy (8.1) and*

$$\|(p^\varepsilon, v^\varepsilon, \psi^\varepsilon)\|_{\mathcal{K}^s(T)} := \sup_{t\in[0,T]} \|(p^\varepsilon, v^\varepsilon)(t)\|_{H^s} + \|\psi^\varepsilon(t)\|_{H^{s+m-1}} < +\infty. \tag{8.2}$$

*Assume further that $\psi^\varepsilon$ converges strongly in $C^0([0,T]; H^\sigma_{loc}(\mathbb{R}^d))$, for some $\sigma > 1 + d/2$, to a limit $\psi$ satisfying, for all $(t,x) \in [0,T] \times \mathbb{R}^d$,*

$$|\psi(t,x)| \leqslant K\,|x|^{-1-\gamma}, \quad |\nabla\psi(t,x)| \leqslant K\,|x|^{-2-\gamma}, \tag{8.3}$$

*for some positive constants $K$ and $\gamma$.*

*Then, $p^\varepsilon \to 0$ strongly in $L^2(0,T; H^{s'}_{loc}(\mathbb{R}^d))$ for all $s' < s$, and $\operatorname{div} v^\varepsilon - F_1(D^m\psi^\varepsilon) \to 0$ strongly in $L^2(0,T; H^{s'}_{loc}(\mathbb{R}^d))$ for all $s' < s - 1$.*

*Proof.* Since $(p^\varepsilon, v^\varepsilon, \psi^\varepsilon)$ is uniformly bounded in $\mathcal{K}^s(T)$ (see (8.2)), using an interpolation argument, it is sufficient to prove that $p^\varepsilon$ and $\operatorname{div} v^\varepsilon - F_1(D^m\psi^\varepsilon)$ converges strongly to 0 in $L^2(0,T; L^2_{loc}(\mathbb{R}^d))$. To prove this result we use the following Theorem proved in [30] (although not explicitly stated in this way). To clarify matters, we mention that: (1) all the functions considered in the following statement are real-valued; (2) all the convergences considered in this proof are strong.

**Theorem 8.3** (from Métivier & Schochet [30]). *Let $T > 0$ and let $u^\varepsilon$ be a bounded sequence in $C^0([0,T]; H^2(\mathbb{R}^d))$ such that*

$$\varepsilon^2\partial_t(a^\varepsilon\partial_t u^\varepsilon) - \operatorname{div}(b^\varepsilon\nabla u^\varepsilon) = c^\varepsilon,$$

*where the source term $c^\varepsilon$ converges to 0 in $L^2(0,T; L^2(\mathbb{R}^d))$. Assume further that the coefficients $(a^\varepsilon, b^\varepsilon)$ are uniformly bounded in $C^0([0,T]; H^\sigma(\mathbb{R}^d))$, for some $\sigma > 1 + d/2$, and converges in $C^0([0,T]; H^\sigma_{loc}(\mathbb{R}^d))$ to a limit $(a,b)$ satisfying the decay estimates*

$$|a(t,x) - \underline{a}| \leqslant K\,|x|^{-1-\gamma}, \quad |\nabla a(t,x)| \leqslant K\,|x|^{-2-\gamma},$$
$$|b(t,x) - \underline{b}| \leqslant K\,|x|^{-1-\gamma}, \quad |\nabla b(t,x)| \leqslant K\,|x|^{-2-\gamma},$$

*for some given positive constants $\underline{a}$, $\underline{b}$, $K$ and $\gamma$.*

*Then, the sequence $u^\varepsilon$ converges to 0 in $L^2(0,T; L^2_{loc}(\mathbb{R}^d))$.*

We can directly apply Theorem 8.3 to prove the first half of Proposition 8.2, that is, the convergence of $p^\varepsilon$ to 0 in $L^2(0,T; L^2_{loc}(\mathbb{R}^d))$. Indeed, applying $\varepsilon^2\partial_t$ to the first equation in (8.1), we compute

$$\varepsilon^2\partial_t(a^\varepsilon\partial_t p^\varepsilon) - \operatorname{div}(b^\varepsilon\nabla p^\varepsilon) = c^\varepsilon := \varepsilon f^\varepsilon, \tag{8.4}$$

with $a^\varepsilon := g_1(\psi^\varepsilon)$, $b^\varepsilon := 1/g_2(\psi^\varepsilon)$ and

$$f^\varepsilon := \partial_t F_1(Y^\varepsilon) + \varepsilon\partial_t F_2(Y^\varepsilon, Z^\varepsilon) - \operatorname{div}\big(b^\varepsilon F_3(Y^\varepsilon, Z^\varepsilon)\big), \tag{8.5}$$

where $Y^\varepsilon := D^m\psi^\varepsilon$ and $Z^\varepsilon := (D^m p^\varepsilon, D^m v^\varepsilon)$.

The equations (8.1) imply that one can rewrite $f^\varepsilon$ as

$$f^\varepsilon = \mathcal{F}(D^{2m}p^\varepsilon, D^{2m}v^\varepsilon, D^{2m}\psi^\varepsilon),$$

where $\mathcal{F}$ is a $C^\infty$ function such that $\mathcal{F}(0) = 0$. Using the usual nonlinear estimate in Sobolev spaces (5.8), the hypothesis (8.2) implies that $f^\varepsilon$ is uniformly bounded in $C^0([0,T]; H^{s-2m}(\mathbb{R}^d))$. Consequently, $c^\varepsilon = \varepsilon f^\varepsilon$ converges to 0 in $L^2(0,T; L^2(\mathbb{R}^d))$ and Theorem 8.3 applies.

To prove the second half of Proposition 8.2, we begin by proving that $\dot{p}^\varepsilon := (\varepsilon\partial_t)p^\varepsilon$ converges to 0 in $L^2(0,T; L^2_{loc}(\mathbb{R}^d))$. To do so we apply $(\varepsilon\partial_t)$ on equation (8.4), to obtain

$$\varepsilon^2\partial_t(a^\varepsilon\partial_t\dot{p}^\varepsilon) - \operatorname{div}(b^\varepsilon\nabla\dot{p}^\varepsilon) = \widetilde{c}^\varepsilon := \varepsilon\widetilde{f}^\varepsilon,$$

with

$$\widetilde{f}^\varepsilon := \varepsilon\partial_t f^\varepsilon - \varepsilon\partial_t\big(\partial_t a^\varepsilon(\varepsilon\partial_t)p^\varepsilon\big) + \operatorname{div}\big(\partial_t b^\varepsilon\nabla p^\varepsilon\big),$$

where $f^\varepsilon$ is given by (8.5). Again one can verify that $\widetilde{f}^\varepsilon$ is a bounded sequence in $C^0([0,T]; L^2(\mathbb{R}^d))$, which proves the desired result.

To complete the proof, observe that

$$\operatorname{div} v^\varepsilon - F_1(Y^\varepsilon) = -g_1(\psi^\varepsilon)(\varepsilon\partial_t)p^\varepsilon + \varepsilon F_2(Z^\varepsilon).$$

Hence, the fact that $\operatorname{div} v^\varepsilon - F_1(Y^\varepsilon)$ converges to 0 in $L^2(0,T; L^2_{loc}(\mathbb{R}^d))$ follows from the previous step and the fact that $g_1(\psi^\varepsilon) - g_1(0)$ and $F_2(Z^\varepsilon)$ are uniformly bounded in $C^0([0,T]; H^s(\mathbb{R}^d))$. □

8.1. **Proof of Theorem 2.9.** To simplify the presentation we concentrate on the hardest case when $\kappa$ is a fixed positive constant.

We first prove the convergences for some sub-sequence of $(p^\varepsilon, v^\varepsilon, \theta^\varepsilon)$.

The equation for $\theta$ implies that $\partial_t\theta$ is bounded in $C^0([0,T]; H^{s-1}(\mathbb{R}^d))$. After extracting a sub-sequence, we can assume that, for all $s' < s$,

$$\theta^\varepsilon \to \theta \quad \text{in } C^0([0,T]; H^{s'+1}_{loc}(\mathbb{R}^d)), \tag{8.6}$$

where the limit $\theta$ belongs to $C^0([0,T]; H^{s'+1}_{loc}(\mathbb{R}^d)) \cap L^\infty(0,T; H^{s+1}(\mathbb{R}^d))$.

Since $(p^\varepsilon, v^\varepsilon)$ is uniformly bounded in $C^0([0,T]; H^s(\mathbb{R}^d))$, after extracting further sub-sequence, we can also assume that

$$(p^\varepsilon, v^\varepsilon) \rightharpoonup (p, v) \quad \text{weakly } \star \text{ in } L^\infty(0,T; H^s(\mathbb{R}^d)). \tag{8.7}$$

Note that the system (8.1) includes system (2.1) as a special case where $\psi^\varepsilon := (\theta^\varepsilon, \varepsilon p^\varepsilon)$ and $m = 2$.

It follows from the very definition of $\psi^\varepsilon$ and the norms $\|\cdot\|_{H^\sigma_\varrho}$ that

$$\begin{aligned}\|(p^\varepsilon, v^\varepsilon)\|_{H^s} + \|\psi^\varepsilon\|_{H^{s+m-1}} &\leqslant \|(p^\varepsilon, v^\varepsilon)\|_{H^s} + \varepsilon\|p^\varepsilon\|_{H^{s+1}} + \|\theta^\varepsilon\|_{H^{s+1}}\\ &\leqslant \|(p^\varepsilon, v^\varepsilon)\|_{H^{s+1}_\varepsilon} + \|\theta^\varepsilon\|_{H^{s+1}}\\ &\lesssim \|(p^\varepsilon, v^\varepsilon)\|_{H^{s+1}_{\varepsilon\nu}} + \|\theta^\varepsilon\|_{H^{s+1}_\nu},\end{aligned}$$

where the implicit constant depends only on the fixed positive value of $\nu := \sqrt{\mu+\kappa}$. It follows that $\|(p^\varepsilon, v^\varepsilon, \psi^\varepsilon)\|_{\mathcal{K}^s(T)} \lesssim \|(p^\varepsilon, v^\varepsilon, \theta^\varepsilon)\|_{\mathcal{H}^s_{(\varepsilon,\mu,\kappa)}(T)}$. As a consequence the first hypothesis of Proposition 8.2 is satisfied. Therefore, in order to apply Proposition 8.2 to the equations (2.1) so as to obtain Theorem 2.9, it only remains to show that $\psi = (\theta, 0)$ satisfies the crucial assumption (8.3).

Multiplying the first equation in (2.1) by $\varepsilon$ and passing to the weak limit (in the sense of distributions) shows that $\operatorname{div} v^\varepsilon - \chi_1(\varepsilon p^\varepsilon) \operatorname{div}(\beta(\theta^\varepsilon)\nabla\theta^\varepsilon)$ converges to 0 in that sense. Therefore, the limit $\theta$ satisfies

$$g_3(\theta, 0)(\partial_t\theta + v\cdot\nabla\theta) - \kappa(\chi_3(0) - \chi_1(0)) \operatorname{div}(\beta(\theta)\nabla\theta) = 0. \tag{8.8}$$

The assumption $\chi_3 > \chi_1$ [see (A3) in Assumption 2.1] implies that (8.8) is parabolic. Hence, the desired spatial decay follows from estimates in weighted Sobolev spaces. Indeed, introduce $\theta_\delta := \langle x\rangle^\delta\theta$ for some given $\delta > 2$. Then

$$\partial_t\theta_\delta - \operatorname{div}\big(k(\theta)\nabla\theta_\delta\big) = \sum V_\alpha(t,x)\partial_x^\alpha\theta_\delta \qquad (\alpha\in\mathbb{N}^d,\ |\alpha|\leqslant 1),$$

for some smooth positive function $k$ and some coefficients $V_\alpha$ such that $\|V_\alpha\|_{H^s} \leqslant C(\|(v,\theta)\|_{H^s})$. Commuting $\Lambda^s$ with the equation and applying the usual $L^2$ estimate, we obtain

$$\|\theta_\delta\|_{L^\infty(0,T;H^s)} \leqslant C\,\|\theta_\delta(0)\|_{H^s}\,, \tag{8.9}$$

for some constant $C$ depending only on $\|(v,\theta)\|_{L^\infty(0,T;H^s)}$ (see (9.2) with $\eta = 1$). The convergence (8.6) implies that $\theta(0)$ is the limit $\theta_0$ of the initial data $\theta^\varepsilon(0)$. Using the hypothesis $\langle x\rangle^\delta\theta_0 \in H^s(\mathbb{R}^d)$, it follows from (8.9) that

$$\|\theta_\delta\|_{L^\infty(0,T;H^s)} < +\infty.$$

Since $s > 1 + d/2$, the Sobolev Theorem implies that $H^s(\mathbb{R}^d) \hookrightarrow W^{1,\infty}(\mathbb{R}^d)$. Hence, assumption (8.3) is satisfied.

A moment's thought shows that the convergences hold for the full sequence. The proof of Theorem 2.9 is complete.

## 9. Some estimates for elliptic, hyperbolic or parabolic systems

First, we briefly recall some standard estimates for hyperbolic or parabolic systems which we used throughout the paper. We state estimates for the solutions $u = u(t,x) \in \mathbb{R}^n$ of systems having the form:

$$\partial_t u + \sum_{|\alpha|=1} A_\alpha\partial_x^\alpha u - \eta\sum_{|\alpha|=2} A_\alpha\partial_x^\alpha u = f_1 + \sqrt{\eta} f_2, \tag{9.1}$$

where each $A_\alpha = A_\alpha(t,x)$ is a $n\times n$ matrix valued-function smooth in its arguments and furthermore symmetric: $A_\alpha = A_\alpha^t$; also, the source terms $f_i = f_i(t,x)$ are smooth functions with values in $\mathbb{R}^n$. In particular, when $\eta = 0$, system (9.1) is symmetric hyperbolic. We assume that for $\eta > 0$ the system is parabolic: there exists a positive $c$ such that for all triple $(t,x,\xi)$ with $\xi\neq 0$,

$$\Re\sum_{|\alpha|=2} A_\alpha(t,x)\xi^\alpha \geqslant c|\xi|^2 I.$$

**Lemma 9.1.** *Let $0 \leqslant \sigma \leqslant s \in (1 + d/2, +\infty)$. There exists a function $C(\cdot)$ such that for all $T \in [0, 1]$, all $u \in C^1([0, T]; H^\infty(\mathbb{D}))$ satisfying (9.1) and all $\eta \in [0, 1]$,*

$$\begin{aligned} &\|u\|_{L^\infty(0,T;H^\sigma)} + \sqrt{\eta}\,\|u\|_{L^2(0,T;H^{\sigma+1})} \\ &\qquad \leqslant K e^{C(R)T} \Big( \|u(0)\|_{H^\sigma} + \|f_1\|_{L^1(0,T:H^\sigma)} + \|f_2\|_{L^2(0,T;H^{\sigma-1})} \Big), \end{aligned} \tag{9.2}$$

*where $K = K(s, d)$ and $R := \sum_{|\alpha|\leqslant 2} \|\nabla A_\alpha\|_{L^\infty(0,T;H^{s-1})}$.*

*Proof.* By commuting the equation (9.1) with $\Lambda^\sigma$, using the commutator estimate (3.9) and the Gronwall's lemma, the proof of (9.2) can be reduced to the special case $\sigma = 0$. The latter case relies upon the usual integrations by parts and duality arguments. See [41] for the details. $\square$

We next prove an estimate used in the proof of Proposition 4.6.

**Lemma 9.2.** *There exists two constants $K_i = K_i(d)$ such that for all Lipschitz functions $\zeta = \zeta(x)$ and $\eta = \eta(x)$ such that $\zeta > 0$ and $\eta + \zeta > 0$, and for all vector field $u \in H^1(\mathbb{D})$,*

$$-\langle\, \zeta\Delta u + \eta\nabla \operatorname{div} u,\, u \,\rangle_{H^{-1}\times H^1} \geqslant K_1 m \|\nabla u\|^2_{L^2} - \frac{K_2 M^2}{m} \|u\|^2_{L^2}, \tag{9.3}$$

*where $m = \inf_{x\in\mathbb{D}}\{\zeta(x), \zeta(x) + \eta(x)\}$ and $M = \|\nabla\zeta\|_{L^\infty} + \|\nabla\eta\|_{L^\infty}$.*

The proof is credited to R. Danchin.

*Proof.* Decompose $u$ as $\underline{u} + \nabla\phi := \mathcal{Q}u + (I - \mathcal{Q})u$ where $\mathcal{Q}$ is the Leray projector onto divergence free vector field, so that

$$\zeta\Delta u + \eta\nabla \operatorname{div} u = \zeta\Delta\underline{u} + (\zeta + \eta)\nabla\Delta\phi.$$

It results from $\operatorname{div} \underline{u} = 0$ that

$$-\langle\, \zeta\Delta u + \eta\nabla \operatorname{div} u,\, u \,\rangle_{H^{-1}\times H^1} = \langle\, \zeta\nabla\underline{u},\, \nabla\underline{u} \,\rangle + \langle\, (\zeta + \eta)\Delta\phi,\, \Delta\phi \,\rangle + R,$$

where $R$ is such that $|R| \leqslant K_3 M \|\nabla u\|_{L^2} \|u\|_{L^2}$.

To infer the desired bound, we use two usual inequalities: (1) for all $\lambda \geqslant 1$, we have $|R| \leqslant (m/\lambda) \|\nabla u\|^2_{L^2} + (\lambda K_3^2 M^2/m) \|u\|^2_{L^2}$; (2) the simplest of all Calderón–Zygmund estimates: $\|\nabla^2\phi\|_{L^2} \leqslant K_4 \|\Delta\phi\|_{L^2}$. $\square$

**Acknowledgement.** I would like to express my sincere gratitude to G. Métivier for introducing me to the analysis of the low Mach number limit as well as for helpful discussions. I warmly thanks D. Bresch and D. Lannes for many advices.

## References

[1] T. Alazard, Incompressible limit of the nonisentropic Euler equations with solid wall boundary conditions, *Adv. in Differential Equations* **10**, 19–44 (2005).

[2] J.-M. Bony, Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires, *Ann. Sci. École Norm. Sup.* **14**, 209–246 (1981).

[3] D. Bresch, B. Desjardins, E. Grenier & C.-K. Lin, Low Mach number limit of viscous polytropic flows: formal asymptotics in the periodic case, *Stud. Appl. Math.* **109**, 125–149 (2002).

[4] D. Bresch, B. Desjardins & C.-K. Lin, On some compressible fluid models: Korteweg, lubrication, and shallow water systems, *Comm. Partial Differential Equations* **28**, 843–868 (2003).

[5] C. Cheverry, O. Guès & G. Métivier, Oscillations fortes sur un champ linéairement dégénéré, *Ann. Sci. École Norm. Sup.* **36**, 691–745 (2003).

[6] R. Danchin, Global existence in critical spaces for compressible Navier-Stokes equations, *Invent. Math.* **141**, 579–614 (2000).

[7] R. Danchin, Global existence in critical spaces for flows of compressible viscous and heat-conductive gases, *Arch. Ration. Mech. Anal.* **160**, 1–39 (2001).

[8] R. Danchin, Zero Mach number limit for compressible flows with periodic boundary conditions, *Amer. J. Math.* **124**, 1153–1219 (2002).

[9] R. Danchin, Zero Mach number limit in critical spaces for compressible Navier-Stokes equations, *Ann. Sci. École Norm. Sup.* **35**, 27–75 (2002).

[10] R. Danchin, Low Mach number limit for viscous compressible flows, *M2AN Math. Model. Numer. Anal.* **39**, 459–475 (2005).

[11] B. Desjardins & E. Grenier, Low Mach number limit of viscous compressible flows in the whole space, *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* **455**, 2271–2279 (1999).

[12] B. Desjardins, E. Grenier, P.-L. Lions & N. Masmoudi, Incompressible limit for solutions of the isentropic Navier-Stokes equations with Dirichlet boundary conditions, *J. Math. Pures Appl.* **78**, 461–471 (1999).

[13] B. Desjardins & C.-K. Lin, A survey of the compressible Navier-Stokes equations, *Taiwanese J. Math.* **3**, 123–137 (1999).

[14] I. Gallagher, Résultats récents sur la limite incompressible, *Séminaire Bourbaki 2003–2004*, num. 926.

[15] E. Grenier, Oscillatory perturbations of the Navier-Stokes equations, *J. Math. Pures Appl.* **76**, 477–498 (1997).

[16] D. Hoff, The zero-Mach limit of compressible flows, *Comm. Math. Phys.* **192**, 543–554 (1998).

[17] L. Hörmander, Lectures on nonlinear hyperbolic differential equations, Springer-Verlag (1997).

[18] H. Isozaki, Singular limits for the compressible Euler equation in an exterior domain, *J. Reine Angew. Math.* **381**, 1–36 (1987).

[19] H. Isozaki, Singular limits for the compressible Euler equation in an exterior domain. II. Bodies in a uniform flow, *Osaka J. Math.* **26**, 399–410 (1989).

[20] S. Kawashima & Y. Shizuta, On the normal form of the symmetric hyperbolic-parabolic systems associated with the conservation laws, *Tohoku Math. J.* **40** 449–464 (1988).

[21] S. Kawashima & W.-A. Yong, Dissipative Structure and Entropy for Hyperbolic Systems of Balance Laws, *Arch. Ration. Mech. Anal.*. **174** 345–364 (2004).

[22] S. Klainerman & A. Majda, Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids, *Comm. Pure Appl. Math.* **34**, 481–524 (1981).

[23] S. Klainerman & A. Majda, Compressible and incompressible fluids, *Comm. Pure Appl. Math.* **35**, 629–651 (1982).

[24] R. Klein, Semi-implicit extension of a Godunov-type scheme based on low Mach number asymptotics. I. One-dimensional flow, *J. Comput. Phys.* **121**, 213–237 (1995).

[25] D. Lannes, Sharp estimates for pseudo-differential operators with symbols of limited smoothness and commutators, Preprint Universitï¿½Bordeaux I (2004).

[26] P.-L. Lions, Mathematical topics in fluid mechanics. Vol. 1, Incompressible models, Oxford Science Publications (1996).

[27] P.-L. Lions & N. Masmoudi, Incompressible limit for a viscous compressible fluid, *J. Math. Pures Appl. (9)* **77**, 585–627 (1998).

[28] A. Majda, Compressible fluid flow and systems of conservation laws in several space variables, Applied Mathematical Sciences **53**, Springer-Verlag (1984).

[29] A. Matsumura & T. Nishida, The initial value problem for the equations of motion of viscous and heat-conductive gases, *J. Math. Kyoto Univ.* **20**, 67–104 (1980).

[30] G. Métivier & S. Schochet, The incompressible limit of the non-isentropic Euler equations, *Arch. Ration. Mech. Anal.* **158**, 61–90 (2001).

[31] G. Métivier & S. Schochet, Limite incompressible des équations d'Euler non isentropiques, Séminaire: Équations aux Dérivées Partielles, Exp. No. X, 17 (2001).

[32] G. Métivier & S. Schochet, Averaging theorems for conservative systems and the weakly compressible Euler equations, *J. Differential Equations* **187**, 106–183 (2003).

[33] Y. Meyer, Remarques sur un théorème de J.-M. Bony, Proceedings of the Seminar on Harmonic Analysis (Pisa, 1980), *Rend. Circ. Mat. Palermo (2)*, 1–20 (1981).

[34] C.-D. Munz, Computational fluid dynamics and aeroacoustics for low Mach number flow, Hyperbolic partial differential equations (Hamburg, 2001), 269–320, Vieweg (2002).

[35] T. Schneider, N. Botta, K. J. Geratz & R. Klein, Extension of finite volume compressible flow solvers to multi-dimensional, variable density zero Mach number flows, *J. Comput. Phys.* **155**, 248–286 (1999).

[36] S. Schochet, The compressible Euler equations in a bounded domain: existence of solutions and the incompressible limit, *Comm. Math. Phys.* **104**, 49–75 (1986).

[37] S. Schochet, Fast singular limits of hyperbolic PDE's, *J. Differential Equations* **114**, 476–512 (1994).

[38] S. Schochet, The mathematical theory of low Mach number flows, *M2AN Math. Model. Numer. Anal.* **39**, 441–458 (2005).

[39] P. Secchi, On the singular incompressible limit of inviscid compressible fluids, *J. Math. Fluid Mech.* **2**, 107–125 (2000).

[40] P. Secchi, On slightly compressible ideal flow in the half-plane, *Arch. Ration. Mech. Anal.*, **161**, 231–255 (2002).

[41] M. E. Taylor, Pseudodifferential operators and nonlinear PDE, Progress in Mathematics **100**, Birkhäuser Boston Inc. (1991).

[42] S. Ukai, The incompressible limit and the initial layer of the compressible Euler equation, *J. Math. Kyoto Univ.*, **26**, 323–331 (1986).

[43] V. N. Vatsa & E. Turkel, Choice of Variables and Preconditioning for Time Dependent Problems, AIAA, 16th Computational Fluid Dynamics Conference (2003).

[44] R. K. Zeytounian, Theory and applications of nonviscous fluid flows, Springer-Verlag, Berlin (2002).

[45] R. K. Zeytounian, Theory and applications of viscous fluid flows, Springer-Verlag, Berlin (2004).