



HAL
open science

Crime pays; homogenized wave equations for long times

Grégoire Allaire, Agnes Lamacz-Keymmling, Jeffrey Rauch

► **To cite this version:**

Grégoire Allaire, Agnes Lamacz-Keymmling, Jeffrey Rauch. Crime pays; homogenized wave equations for long times. *Asymptotic Analysis*, 2022, 128 (3), pp.295-336. 10.3233/ASY-211707 . hal-04431158

HAL Id: hal-04431158

<https://hal.science/hal-04431158v1>

Submitted on 1 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CRIME PAYS; HOMOGENIZED WAVE EQUATIONS FOR LONG TIMES

GRÉGOIRE ALLAIRE¹, AGNES LAMACZ², AND JEFFREY RAUCH³

ABSTRACT. This article examines the accuracy for large times of asymptotic expansions from periodic homogenization of wave equations. As usual, ϵ denotes the small period of the coefficients in the wave equation. We first prove that the standard two scale asymptotic expansion provides an accurate approximation of the exact solution for times t of order $\epsilon^{-2+\delta}$ for any $\delta > 0$. Second, for longer times, we show that a different algorithm, that is called criminal because it mixes different powers of ϵ , yields an approximation of the exact solution with error $O(\epsilon^N)$ for times ϵ^{-N} with N as large as one likes. The criminal algorithm involves high order homogenized equations that, in the context of the wave equation, were first proposed by Santosa and Symes and analyzed by Lamacz. The high order homogenized equations yield dispersive corrections for moderate wave numbers. We give a systematic analysis for all time scales and all high order corrective terms.

Key words. homogenization, secular growth, dispersive effects, asymptotic crimes, wave equations

2010 Mathematics Subject Classification: 35B27, 35L05

¹ Centre de Mathématiques Appliquées, École Polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau, France.

Email: gregoire.allaire@polytechnique.fr

² Department of Mathematics, University of Duisburg-Essen, 45127 Essen, Germany.

Email: agnes.lamacz@uni-due.de

³ Department of Mathematics, University of Michigan, Ann Arbor 48109 MI, USA.

Email: rauch@umich.edu

1. INTRODUCTION

1.1. Traditional homogenization, secular growth, and long times. This paper studies the long time behavior of the wave equation in an infinite periodic medium,

$$(1.1) \quad \rho(x/\epsilon) \partial_t^2 u^\epsilon - \operatorname{div}(a(x/\epsilon) \operatorname{grad} u^\epsilon) = f(t, x), \quad u^\epsilon = f = 0 \text{ for } t < 0,$$

where ρ, a are periodic functions. Denote by \mathbb{T}^d the unit torus. The unknown u^ϵ is real valued as is $\rho \in L^\infty(\mathbb{T}^d)$, while $a \in L^\infty(\mathbb{T}^d)$ has values that are real symmetric matrices. The coefficients are positive definite in the sense that there is a constant $m_1 > 0$ so that

for all $\xi \in \mathbb{R}^d$,

$$a(y)\xi \cdot \xi \geq m_1|\xi|^2, \quad \text{and} \quad \rho(y) \geq m_1, \quad \text{for a.e. } y \in \mathbb{T}^d.$$

The source term f is smooth with $\partial_{t,x}^\alpha f \in L^2(\mathbb{R}^{1+d})$ for all $\alpha \in \mathbb{N}^{d+1}$ and is supported in $\{0 \leq t \leq 1\}$ until Section 5.

The motivation comes from the articles, in chronological order, [23], [18], [14], [3], [9] that describe the behavior of solutions to the wave equation on the very long time scale $t \sim 1/\epsilon^2$ and beyond (see also the engineering literature, including [7], [15], and the numerical literature [1], [2], [6]). The descriptions on these time scales require modifications of the traditional two scale homogenization *ansatz*, which is the following:

$$(1.2) \quad U(\epsilon, t, x, x/\epsilon), \quad U(\epsilon, t, x, y) \sim \sum_{n=0}^{\infty} \epsilon^n u_n(t, x, y), \quad u_n(t, x, y) \text{ periodic in } y.$$

The right hand side is a formal power series in ϵ . No convergence is expected (see Appendix A). The sign \sim represents equality in the sense of formal power series. The coefficient functions u_n belong to the space of smooth functions of $(t, x, y) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{T}^d$ supported in $t \geq 0$ that are periodic in y .

The classical homgenization approach [8], [10], [11], [16], [22] shows that the ansatz (1.2) provides a good approximation on bounded time intervals. Section 3 proves that the traditional construction (1.2) yields in fact a good approximation on time intervals $0 \leq t \leq C\epsilon^{-2+\delta}$ with C as large and $\delta > 0$ as small as one likes.

It was first observed by Santosa and Symes in [23], and then proved in [18] (see also [14], [3], [9]), that a different ansatz that we call *criminal*, yields a good approximation for times of order ϵ^{-2} . In the elliptic setting, the criminal ansatz was first proposed by Bakhvalov and Panasenko [8]. They recognized that an analogous approach formally works for the wave equation but did not discover its utility for large time asymptotics.

To analyze the two scale ansatz (1.2), each profile u_n is written as the sum of its non oscillating contribution πu_n and its oscillating part $(I - \pi)u_n$, defined as ¹

$$u_n(t, x, y) = \pi u_n + (I - \pi)u_n, \quad (\pi u_n)(t, x) := \frac{1}{|\mathbb{T}^d|} \int_{\mathbb{T}^d} u_n(t, x, y) dy.$$

Introduce the traditional second order partial differential operators

$$(1.3) \quad \begin{aligned} \mathcal{A}_{yy} &:= \operatorname{div}_y a(y) \operatorname{grad}_y, \\ \mathcal{A}_{xx} &:= \operatorname{div}_x a(y) \operatorname{grad}_x, \\ \mathcal{A}_{xy} &:= \operatorname{div}_x a(y) \operatorname{grad}_y + \operatorname{div}_y a(y) \operatorname{grad}_x. \end{aligned}$$

¹This decomposition of the two scale hierarchy emphasizing the projector π follows modern developments in hyperbolic geometric optics, see [21].

Then

$$\left[\rho(x/\epsilon) \partial_t^2 - \operatorname{div} a(x/\epsilon) \operatorname{grad} \right] U(\epsilon, t, x, x/\epsilon) \sim W(\epsilon, t, x, x/\epsilon),$$

where W is the formal Laurent series in ϵ defined by

$$(1.4) \quad \left[\rho(y) \partial_t^2 - \frac{1}{\epsilon^2} \mathcal{A}_{yy} - \frac{1}{\epsilon} \mathcal{A}_{xy} - \mathcal{A}_{xx} \right] U(\epsilon, t, x, y) \sim W(\epsilon, t, x, y) := \sum_{n=-2}^{\infty} \epsilon^n w_n(t, x, y).$$

In the traditional approach the u_n are chosen so that $W = f$.

Since the source term $f(t, x)$ does not depend on y , it has no oscillating part, $(I - \pi)f = 0$, and thus it is natural to seek the u_n so that $(I - \pi)w_n = 0$. The formal power series U , satisfying (1.4), for which $(I - \pi)w_n = 0$ for all n has a very rigid structure that steers our analysis. For $k \geq 1$, Definition 2.2 introduces differential operators

$$(1.5) \quad \chi_k(y, \partial_t, \partial_x) = \sum_{\alpha \in \mathbb{N}^{d+1}, |\alpha|=k} c_{\alpha,k}(y) \partial_{t,x}^\alpha.$$

The coefficients $c_{\alpha,k}$ are solutions of periodic cell problems. The coefficients of the pure x derivatives in (1.5) are the classical k^{th} order correctors in elliptic homogenization [8], [10], [22]. Theorem 2.5 proves that the formal power series $U(t, x, y)$, that yield profiles w_n satisfying $(I - \pi)w_n = 0$ for all n , are characterized by u_n satisfying

$$(1.6) \quad \forall n \geq 0, \quad (I - \pi)u_n = \sum_{k=1}^n \chi_k(y, \partial_t, \partial_x) \pi u_{n-k}.$$

In particular the oscillating part is given in terms of the non-oscillating parts of lower order.

The second structural identity concerning the U satisfying $(I - \pi)w_n = 0$ is a formula for πw_n that involves homogenized differential operators $a_k^*(\partial_t, \partial_x)$ with constant coefficients. The operator a_2^* is the standard homogenized wave operator [8], [10], [11], [16], [22] (see (2.19) for its precise definition). For $k \geq 3$, the a_k^* are called high order homogenized operators [8]. By establishing a combinatorial formula for the a_k^* , Theorem 2.13 proves that the odd order homogenized operators vanish, $a_{2n+1}^* = 0$. This is a classical result for $a_1^* = 0$ and $a_3^* = 0$ (see e.g. [4] and references therein). It was already known for all odd orders in the elliptic case [24]. Theorem 2.10 proves that

$$(1.7) \quad \forall n, \quad (I - \pi)w_n = 0, \quad \iff \quad \forall n, \quad \pi w_n = \sum_{0 \leq 2j \leq n} a_{2j+2}^*(\partial_t, \partial_x) \pi u_{n-2j}.$$

Equation (1.7) expresses πw_n in terms of πu_m with $m \leq n$ and having the same parity as n . This is the *leap frog structure*.

The traditional algorithm [8], [10], [22] sets $W = f$ enforcing $w_0 = \pi w_0 = f$ and $w_n = 0$ for $n \neq 0$. The first equation $w_0 = f$ yields the homogenized wave equation $a_2^*(\partial_t, \partial_x) \pi u_0 = f$ whose solution $u_0 = \pi u_0$ has energy independent of t for t beyond the support of f . The

leading profile πu_0 does not grow with time. Demanding that $\pi w_n = 0$, for $n \geq 1$ leads to wave equations for πu_n with a source term given in terms of πu_m , with m smaller than n and having the same parity as n . One finds $\pi u_n = 0$ for n odd. One can quickly assess the rate of growth of the profiles u_{2n} in time. This is called *secular growth* (see Section 2.3). For u_2 one has $a_2^*(\partial_t, \partial_x)\pi u_2 + a_4^*(\partial_t, \partial_x)\pi u_0 = 0$, a wave equation for πu_2 with source that does not grow in time. Therefore u_2 cannot grow faster than t . Continuing one finds that πu_4 grows no faster than t^2 and πu_{2k} no faster than t^k . The leap frog structure shows that πu_n grows no faster than $t^{n/2}$. Without the leap frog structure one would have found t^n . The $(2n)^{\text{th}}$ term in the ansatz (1.2) is of size $\epsilon^{2n} t^n$. For times $t \sim 1/\epsilon^2$ the higher order terms can no longer be understood as correction terms. The slow secular growth from the leap frog structure explains why $t \sim 1/\epsilon^2$ is a critical time scale for the traditional expansion.

The secular growth estimate implies Theorem 3.1 asserting that *for any $N, \delta > 0$ choosing a sufficiently large number of terms in the traditional ansatz (1.2) guarantees that the error is $O(\epsilon^N)$ for times $t \leq 1/\epsilon^{2-\delta}$.*

Appendix A.2 contains an example showing that the classical approximation is not accurate for times $t \sim 1/\epsilon^{2+\delta}$ for any $\delta > 0$.

Remark 1.1. *It is interesting to contrast our results with those of Pastukhova [20], who showed that the large time behavior of periodic parabolic equations is given, at least to leading order, by the homogenized equation. In contrast, our example in Appendix A.2 shows that in the hyperbolic setting the secular growth renders the homogenized equation inaccurate for times longer than $1/\epsilon^2$. The method of Pastukhova [20], using Bloch waves, shares some features with ours, including high order homogenized equations and filtering. It is completely different because the time asymptotics of the parabolic equation is purely decaying while the wave equation oscillates indefinitely.*

1.2. Asymptotic crimes and longer times. To find approximate solutions for longer times we abandon the classical *ansatz* (1.2) that requires $w_0 = \Pi w_0 = f$ and $w_n = 0$ for $n \neq 0$ in two ways. Firstly, we allow the terms u_n in the series (1.2) to depend on ϵ and secondly we demand $w_n = 0$ only up to some precision. We call it an *asymptotic crime* in the spirit of the *variational crimes* of Strang for non-conforming finite elements [25]. In addition to homogenization theory, asymptotic crimes have a long history in fluid dynamics and geometric optics, see [19].

Bakhvalov and Panasenko [8] used the strategy to find a high order elliptic approximation in one shot. The present paper analyses the hyperbolic case showing that the analogous ansatz is accurate for very long times, a result that has no elliptic analogue.

To construct the criminal approximation, one starts with a two-scale series as in the classical ansatz. The main difference is to allow the terms in the series to depend on ϵ ". To emphasize the fact that the new profiles are not the same as the old ones we call them v_n and set

$$V(\epsilon, t, x, y) \sim \sum_{n=0}^{\infty} \epsilon^n v_n.$$

In order that the computations retain much of the structure from the traditional *ansatz* we demand that $(1 - \pi)v_n = 0$ for all n . That yields (1.6) and (1.7) and, in particular, the leading term is non oscillating, $v_0 = \pi v_0$.

Then (1.7) implies that the discrepancy $W - f = \pi(W - f)$ is equal to the sum of the lines,

$$(1.8) \quad \begin{aligned} & \epsilon^0 [a_2^*(\partial_t, \partial_x)\pi v_0 - f] + \\ & \epsilon^1 [a_2^*(\partial_t, \partial_x)\pi v_1] + \\ & \epsilon^2 [a_2^*(\partial_t, \partial_x)\pi v_2 + a_4^*(\partial_t, \partial_x)\pi v_0] + \\ & \epsilon^3 [a_2^*(\partial_t, \partial_x)\pi v_3 + a_4^*(\partial_t, \partial_x)\pi v_1] + \\ & \epsilon^4 [a_2^*(\partial_t, \partial_x)\pi v_4 + a_4^*(\partial_t, \partial_x)\pi v_2 + a_6^*(\partial_t, \partial_x)\pi v_0] + \\ & \epsilon^5 [a_2^*(\partial_t, \partial_x)\pi v_5 + a_4^*(\partial_t, \partial_x)\pi v_3 + a_6^*(\partial_t, \partial_x)\pi v_1] + \dots \end{aligned}$$

The problems of secular growth came from setting all the rows equal to zero. That yields equations for the corrector terms that have the preceding profiles as sources. The criminal strategy requires only that the sum of the lines vanishes. That can be achieved setting $\pi v_n = 0$ for all $n > 0$ and demanding that

$$(1.9) \quad \left[a_2^*(\partial_{t,x}) + \epsilon^2 a_4^*(\partial_{t,x}) + \epsilon^4 a_6^*(\partial_{t,x}) + \dots \right] v_0 = f, \quad v_0 = 0 \quad \text{for } t < 0.$$

The coefficients in Equation (1.9) depend on ϵ . To solve this equation with accuracy $O(\epsilon^N)$, v_0 must depend on ϵ , $v_0 = v_0(\epsilon, t, x)$. Including oscillatory correction terms, the approximation takes the new form

$$(1.10) \quad V(\epsilon, t, x, y) \sim \sum_{n=0}^{\infty} \epsilon^n v_n(\epsilon, t, x, y), \quad v_n(\epsilon, t, x, y) \text{ periodic in } y,$$

where $v_n = (I - \pi)v_n$, for $n \geq 1$, and each profile depends on ϵ . The series is different from (1.2). First v_n depends on ϵ . Second $\pi v_n = 0$ for $n \geq 1$. Expanding v_n with respect to ϵ shows that formally $\sum \epsilon^n v_n$ is equal to $\sum \epsilon^n u_n$ as power series in ϵ . It is crucial to note that none of the infinite series converges in any sense at all so resummation is not justified. The traditional ansatz $\sum_0^N \epsilon^n u_n$ is not a good approximation for times larger than $1/\epsilon^2$. The $v_n(\epsilon, t, x, y)$ that we construct yield good approximations where the traditional summation is innaccurate.

Neither we nor anyone else solves (1.9). To construct v_0 , (1.9) is modified in several ways. The first difficulty is that the terms $\epsilon^{2j-2} a_{2j}^*(\partial_t, \partial_x)$ are typically of order $2j$ in ∂_t . The more terms one keeps the higher order is the equation in ∂_t . The truncated operators usually define ill posed initial value problems, see our discussion just after (1.13). The first thing that we do is perform a normal form transformation that converts the operators a_{2j}^* with $j \geq 2$ to operators in ∂_x only. The normal form removes all the time derivatives other than those in a_2^* . In Section 4.2 it is proved that there are uniquely determined homogeneous operators $R_{2j}(\partial_{t,x})$ and $\tilde{a}_{2j}(\partial_x)$ of degree $2j$ so that as formal power series in ∂_t, ∂_x ,

$$(1.11) \quad \left[1 + \sum_{j=1}^{\infty} R_{2j}(\partial_{t,x}) \right] \left[\sum_{j=1}^{\infty} a_{2j}^*(\partial_{t,x}) \right] = a_2^*(\partial_{t,x}) + \sum_{j=2}^{\infty} \tilde{a}_{2j}(\partial_x).$$

The operators R_{2j} and \tilde{a}_{2j} are computable by a rapid recursive algorithm. This step has no analogue in the elliptic context. Applying $1 + \sum_{j=1}^{\infty} R_{2j}(\epsilon \partial_{t,x})$ to (1.9) yields the equivalent equation

$$(1.12) \quad \left[a_2^*(\partial_{t,x}) + \sum_{j=2}^{\infty} \epsilon^{2j-2} \tilde{a}_{2j}(\partial_x) \right] v_0(\epsilon, t, x) = \left[1 + \sum_{j=1}^{\infty} \epsilon^{2j} R_{2j}(\partial_{t,x}) \right] f.$$

To construct the criminal approximation the sums in (1.12) are first truncated to finite sums. The corresponding equation depends on the number of terms retained and the unknown function is denoted \underline{v}_0^k . The truncated equation of order k is, with $R^k := \sum_{j=1}^k R_{2j}$,

$$(1.13) \quad \left[a_2^*(\partial_{t,x}) + \sum_{j=2}^{k+1} \epsilon^{2j-2} \tilde{a}_{2j}(\partial_x) \right] \underline{v}_0^k(\epsilon, t, x) = \left[1 + R^k(\epsilon \partial_{t,x}) \right] f, \quad \underline{v}_0^k = 0 \text{ for } t < 0.$$

The initial value problem (1.13) is usually ill posed so does not define a profile \underline{v}_0^k . The typical reason for this ill-posedness is that the coefficient of the highest order derivative operator has the wrong sign. For example, it is known [13], [18], [4] that, at least when ρ is constant, the operator \tilde{a}_4 has the wrong sign so that (1.13) is ill posed for $k = 2$. Surprisingly, that is not a fatal flaw.

A classical idea (at least for $k = 1$) [18], [14], [3], [4] to overcome this obstacle is to rely on a Boussinesq's trick to obtain a suitable sign of the highest order space derivative by exchanging some space derivatives with time derivatives. Here, we rather use a filtering approach. The correctors $\epsilon^{2j-2} \tilde{a}_{2j}(\partial_x)$ added to $a_2^*(\partial_t, \partial_x)$ are only small compared to a_2^* when applied to functions whose Fourier transform is supported where $\epsilon \xi$ is small (ξ being the Fourier variable). The idea is to filter the source term. Choose $\psi_1 \in C_0^\infty(\mathbb{R}^d)$ equal to 1 on a neighborhood of the origin. Choose $0 < \alpha < 1$. The operator $\psi_1(\epsilon^\alpha D)$

is the Fourier multiplier $g \mapsto \mathcal{F}^{-1}(\psi_1(\epsilon^\alpha \xi) \mathcal{F}g)$. Equivalently $D := (1/i)\partial_x$. The filtered equation is

$$(1.14) \quad \left[a_2^*(\partial_{t,x}) + \sum_{j=2}^{k+1} \epsilon^{2j-2} \tilde{a}_{2j}(\partial_x) \right] v_0^k = \psi_1(\epsilon^\alpha D)(1 + R^k(\epsilon \partial_{t,x}))f, \quad v_0^k = 0 \text{ for } t < 0.$$

The right hand side has Fourier transform supported in $|\xi| \leq C\epsilon^{-\alpha} \ll 1/\epsilon$. Equation (1.14) is the one that is solved to determine a profile v_0^k . The filtered equation (1.14) has a unique tempered solution. That solution has spatial Fourier transform supported in $\epsilon^{-\alpha} \text{supp } \psi_1$. Energy bounds like those for a_2^* are proved in Section 4.4. The operator on the left in (1.13) is the sum of the homogenized operator and small higher order terms. The higher order terms are sometimes thought of as dispersive correctors. This is at least the original interpretation of \tilde{a}_4 in [23]. Equations (1.9), (1.12), (1.13) and (1.14) are called *high order homogenized equations*, introduced in the elliptic setting in [8]. Since then, many authors have used it. For the wave equation, see [23], [18], [14], [3], [9], [4]. Note that, as remarked in [3], the high order homogenized equation (1.13) of a given order k is not unique, since some x and t derivatives can be exchanged.

The next definition summarizes the recipe for the criminal approximate solution.

Definition 1.2 (Criminal approximation). Fix the choice of $\psi_1 \in C_0^\infty(\mathbb{R}^d)$, $0 < \alpha < 1$, and $0 \leq k \in \mathbb{N}$. Define profiles $v_n^k(\epsilon, t, x, y)$ for $0 \leq n \leq 2k+2$ as follows.

- *Nonoscillatory parts.* For $1 \leq n \leq 2k+2$ set $\pi v_n^k = 0$. For $n = 0$, πv_0^k is the unique tempered solution of the high order homogenized equation (1.14).
- *Oscillatory parts.* For $1 \leq n \leq 2k+2$ set $(I - \pi)v_n^k = \chi_n \pi v_0^k$, where χ_n is defined by (1.5), and $(I - \pi)v_0^k = 0$ (or equivalently, $v_0^k = \pi v_0^k$).

Define

$$(1.15) \quad V^k(\epsilon, t, x, y) := \sum_{n=0}^{2k+2} \epsilon^n v_n^k(\epsilon, t, x, y) = \left(I + \sum_{n=1}^{2k+2} \epsilon^n \chi_n(y, \partial_t, \partial_x) \right) v_0^k(\epsilon, t, x).$$

The criminal approximate solution is $V^k(\epsilon, t, x, x/\epsilon)$.

The main result of the present paper is the following approximation theorem.

Theorem 1.3 (Criminal error). Suppose that u^ϵ is the exact solution of (1.1) and V^k is given by (1.15). For each $k \geq 1$ there are positive constants C, ϵ_0 so that for $0 < \epsilon < \epsilon_0$ and $t \geq 0$, the error in energy satisfies

$$\left\| \nabla_{t,x}(u^\epsilon(t, x) - V^k(\epsilon, t, x, x/\epsilon)) \right\|_{L^2(\mathbb{R}_x^d)} \leq C \epsilon^{2k+2} \langle t \rangle, \quad \text{with } \langle t \rangle := \sqrt{1+t^2}.$$

Remark 1.4. i. If one wants the error to decrease as ϵ^{N_1} on time intervals $0 \leq t \leq C/\epsilon^{N_2}$, it suffices to choose k so that $N_1 + N_2 \leq 2k+2$.

ii. *The initial value problem defining v_0^k has constant coefficients. Its spatial Fourier Transform is given by an explicit formula. A spectrally accurate approximate solution is computable by FFT.*

Writing $u = \int_0^t \partial_t u dt$, with $u(0) = 0$, yields the following corollary.

Corollary 1.5. *With the assumptions and notations in Theorem 1.3, the error measured in $L^2(\mathbb{R}^d)$ satisfies*

$$\|u^\epsilon(t) - V^k(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}^d)} \leq C \epsilon^{2k+2} \langle t \rangle^2.$$

Remark 1.6. *Since the L^2 -norm of $\nabla_{t,x} u^\epsilon$ is of order 1 and independent of time t , the absolute error of Theorem 1.3 is also a relative error. In contrast, the L^2 -norm of u^ϵ is $O(\langle t \rangle)$ and can grow nearly this fast. Therefore, the relative error is better than the absolute error of Corollary 1.5 and is similar to that of Theorem 1.3.*

A more subtle corollary is that the oscillating part of the approximate solution is not necessary for the long time asymptotics if one is content with an error of the order of ϵ .

Corollary 1.7. *With the assumptions and notations in Theorem 1.3, the error from the leading term v_0^k satisfies*

$$\|u^\epsilon(t) - v_0^k(\epsilon, t, x)\|_{L^2(\mathbb{R}^d)} \leq C (\epsilon + \epsilon^{2k+2} \langle t \rangle^2).$$

Corollary 1.7 shows that for N as large as one likes, if one takes $k \geq N$, then uniformly on $0 \leq t \leq C/\epsilon^N$ the $L^2(\mathbb{R}^d)$ -error is smaller than ϵ using only the leading nonoscillatory term $v_0^k(\epsilon, t, x)$ in the approximate solution. Corollary 1.7 was proved in [9] in a more general context (almost periodic or random coefficients) with a proof based on Bloch waves. Theorem 1.3 improves previous results since, not only the approximation error is valid for times as large as one wants, but the error is as high order in ϵ as one wants. Our results also improve those of [18], [14] which were restricted to times of order $1/\epsilon^2$ with an error of order ϵ . The first paper [18] relies on two scale asymptotic expansions, while the second one [14] uses Bloch waves.

The previous works [18], [14] considered (1.1) with $f = 0$ and nonvanishing initial data. In addition in [14], $\partial_t u(0) = 0$. In [9] a right hand side and nonvanishing initial data are allowed, but the error estimates are restricted to order ϵ . One of the reasons that we can push the analysis further is that our choice simplifies some things. We next expand a little on this choice. The solutions of (1.1) with f smooth in time with values in $L^2(\mathbb{R}^d)$ and supported in a compact time interval satisfy

$$(1.16) \quad \forall j \in \mathbb{N}, \quad \sup_{0 < \epsilon < 1} \sup_{t \in \mathbb{R}} \|\nabla_{t,x} \partial_t^j u^\epsilon(t)\|_{L^2(\mathbb{R}^d)} < \infty.$$

The coefficients vary on the small scale ϵ but the solutions do not oscillate in time. For smooth solutions with $f = 0$ and Cauchy data $u^\epsilon(0), \partial_t u^\epsilon(0)$ the initial derivatives satisfy

$$\forall j \geq 0, \ell \in \{0, 1\}, \quad \partial_t^{2j+\ell} u^\epsilon(0) = \left(\frac{1}{\rho(x/\epsilon)} (\operatorname{div} a(x/\epsilon) \operatorname{grad}) \right)^j \partial_t^\ell u^\epsilon(0).$$

These yield formulas for $\nabla_{t,x} \partial_t^j u^\epsilon|_{t=0} := H_j^\epsilon(u^\epsilon(0), \partial_t u^\epsilon(0))$. The initial data corresponding to solutions satisfying (1.16) are those for which

$$(1.17) \quad \forall j, \quad \sup_\epsilon \|H_j^\epsilon(u^\epsilon(0), \partial_t u^\epsilon(0))\|_{L^2(\mathbb{R}^d)} < \infty.$$

For the $f = 0$ problem with Cauchy data satisfying (1.17) the approximation properties for both classical and criminal strategies are as in our Theorems. For the $f = 0$ problem the accuracy of the approximation is determined by how well the initial data can be approximated by data satisfying (1.17).

The condition (1.17) is awkward to use. For example when ρ, a are just $L^\infty(\mathbb{T}^d)$ and at least one of them is not constant it is true but not immediately obvious that no family of initial data that is independent of ϵ can satisfy this condition. Without performing a nontrivial computation it is not clear that there are initial data given by two scale expansions that satisfy this condition. The solutions from traditional homogenization with our source term $f(t, x)$ viewed for t beyond $\operatorname{supp} f$ show that there are many such two scale data.

Equation (1.1) shows that solutions that do not oscillate in time are important. Their description via Cauchy data is awkward. We study problem (1.1).

Remark 1.8. *Theorem 1.3 has no analogue in the elliptic setting. In the elliptic setting, high accuracy by criminal methods dates to Bahkvalov and Panashenko [8]. Smyshlyaev and Cherednichenko [24] present a different but related high order elliptic strategy. If one is only interested in local averages of the difference between the exact solution and its two-scale ansatz, then the oscillating terms $(I - \pi)u_n$ disappear by averaging. The non oscillating solutions of the high order homogenized equation give an approximation with high order accuracy to such averages.*

1.3. Outline of the present paper. In Section 2 the classical two scale asymptotic expansion for wave equations is analyzed. Theorem 2.13 proves that the odd order homogenized operators vanish and Theorem 2.15 shows that secular growth of the profiles is half as fast as one might expect. We call this *the leap frog structure* of the asymptotic expansion.

Section 3 studies the accuracy of the classical expansion. Classical proofs show that for bounded time and any N the error is $O(\epsilon^N)$. We prove that taking more corrector terms one has $O(\epsilon^N)$ accuracy for times of order $\epsilon^{-2+\delta}$, for any $\delta > 0$.

Section 4 presents the details of the derivation of the criminal asymptotic expansion and proves Theorem 1.3.

Section 5 shows that our results for sources f compactly supported in time suffice, by a simple argument, to treat sources that grow at most polynomially in time.

Section 6 discusses the application of our ideas to Schrödinger's equations. It also contains a discussion of differences that arise when considering systems of wave equations. These include challenges not yet resolved.

Appendix A gives an example in dimension $d = 1$ for which the upper bound on the secular growth predicted in Section 2 is attained. For the same example, the classical two scale asymptotic expansion (1.2) does not yield a good approximation for times $t \sim 1/\epsilon^{2+\delta}$.

Appendix B provides a classical a priori estimate for two scale oscillating functions.

Appendix C proves that solutions of the wave equation have finite energy for sources less regular in x but more regular in t than the standard condition $f \in L^1_{loc}(\mathbb{R}; L^2(\mathbb{R}^d))$.

2. ANALYSIS OF THE TWO SCALE ANSATZ (1.2)

This section analyses the classical ansatz (1.2) and not the criminal strategy with ϵ -dependent coefficients. Revisit the standard method of two scale asymptotic expansions for the wave equation (1.1). We depart from the textbooks [8], [10], [22] in several ways. First, in these books the method is usually applied to an elliptic equation and the wave equation is only said to be treated similarly and only for time bounded independently of ϵ . Second, we recognize that much information can be extracted from the part $(I - \pi)w_n = 0$ of equation (1.4). Exact combinatorial formulas are given for terms of all orders in the ansatz (1.2) and its leap frog structure is exhibited. All these results prepare the way for the criminal approach of Section 4.

Infinite order asymptotic expansions require that the source term f be infinitely smooth with respect to time. The periodic coefficients ρ and a are assumed to be in $L^\infty(\mathbb{T}^d)$.

2.1. Ansatz. Let $\mathcal{A}_{yy}, \mathcal{A}_{xy}, \mathcal{A}_{xx}$ be the second order partial differential operators defined in (1.3). Consider the two scale power series (1.2), and the formula (1.4) for the action of the differential operator. All terms $u_n(t, x, y)$ and $w_n(t, x, y)$ are periodic in y , equivalently defined for $y \in \mathbb{T}^d$. The relation (1.4) is equivalent to

$$(2.1) \quad \left[\rho(y) \partial_t^2 - \frac{1}{\epsilon^2} \mathcal{A}_{yy} - \frac{1}{\epsilon} \mathcal{A}_{xy} - \mathcal{A}_{xx} \right] \sum_{n=0}^{\infty} \epsilon^n u_n(t, x, y) = \sum_{n=-2}^{\infty} \epsilon^n w_n(t, x, y)$$

as formal Laurent series in ϵ . Equation (2.1) at order ϵ^n reads

$$(2.2) \quad \begin{aligned} \epsilon^{-2} : & \quad -\mathcal{A}_{yy} u_0 = w_{-2}, \\ \epsilon^{-1} : & \quad -(\mathcal{A}_{yy} u_1 + \mathcal{A}_{xy} u_0) = w_{-1}, \end{aligned}$$

and, for $k \geq 0$, the coefficient of ϵ^k is

$$(2.3) \quad \rho(y)\partial_t^2 u_k - (\mathcal{A}_{yy}u_{k+2} + \mathcal{A}_{xy}u_{k+1} + \mathcal{A}_{xx}u_k) = w_k.$$

To homogenize problem (1.1) the usual strategy is to choose the series U such that $W \sim f$, i.e to choose the profiles u_n such that $w_0 = f$ and $w_n = 0$ for all $0 \neq n \geq -2$. We perform a more subtle analysis. The equation $W - f \sim 0$ is satisfied if and only if $\pi(W - f) \sim 0$ and $(I - \pi)(W - f) \sim 0$. The source term $f(t, x)$ is smooth and non oscillatory, $(I - \pi)f = 0$. Therefore, it follows that

$$(2.4) \quad (I - \pi)w_k = 0 \quad \text{for all } k \geq -2.$$

Using (2.4) already yields a lot of information about the two-scale asymptotic expansion (1.2). The next subsection relies only on assumption (2.4). Both classical and criminal strategies use (2.4).

2.2. Projections and the hierarchy. The analysis of (2.2), (2.3) pivots around the second order symmetric elliptic operator $\mathcal{A}_{yy} : H^1(\mathbb{T}^d) \rightarrow H^{-1}(\mathbb{T}^d)$. Denote by π the $L^2(\mathbb{T}^d)$ orthogonal projection on constants,

$$\pi g := \frac{1}{|\mathbb{T}^d|} \int_{\mathbb{T}^d} g(y) dy.$$

This operator π coincides with the action of g as a distribution on the test function 1. It is therefore a well defined operator on all periodic distributions. This operator extends to functions of t, x, y by acting only on the last variable,

$$(\pi g)(t, x) := \frac{1}{|\mathbb{T}^d|} \int_{\mathbb{T}^d} g(t, x, y) dy.$$

Lemma 2.1 (Cell Problem). *The operators in (1.3) satisfy*

$$\pi \mathcal{A}_{yy} = 0, \quad \mathcal{A}_{yy} \pi = 0 \quad \text{and} \quad \pi \mathcal{A}_{xy} \pi = 0.$$

The nullspace of \mathcal{A}_{yy} is equal to the space of constant functions, i.e. $\pi H^1(\mathbb{T}^d)$. The image, Range \mathcal{A}_{yy} , is the subspace of mean zero functions, i.e. $(I - \pi)H^{-1}(\mathbb{T}^d)$. Therefore \mathcal{A}_{yy} is a bijection $(I - \pi)H^1(\mathbb{T}^d) \rightarrow (I - \pi)H^{-1}(\mathbb{T}^d)$ and thus has an inverse denoted by \mathcal{A}_{yy}^{-1} .

Proof. A classical application of the Lax-Milgram Lemma. □

To solve (2.2) and (2.3), these equations will be projected by π (yielding the non-oscillatory hierarchy) and $(I - \pi)$ (leading to the oscillatory hierarchy) and solved separately.

2.2.1. *The oscillatory hierarchy.* Consider power series U and W for which $(I - \pi)w_n = 0$ for all $n \geq -2$. Equations (2.2) and (2.3) are multiplied on the left by $(I - \pi)$. Using $(I - \pi)w_{-2} = 0$ the first line of (2.2) becomes

$$0 = (I - \pi)\mathcal{A}_{yy}u_0 = (I - \pi)\mathcal{A}_{yy}(I - \pi)u_0.$$

Lemma 2.1 shows that this is equivalent to $(I - \pi)u_0 = 0$, the oscillatory part of u_0 vanishes.

Since $\pi\mathcal{A}_{xy}\pi = 0$, one has $\pi\mathcal{A}_{xy}u_0 = \pi\mathcal{A}_{xy}\pi u_0 = 0$. Thus, the second line of (2.2) shows that $(I - \pi)w_{-1} = 0$ if and only if

$$(2.5) \quad (I - \pi)u_1 = -\mathcal{A}_{yy}^{-1}(I - \pi)\mathcal{A}_{xy}\pi u_0 = -\mathcal{A}_{yy}^{-1}\mathcal{A}_{xy}\pi u_0 := \chi_1(y, \partial_x)\pi u_0.$$

Next, derive analogous formulas expressing $(I - \pi)u_k$ in terms of the πu_j with $j < k$. Since by assumption $(I - \pi)w_k = 0$ for all $k \geq 0$, (2.3) leads to

$$(I - \pi)\left[\rho(y)\partial_t^2 u_k - \mathcal{A}_{yy}u_{k+2} - \mathcal{A}_{xy}u_{k+1} - \mathcal{A}_{xx}u_k\right] = 0.$$

By Lemma 2.1 this is equivalent to

$$(2.6) \quad (I - \pi)u_{k+2} = -\mathcal{A}_{yy}^{-1}(I - \pi)\left[\mathcal{A}_{xy}u_{k+1} + (\mathcal{A}_{xx} - \rho(y)\partial_t^2)u_k\right].$$

Equation (2.6) expresses the oscillatory part of u_{k+2} in terms of earlier profiles. It can be further simplified by rewriting the earlier profiles as $u_j = \pi u_j + (1 - \pi)u_j$, the sum of non oscillatory and oscillatory parts. Then express the $(1 - \pi)u_j$ parts in terms of still earlier profiles, and so on. In this way the oscillatory parts can be eliminated yielding a relation determining the oscillatory parts in terms of the nonoscillatory parts that is made explicit in Theorem 2.5. In (2.5) an operator χ_1 was introduced. This definition is now extended to higher order.

Definition 2.2. Set $\chi_{-1} := 0$, $\chi_0 := I$. For $k \geq 1$ define operators mapping functions of t, x to functions of t, x, y by

$$(2.7) \quad \chi_k(y, \partial_t, \partial_x) := -\mathcal{A}_{yy}^{-1}(I - \pi)\left[\mathcal{A}_{xy}\chi_{k-1} + (\mathcal{A}_{xx} - \rho(y)\partial_t^2)\chi_{k-2}\right] = (I - \pi)\chi_k.$$

This recovers the previous definition of $\chi_1 = -\mathcal{A}_{yy}^{-1}\mathcal{A}_{xy} = -\mathcal{A}_{yy}^{-1}(I - \pi)\mathcal{A}_{xy}$, where the last equality follows from Lemma 2.1. The operators χ_k are differential in t, x with coefficients that depend on y . The y -dependence arises only from the coefficients $a(y), \rho(y)$ and is analysed in the next Lemma. To show that the above definition makes sense, it suffices to prove that for any smooth function $\varphi \in C^\infty(\mathbb{R}^{1+d})$ and every $(t, x) \in \mathbb{R}^{1+d}$ the argument of \mathcal{A}_{yy}^{-1} , namely

$$(I - \pi)\left[\mathcal{A}_{xy}\chi_{k-1} + (\mathcal{A}_{xx} - \rho\partial_t^2)\chi_{k-2}\right]\varphi(t, x),$$

belongs to the range of \mathcal{A}_{yy} . This is verified in the next Lemma.

Lemma 2.3. *For all $k \geq 1$ the following holds. For every function $\varphi \in C^\infty(\mathbb{R}^{1+d})$ and every $(t, x) \in \mathbb{R}^{1+d}$ one has that*

$$\left[\mathcal{A}_{xy} \chi_{k-1} + (\mathcal{A}_{xx} - \rho \partial_t^2) \chi_{k-2} \right] \varphi(t, x)$$

belongs to $H^{-1}(\mathbb{T}^d)$. In particular $\chi_k \varphi(t, x) \in (I - \pi)H^1(\mathbb{T}^d)$. Furthermore, for any $k \geq 1$, there exist coefficients $c_{\beta, k} \in (I - \pi)H^1(\mathbb{T}^d)$ such that

$$(2.8) \quad \chi_k(y, \partial_t, \partial_x) = \sum_{|\beta|=k} c_{\beta, k}(y) \partial_{t, x}^\beta.$$

In particular, χ_k is a homogeneous operator of degree k in $\partial_{t, x}$.

Proof. The proof is by induction on k . For $k = 1$ one directly observes that $\mathcal{A}_{xy} \varphi$ belongs to $H^{-1}(\mathbb{T}^d)$ and thus $\chi_1 = -\mathcal{A}_{yy}^{-1}(I - \pi)\mathcal{A}_{xy}$ is a first-order operator in x with $(I - \pi)H^1(\mathbb{T}^d)$ coefficients. The case $k = 2$ is analogous. One has that $[\mathcal{A}_{xy} \chi_1 + (\mathcal{A}_{xx} - \rho \partial_t^2)] \varphi(t, x) \in H^{-1}(\mathbb{T}^d)$, since χ_1 satisfies (2.8). In particular $\chi_2 \varphi(t, x) \in (I - \pi)H^1(\mathbb{T}^d)$, which proves (2.8) for $k = 2$.

Assume the statement for $k \geq 1$ and prove it for $k + 1$. For a function $\varphi \in C^\infty(\mathbb{R}^{1+d})$ compute

$$(2.9) \quad \begin{aligned} & \left[\mathcal{A}_{xy} \chi_k + (\mathcal{A}_{xx} - \rho(y) \partial_t^2) \chi_{k-1} \right] \varphi(t, x) \\ &= \operatorname{div}_x (a(y) \operatorname{grad}_y \chi_k \varphi(t, x)) + \operatorname{div}_y (a(y) \operatorname{grad}_x \chi_k \varphi(t, x)) \\ & \quad + \operatorname{div}_x (a(y) \operatorname{grad}_x \chi_{k-1} \varphi(t, x)) - \rho(y) \partial_t^2 (\chi_{k-1} \varphi(t, x)). \end{aligned}$$

By the induction hypothesis $\chi_k \varphi(t, x)$ and $\chi_{k-1} \varphi(t, x)$ are in $H^1(\mathbb{T}^d)$. Therefore, all terms on the right hand side of (2.9) are in $H^{-1}(\mathbb{T}^d)$. In particular $\chi_{k+1} \varphi(t, x) \in (I - \pi)H^1(\mathbb{T}^d)$, which is the claimed result.

Since the operator \mathcal{A}_{xy} is homogeneous of degree one and $(\mathcal{A}_{xx} - \rho \partial_t^2)$ is homogeneous of degree two, it follows that χ_{k+1} is homogeneous of degree $k + 1$. \square

Remark 2.4. i. *If ρ is independent of y , the fact that $(I - \pi)\chi_0 = 0$ implies that $\chi_2(y, \partial_t, \partial_x)$ does not depend on ∂_t .* **ii.** *In this case an induction on k shows that for any $k \geq 1$, $\chi_k(y, \partial_t, \partial_x)$ contains only time derivatives of order $\leq k - 2$.*

The first structural result concerns formal power series U for which the oscillatory parts $(I - \pi)w_n$ vanish.

Theorem 2.5. *Fix $-2 \leq k \in \mathbb{Z}$. For a formal power series U and corresponding W the following are equivalent.*

i. *For $-2 \leq j \leq k$ one has*

$$(2.10) \quad (I - \pi)w_j = 0.$$

ii. For $0 \leq \ell \leq k + 2$ one has

$$(2.11) \quad (I - \pi)u_\ell = \sum_{n=1}^{\ell} \chi_n \pi u_{\ell-n}.$$

Remark 2.6. Relation (2.11) implies that the first term in the formal power series is not oscillating, namely $u_0 = (I - \pi)u_0$.

Theorem 2.5 has a particularly elegant form for profiles so that $(1 - \pi)w_\ell = 0$ for all ℓ . This holds if and only if the formal power series in ϵ for u is given in terms of the series for πu by

$$\sum_{n=0}^{\infty} \epsilon^n u_n = \left(\sum_{\ell=0}^{\infty} \epsilon^\ell \chi_\ell \right) \left(\sum_{k=0}^{\infty} \epsilon^k \pi u_k \right).$$

The elliptic analogue is in [8].

Proof. For $k = -2$ the statement follows directly by recalling that $(I - \pi)w_{-2} = 0$ if and only if $(I - \pi)u_0 = 0$. For $k = -1$ one has

$$(2.12) \quad (I - \pi)w_{-1} = -(I - \pi) \left(\mathcal{A}_{yy} u_1 + \mathcal{A}_{xy} u_0 \right).$$

Lemma 2.1 implies that

$$(2.13) \quad \pi \mathcal{A}_{yy} = 0 \quad \text{and} \quad (1 - \pi) \mathcal{A}_{yy} u_k = \mathcal{A}_{yy} (1 - \pi) u_k \quad \text{for all } k \geq 0.$$

Using (2.13) along with $u_0 = \pi u_0$ and applying $-\mathcal{A}_{yy}^{-1}$ to (2.12) yields

$$-\mathcal{A}_{yy}^{-1} (I - \pi) w_{-1} = (I - \pi) u_1 + \mathcal{A}_{yy}^{-1} (I - \pi) \mathcal{A}_{xy} \pi u_0 = (I - \pi) u_1 - \chi_1 \pi u_0.$$

Since $(I - \pi)w_{-1} = 0$ is equivalent to $-\mathcal{A}_{yy}^{-1} (I - \pi) w_{-1} = 0$, this proves the case $k = -1$ of the Theorem.

For $k \geq 0$ reason by induction. Assume the case $k - 1$ and prove the case k . The induction hypothesis is

$$(2.14) \quad (I - \pi)u_\ell = \sum_{n=1}^{\ell} \chi_n \pi u_{\ell-n} \quad \text{for } 0 \leq \ell \leq k + 1,$$

if and only if $(I - \pi)w_j = 0$ for $-2 \leq j \leq k - 1$. For the inductive step one needs to treat $j = k$ and $l = k + 2$. For $k \geq 0$ one has

$$(2.15) \quad (I - \pi)w_k = -(I - \pi) \left(\mathcal{A}_{yy} u_{k+2} + \mathcal{A}_{xy} u_{k+1} + (\mathcal{A}_{xx} - \rho(y) \partial_t^2) u_k \right).$$

Exploiting (2.13) and applying $-\mathcal{A}_{yy}^{-1}$ yields

$$(2.16) \quad -\mathcal{A}_{yy}^{-1} (I - \pi) w_k = (I - \pi) u_{k+2} + \mathcal{A}_{yy}^{-1} (I - \pi) \left(\mathcal{A}_{xy} u_{k+1} + (\mathcal{A}_{xx} - \rho(y) \partial_t^2) u_k \right).$$

Expressing each profile in (2.16) as a sum of its oscillatory and non-oscillatory part and recalling the definition $\chi_1 = -\mathcal{A}_{yy}^{-1}(I - \pi)\mathcal{A}_{xy}$, yields (2.16) as

$$(2.17) \quad \begin{aligned} -\mathcal{A}_{yy}^{-1}(I - \pi)w_k &= (I - \pi)u_{k+2} - \chi_1\pi u_{k+1} + \mathcal{A}_{yy}^{-1}(I - \pi)\left((\mathcal{A}_{xx} - \rho(y)\partial_t^2)\pi u_k\right) \\ &+ \mathcal{A}_{yy}^{-1}(I - \pi)\left(\mathcal{A}_{xy}(I - \pi)u_{k+1} + (\mathcal{A}_{xx} - \rho(y)\partial_t^2)(I - \pi)u_k\right). \end{aligned}$$

Use that $(I - \pi)w_k = 0$ if and only if $\mathcal{A}_{yy}^{-1}(I - \pi)w_k = 0$. Thus $(I - \pi)w_k = 0$ if and only if the right hand side of (2.17) vanishes. Using the induction hypothesis, (2.14) holds for $(I - \pi)u_{k+1}$ and $(I - \pi)u_k$. This yields

$$\begin{aligned} (I - \pi)u_{k+2} &= \chi_1\pi u_{k+1} - \mathcal{A}_{yy}^{-1}(I - \pi)\left((\mathcal{A}_{xx} - \rho(y)\partial_t^2)\pi u_k\right) \\ &- \mathcal{A}_{yy}^{-1}(I - \pi)\left(\mathcal{A}_{xy}\sum_{n=1}^{k+1}\chi_n\pi u_{k+1-n} + (\mathcal{A}_{xx} - \rho(y)\partial_t^2)\sum_{n=1}^k\chi_n\pi u_{k-n}\right). \\ &= \chi_1\pi u_{k+1} - \mathcal{A}_{yy}^{-1}(I - \pi)\left((\mathcal{A}_{xx} - \rho(y)\partial_t^2)\pi u_k + \mathcal{A}_{xy}\chi_1\pi u_k\right) \\ &- \mathcal{A}_{yy}^{-1}(I - \pi)\left(\mathcal{A}_{xy}\sum_{n=2}^{k+1}\chi_n\pi u_{k+1-n} + (\mathcal{A}_{xx} - \rho(y)\partial_t^2)\sum_{n=1}^k\chi_n\pi u_{k-n}\right). \end{aligned}$$

By definition $-\mathcal{A}_{yy}^{-1}(I - \pi)\left((\mathcal{A}_{xx} - \rho(y)\partial_t^2)\pi u_k + \mathcal{A}_{xy}\chi_1\pi u_k\right) = \chi_2\pi u_k$. Therefore

$$\begin{aligned} (I - \pi)u_{k+2} &= \chi_1\pi u_{k+1} + \chi_2\pi u_k - \mathcal{A}_{yy}^{-1}(I - \pi)\left(\mathcal{A}_{xy}\sum_{n=2}^{k+1}\chi_n\pi u_{k+1-n} + (\mathcal{A}_{xx} - \rho(y)\partial_t^2)\sum_{n=1}^k\chi_n\pi u_{k-n}\right) \\ &= \chi_1\pi u_{k+1} + \chi_2\pi u_k + \sum_{n=1}^k\left(-\mathcal{A}_{yy}^{-1}(I - \pi)\right)\left[\mathcal{A}_{xy}\chi_{n+1} + (\mathcal{A}_{xx} - \rho(y)\partial_t^2)\chi_n\right]\pi u_{k-n} \\ &= \chi_1\pi u_{k+1} + \chi_2\pi u_k + \sum_{n=1}^k\chi_{n+2}\pi u_{k-n} = \sum_{n=1}^{k+2}\chi_n\pi u_{k+2-n}, \end{aligned}$$

where the last line uses the definition (2.7) of χ_{n+2} . The last identity is the desired formula for $(I - \pi)u_{k+2}$. The proof is complete. \square

2.2.2. The nonoscillatory hierarchy. Next analyse the equations determining the non oscillatory parts πu_n of the profiles. Equations (2.2) and (2.3) are multiplied on the left by π . Since $\pi\mathcal{A}_{yy} = 0$ and $\pi\mathcal{A}_{xy}\pi = 0$, one has $\pi w_{-2} = 0$ and $\pi w_{-1} = 0$. For $k \geq 0$ the \mathcal{A}_{yy} terms are eliminated and (2.3) with $k \geq 0$ simplifies to

$$(2.18) \quad \pi w_k = \pi\left[\rho(y)\partial_t^2 u_k - \mathcal{A}_{xx}u_k - \mathcal{A}_{xy}u_{k+1}\right].$$

Exploiting (2.18) with $k = 0$, writing $u_1 = \pi u_1 + (1 - \pi)u_1$ and using the recurrence from (2.11) yields

$$\pi \mathcal{A}_{xy} u_1 \stackrel{\pi \mathcal{A}_{xy} \pi = 0}{=} \pi \mathcal{A}_{xy} (I - \pi) u_1 \stackrel{(2.11)}{=} \pi \mathcal{A}_{xy} \chi_1 \pi u_0 \stackrel{(2.7)}{=} -\pi \mathcal{A}_{xy} \mathcal{A}_{yy}^{-1} (I - \pi) \mathcal{A}_{xy} \pi u_0.$$

Since $u_0 = \pi u_0$, this yields

$$\pi w_0 = a_2^*(\partial_t, \partial_x) \pi u_0$$

with the homogenized wave operator defined as

$$(2.19) \quad a_2^*(\partial_t, \partial_x) := (\pi \rho) \partial_t^2 - \operatorname{div}_x (\pi a) \operatorname{grad}_x + \pi \mathcal{A}_{xy} (I - \pi) \mathcal{A}_{yy}^{-1} (I - \pi) \mathcal{A}_{xy} \pi.$$

Remark 2.7. *The homogenized wave operator a_2^* coincides with the formula from classical homogenization theory [8], [10], [16], [22].*

Definition 2.8. *Scalar partial differential operators $a_n^*(\partial_t, \partial_x)$ mapping functions of t, x to functions of t, x are defined for $n \geq 1$ by*

$$(2.20) \quad a_n^*(\partial_t, \partial_x) = \pi \left((\rho(y) \partial_t^2 - \mathcal{A}_{xx}) \chi_{n-2} - \mathcal{A}_{xy} \chi_{n-1} \right).$$

Remark 2.9. **i.** *The operators a_n^* have constant coefficients.* **ii.** *The operator a_n^* is homogeneous of degree n .* **iii.** *The symbol $a_n^*(\partial_t, \partial_x)$ contains only even powers of ∂_t .* **iv.** *The definitions of χ_0, χ_{-1} imply that $a_1^* = 0$.*

Theorem 2.10. *Suppose that the formal power series U and corresponding W satisfy the conditions of Theorem 2.5 for some $k \in \mathbb{Z}$ with $k \geq -2$. Then $\pi w_{-2} = \pi w_{-1} = 0$ and for $0 \leq j \leq k + 1$,*

$$(2.21) \quad \pi w_j = \sum_{n=0}^j a_{n+2}^*(\partial_t, \partial_x) \pi u_{j-n}.$$

Remark 2.11. *The result is particularly elegant for profiles so that $(1 - \pi)w_n = 0$ for all n . In that case the formal power series in ϵ for the residual is given in terms of the nonoscillating parts by*

$$\sum_{j=0}^{\infty} \epsilon^j \pi w_j = \left(\sum_{n=0}^{\infty} \epsilon^n a_{n+2}^* \right) \left(\sum_{m=0}^{\infty} \epsilon^m \pi u_m \right).$$

The elliptic analogue was observed in [8].

Proof. The cases $k = -2$ and $k = -1$ have already been discussed at the beginning of Subsection 2.2.2. Let $k \geq 0$ and fix $0 \leq j \leq k + 1$. Using $\pi \mathcal{A}_{xy} \pi = 0$ and $\pi \mathcal{A}_{yy} = 0$ provides

$$(2.22) \quad \pi w_j = \pi \left((\rho(y) \partial_t^2 - \mathcal{A}_{xx}) \pi u_j + (\rho(y) \partial_t^2 - \mathcal{A}_{xx}) (I - \pi) u_j - \mathcal{A}_{xy} (I - \pi) u_{j+1} \right).$$

Since we assumed that the conditions of Theorem 2.5 hold for k and since $j, j+1 \leq k+2$, we can replace $(I - \pi)u_l$ in (2.22), for $l = j, j+1$ according to formula (2.11). This yields

$$\begin{aligned} \pi w_j &= \pi(\rho(y)\partial_t^2 - \mathcal{A}_{xx})\pi u_j + \pi\left((\rho(y)\partial_t^2 - \mathcal{A}_{xx})\sum_{n=1}^j \chi_n \pi u_{j-n} - \mathcal{A}_{xy}\sum_{n=1}^{j+1} \chi_n \pi u_{j+1-n}\right) \\ &= \pi\left((\rho(y)\partial_t^2 - \mathcal{A}_{xx}) - \mathcal{A}_{xy}\chi_1\right)\pi u_j + \pi\left(\sum_{n=1}^j (\rho(y)\partial_t^2 - \mathcal{A}_{xx})\chi_n \pi u_{j-n} - \sum_{n=2}^{j+1} \mathcal{A}_{xy}\chi_n \pi u_{j+1-n}\right) \end{aligned}$$

Regrouping terms and recalling that $\chi_0 = I$ yields

$$(2.23) \quad \pi w_j = \pi \sum_{n=0}^j \left((\rho(y)\partial_t^2 - \mathcal{A}_{xx})\chi_n - \mathcal{A}_{xy}\chi_{n+1} \right) \pi u_{j-n}.$$

By definition of the effective operators a_{n+2}^* , Equation (2.23) is equivalent to

$$(2.24) \quad \pi w_j = \sum_{n=0}^j a_{n+2}^*(\partial_t, \partial_x) \pi u_{j-n}.$$

This completes the proof. \square

Remark 2.12. a_n^* is a homogeneous polynomial of degree n in (∂_t, ∂_x) . Formula (2.20) shows that the highest degree of ∂_t in a_n^* comes from χ_{n-1} or $\partial_t^2 \chi_{n-2}$. When ρ is independent of y , Remark 2.4 yields that χ_{n-1} and χ_{n-2} are of degree $\leq n-3$ and $\leq n-4$, respectively, with respect to time t . Therefore, when ρ is constant, a_n^* is of order $\leq n-2$ in ∂_t for $n > 2$.

The next result shows that the equation (2.24) has half as many terms as it seems. The proof depends on a precise combinatorial formula for χ_k . The elliptic analogue of Theorem 2.13 was proved by a quite different variational argument in [24].

Theorem 2.13. *For any odd $n \geq 1$, the homogenized operator a_n^* vanishes. That is for $m \geq 0$, $a_{2m+1}^* = 0$.*

Proof. Introduce

$$C_1 := -\mathcal{A}_{yy}^{-1}(I - \pi)\mathcal{A}_{xy} \quad \text{and} \quad C_2 := -\mathcal{A}_{yy}^{-1}(I - \pi)(\mathcal{A}_{xx} - \rho(y)\partial_t^2).$$

The operator $\mathcal{A}_{yy}^{-1}(I - \pi)$ acts only on the y variable and is continuous from $H^{-1}(\mathbb{T}^d) \rightarrow H^1(\mathbb{T}^d)$. The operators C_j are homogeneous polynomials of degree j in (∂_t, ∂_x) , whose coefficients are operators in y . In the proof we integrate by parts with respect to y and not with respect to t, x . With these C_j , 2.7 yields

$$(2.25) \quad \chi_k = C_1 \chi_{k-1} + C_2 \chi_{k-2}, \quad k \geq 1.$$

Replace χ_{k-1} and χ_{k-2} using the two earlier instances of the recurrence. Continuing leads to an expression

$$\chi_k = W_k \chi_0,$$

where only the earliest operator χ_0 appears. Equation (2.25) implies that

$$(2.26) \quad W_k = C_1 W_{k-1} + C_2 W_{k-2} .$$

Equation (2.26) implies that W_k is the sum of all words written with the two "letters" C_1 and C_2 such that the number of letters satisfies $\#C_1 + 2\#C_2 = k$. Each word is a homogeneous differential operator of degree k in ∂_x . Separating the words into two groups, those that end in C_1 and those that end in C_2 implies that

$$W_k = W_{k-1} C_1 + W_{k-2} C_2 .$$

Denote with an exponent T the $L^2(\mathbb{T}^d)$ adjoint. Integration by parts in y shows that $\mathcal{A}_{xy}^T = -\mathcal{A}_{xy}$, while \mathcal{A}_{yy} and \mathcal{A}_{xx} are selfadjoint. Define operators

$$\begin{aligned} D_1 &:= -\mathcal{A}_{xy} \mathcal{A}_{yy}^{-1} (I - \pi) = -C_1^T, \\ D_2 &:= -(\mathcal{A}_{xx} - \rho(y) \partial_t^2) \mathcal{A}_{yy}^{-1} (I - \pi) = C_2^T. \end{aligned}$$

An induction shows that

$$W_k^T = (-1)^k Z_k \quad \text{with} \quad Z_k = D_1 Z_{k-1} + D_2 Z_{k-2} .$$

Therefore Z_k is the sum of all words written with the two letters D_1 and D_2 such that the number of letters satisfy $\#D_1 + 2\#D_2 = k$.

Introduce $G := \mathcal{A}_{yy}^{-1} (I - \pi)$ that satisfies

$$C_1 G = G D_1, \quad C_2 G = G D_2, \quad W_k G = G Z_k .$$

Since $\chi_0(y) = I$, definition (2.20) can be rewritten, by using $\mathcal{A}_{xy}^T = -\mathcal{A}_{xy}$, $(\mathcal{A}_{xx} - \rho \partial_t^2)^T = (\mathcal{A}_{xx} - \rho \partial_t^2)$ and $W_k^T = (-1)^k Z_k$, as

$$\begin{aligned} a_k^*(\partial_t, \partial_x) &= \int_{\mathbb{T}^d} \left((\rho \partial_t^2 - \mathcal{A}_{xx}) W_{k-2} \chi_0(y) - \mathcal{A}_{xy} W_{k-1} \chi_0(y) \right) \chi_0(y) dy \\ &= \int_{\mathbb{T}^d} \left(W_{k-2} \chi_0(y) (\rho \partial_t^2 - \mathcal{A}_{xx}) \chi_0(y) + W_{k-1} \chi_0(y) \mathcal{A}_{xy} \chi_0(y) \right) dy \\ &= (-1)^k \int_{\mathbb{T}^d} \chi_0(y) \left(Z_{k-2} (\rho \partial_t^2 - \mathcal{A}_{xx}) \chi_0(y) - Z_{k-1} \mathcal{A}_{xy} \chi_0(y) \right) dy . \end{aligned}$$

The properties of Z_k and W_k imply that

$$\begin{aligned} -Z_{k-1} \mathcal{A}_{xy} &= (D_1 Z_{k-2} + D_2 Z_{k-3}) \mathcal{A}_{xy} \\ &= ((\rho(y) \partial_t^2 - \mathcal{A}_{xx}) G Z_{k-3} - \mathcal{A}_{xy} G Z_{k-2}) \mathcal{A}_{xy} \\ &= ((\rho(y) \partial_t^2 - \mathcal{A}_{xx}) W_{k-3} G - \mathcal{A}_{xy} W_{k-2} G) \mathcal{A}_{xy} \\ &= (\rho(y) \partial_t^2 - \mathcal{A}_{xx}) W_{k-3} C_1 - \mathcal{A}_{xy} W_{k-2} C_1 . \end{aligned}$$

Similarly,

$$\begin{aligned}
Z_{k-2}(\rho(y)\partial_t^2 - \mathcal{A}_{xx}) &= (D_1 Z_{k-3} + D_2 Z_{k-4})(\rho(y)\partial_t^2 - \mathcal{A}_{xx}) \\
&= ((\rho(y)\partial_t^2 - \mathcal{A}_{xx})GZ_{k-4} - \mathcal{A}_{xy}GZ_{k-3})(\rho(y)\partial_t^2 - \mathcal{A}_{xx}) \\
&= ((\rho(y)\partial_t^2 - \mathcal{A}_{xx})W_{k-4}G - \mathcal{A}_{xy}W_{k-3}G)(\rho(y)\partial_t^2 - \mathcal{A}_{xx}) \\
&= (\rho(y)\partial_t^2 - \mathcal{A}_{xx})W_{k-4}C_2 - \mathcal{A}_{xy}W_{k-3}C_2.
\end{aligned}$$

Summing yields

$$\begin{aligned}
Z_{k-2}(\rho(y)\partial_t^2 - \mathcal{A}_{xx}) - Z_{k-1}\mathcal{A}_{xy} &= (\rho(y)\partial_t^2 - \mathcal{A}_{xx})(W_{k-3}C_1 + W_{k-4}C_2) \\
&\quad - \mathcal{A}_{xy}(W_{k-2}C_1 + W_{k-3}C_2) + (\rho(y)\partial_t^2 - \mathcal{A}_{xx})W_{k-2} - \mathcal{A}_{xy}W_{k-1}.
\end{aligned}$$

Therefore

$$\begin{aligned}
a_k^*(\partial_t, \partial_x) &= (-1)^k \int_{\mathbb{T}^d} \chi_0(y) \left((\rho\partial_t^2 - \mathcal{A}_{xx})W_{k-2}\chi_0(y) - \mathcal{A}_{xy}W_{k-1}\chi_0(y) \right) dy \\
&= (-1)^k a_k^*(\partial_t, \partial_x).
\end{aligned}$$

For odd k , this implies $a_k^*(\partial_t, \partial_x) = 0$. □

2.3. Leap frog and secular growth. So far our only assumption was (2.4), namely $(I - \pi)w_k = 0$ for all $k \geq -2$. Theorem 2.5 then shows that the nonoscillating part of (1.2) determines the oscillatory part by

$$(2.27) \quad (I - \pi)u_\ell = \sum_{n=1}^{\ell} \chi_n \pi u_{\ell-n} \quad \text{for all } \ell \geq 0.$$

Theorem 2.10 shows that when the non oscillatory parts πu_n satisfy (2.27), then

$$\pi w_j = \sum_{n=0}^j a_{n+2}^*(\partial_t, \partial_x) \pi u_{j-n}.$$

Theorem 2.13 implies that only terms of the same parity appear,

$$\pi w_j = a_2^*(\partial_t, \partial_x) \pi u_j + \sum_{\substack{2 \leq n \\ 2n+k=j+2}} a_{2n}^*(\partial_t, \partial_x) \pi u_k.$$

To go further in the classical strategy assume that for all t, x, y

$$w_0(t, x, y) = f(t, x), \quad \forall 0 \neq n \geq -2, \quad w_n = 0.$$

It yields the following hierarchy of equations for the πu_k ,

$$\begin{aligned}
(2.28) \quad \epsilon^0 : \quad & a_2^*(\partial_t, \partial_x)\pi u_0 = f \\
\epsilon^1 : \quad & a_2^*(\partial_t, \partial_x)\pi u_1 = 0 \\
\epsilon^2 : \quad & a_2^*(\partial_t, \partial_x)\pi u_2 = -a_4^*(\partial_t, \partial_x)\pi u_0 \\
\epsilon^3 : \quad & a_2^*(\partial_t, \partial_x)\pi u_3 = -a_4^*(\partial_t, \partial_x)\pi u_1 \\
\epsilon^4 : \quad & a_2^*(\partial_t, \partial_x)\pi u_4 = -a_4^*(\partial_t, \partial_x)\pi u_2 - a_6^*(\partial_t, \partial_x)\pi u_0 \\
\epsilon^5 : \quad & a_2^*(\partial_t, \partial_x)\pi u_5 = -a_4^*(\partial_t, \partial_x)\pi u_3 - a_6^*(\partial_t, \partial_x)\pi u_1 \\
\epsilon^6 : \quad & a_2^*(\partial_t, \partial_x)\pi u_6 = -a_4^*(\partial_t, \partial_x)\pi u_4 - a_6^*(\partial_t, \partial_x)\pi u_2 - a_8^*(\partial_t, \partial_x)\pi u_0 \\
\epsilon^7 : \quad & a_2^*(\partial_t, \partial_x)\pi u_7 = -a_4^*(\partial_t, \partial_x)\pi u_5 - a_6^*(\partial_t, \partial_x)\pi u_3 - a_8^*(\partial_t, \partial_x)\pi u_1
\end{aligned}$$

In addition, for all n , $\pi u_n = 0$ for $t < 0$. The ϵ^0 -order equation yields the classical homogenized wave equation $a_2^*(\partial_t, \partial_x)\pi u_0 = f$.

The equations for the odd subscripts are decoupled from those with even subscripts in (2.28). The equations repeat in pairs. This is the *leap frog structure of the non oscillatory hierarchy*.

Lemma 2.14 (Leap frog). *For every $0 \leq k \in \mathbb{Z}$, $\pi u_{2k+1} = 0$.*

Proof. The statement follows by inspection of (2.28) and using that the initial data are zero. Starting with $n = 1$ one concludes by induction in steps of two, that $\pi u_n = 0$ for all odd n . \square

The leap frog structure implies that secular growth is slow. Without the leap frog structure one would have $\|u_n\| \lesssim t^n$ instead of the $t^{n/2}$ in next theorem.

Theorem 2.15 (Secular growth). *If there is a $\underline{t} > 0$ so that $f = 0$ for $t > \underline{t}$, then for each non zero $\alpha \in \mathbb{N}^{1+d} \setminus \{0\}$ and every $k = 0, 1, 2, \dots$ there exists a constant C depending on f, α and k so that for all $t \geq 0$,*

$$(2.29) \quad \left\| \partial_{t,x}^\alpha u_{2k}(t), \partial_{t,x}^\alpha u_{2k+1}(t) \right\|_{L^2(\mathbb{R}^d \times \mathbb{T}^d)} \leq C \langle t \rangle^k, \quad \langle t \rangle := \sqrt{1+t^2}.$$

Remark 2.16. *Estimate (2.29) provides a bound on the derivatives of the u_n but not on the u_n themselves. To estimate u_{2k} or u_{2k+1} use $u = u(0) + \int_0^t \partial_t u dt$ to find*

$$(2.30) \quad \left\| u_{2k}(t), u_{2k+1}(t) \right\|_{L^2(\mathbb{R}^d \times \mathbb{T}^d)} = O(\langle t \rangle^{k+1}).$$

Proof. By Theorem 2.5 the leading term $u_0(t, x)$ satisfies $u_0 = \pi u_0$, and by Theorem 2.10 for $w_0 = f$

$$a_2^*(\partial_t, \partial_x)u_0 = f, \quad u_0 = 0 \quad \text{for } t < 0.$$

²The notation $A \lesssim B$ means that there is a constant $C > 0$, independent of A and B , so that $A \leq CB$.

Since $f \in C_0^\infty(\mathbb{R}; H^s(\mathbb{R}^d))$ for all s , it follows that for $0 \neq \alpha \in \mathbb{N}^{1+d}$, $\partial_{t,x}^\alpha u_0 \in L^\infty(\mathbb{R}; L^2(\mathbb{R}^d \times \mathbb{T}^d))$.

One has $\pi u_1 = 0$. Equation (2.5) implies that the oscillatory part of u_1 satisfies

$$\partial_{t,x}^\alpha (I - \pi)u_1 = -\partial_{t,x}^\alpha \mathcal{A}_{yy}^{-1} \mathcal{A}_{xy} u_0 \in L^\infty(\mathbb{R}; L^2(\mathbb{R}^d \times \mathbb{T}^d)).$$

This completes the analysis of u_1 and therefore the case $k = 0$ of the Theorem.

The proof is by induction on k . Assuming the result for indices $\leq k$ it suffices to prove the case $k + 1$.

First estimate the π projections. Since $2(k + 1) + 1$ is odd, $\pi u_{2(k+1)+1} = 0$. To estimate πu_{2k+2} , use the equation

$$a_2^*(\partial_t, \partial_x)\pi u_{2k+2} = - \sum_{\substack{2 \leq n \\ n+j=k+2}} a_{2n}^*(\partial_t, \partial_x)\pi u_{2j}.$$

The case k of (2.29) bounds the right hand side. Since a_{2n}^* is a sum of derivatives, the inductive hypothesis implies that for all β including $\beta = 0$,

$$\left\| \partial_{t,x}^\beta a_2^*(\partial_t, \partial_x)\pi u_{2k+2} \right\|_{L^2(\mathbb{R}^d \times \mathbb{T}^d)} = O(\langle t \rangle^k).$$

It is important that the right hand side of the equation determining $a_2^*(\partial_t, \partial_x)\pi u_{2(k+1)}$ involves only derivatives of the earlier profiles and not the profiles themselves. The standard energy estimate for $a_2^*(\partial_t, \partial_x)$ implies that for $\alpha \neq 0$,

$$\left\| \partial_{t,x}^\alpha \pi u_{2k+2} \right\|_{L^2(\mathbb{R}^d \times \mathbb{T}^d)} = O(\langle t \rangle^{k+1}).$$

It remains to estimate $(I - \pi)u_n$ for $n = 2k + 2$ and $2k + 3$.

Equation (2.6) with index $2k + 2$ in place of $k + 2$ expresses $(I - \pi)u_{2k+2}$ in terms of the profiles with indices $\leq 2k + 1$. Those profiles are $O(\langle t \rangle^k)$ by the inductive hypothesis. This yields

$$\left\| \partial_{t,x}^\alpha (I - \pi)u_{2k+2} \right\|_{L^2(\mathbb{R}^d \times \mathbb{T}^d)} = O(\langle t \rangle^k)$$

an estimate stronger than the $O(\langle t \rangle^{k+1})$ required by the Theorem.

Equation (2.6) with index $2k + 3$ in place of $k + 2$ expresses $(I - \pi)u_{2k+3}$ in terms of the profiles with indices $\leq 2k + 2$. Those with index $\leq 2k + 1$ are $O(\langle t \rangle^k)$ by the inductive hypothesis. All the derivatives of the profile u_{2k+2} have just been shown to be $O(\langle t \rangle^{k+1})$. It follows that

$$\left\| \partial_{t,x}^\alpha (I - \pi)u_{2k+3} \right\|_{L^2(\mathbb{R}^d \times \mathbb{T}^d)} = O(\langle t \rangle^{k+1}).$$

□

3. HIGH ACCURACY ON $t \sim \epsilon^{-2+\delta}$ WITHOUT CRIMES

This section is devoted to a proof of the correctness of the traditional two scale ansatz for times strictly smaller than ϵ^{-2} . Even in the classical strategy, this result is new. The example of Appendix A shows that this is optimal. For the example, the traditional two scale ansatz is not a good approximation for longer times.

Theorem 3.1. *For $k \in \mathbb{N}$, define a truncated ansatz, constructed from the first non oscillating profiles $\pi u_0, \pi u_2, \dots, \pi u_{2k}$ by*

$$U^k(\epsilon, t, x, y) := \sum_{n=0}^{2k} \epsilon^n u_n(t, x, y) + \epsilon^{2k+1}(I - \pi)u_{2k+1} + \epsilon^{2k+2}(I - \pi)u_{2k+2}.$$

The approximate solution is $U^k(\epsilon, t, x, x/\epsilon)$. Denote by u^ϵ the exact solution of (1.1). There is a constant C , independent of $0 < \epsilon \leq 1$ and $t \geq 0$, so the energy of the error is bounded by

$$(3.1) \quad \|\nabla_{t,x}[u^\epsilon(t, x) - U^k(\epsilon, t, x, x/\epsilon)]\|_{L^2(\mathbb{R}_x^d)} \leq C \min \{ \epsilon^{2k+1} \langle t \rangle^{k+1}, \epsilon^{2k+2} \langle t \rangle^{k+2} \}.$$

Remark 3.2. i. *The energy of u^ϵ is bounded uniformly in time so the right hand side of (3.1) bounds the relative energy error.*

ii. *By choosing k large one gets arbitrarily high order accuracy on time intervals that grow as $1/\epsilon^{2-\delta}$ for any $\delta > 0$. Indeed, on the time interval $0 \leq t \leq \epsilon^{-\gamma}$ with $\gamma(k+1) < 2k+1$ the term $\epsilon^{2k+1} \langle t \rangle^{k+1}$ is of order $\epsilon^{2k+1-\gamma(k+1)}$ and tends to zero as $\epsilon \rightarrow 0$. Analogously for $\gamma(k+2) < 2k+2$ the term $\epsilon^{2k+2} \langle t \rangle^{k+2}$ is of order $\epsilon^{2k+2-\gamma(k+2)} \rightarrow 0$ as $\epsilon \rightarrow 0$.* **iii.** *There are two terms in the error estimate (3.1). The first is the smaller for large times $\langle t \rangle > \epsilon^{-1}$. The second is smaller for moderate times $\langle t \rangle < \epsilon^{-1}$.*

iv. *The problem (1.1) is invariant by differentiation in time. The derivative $\partial_t^j u^\epsilon$ is the solution of the same problem with source term $\partial_t^j f$. The profiles of the two scale asymptotic solution of that problem are equal to the functions $\partial_t^j u_n(t, x, y)$. Theorem 3.1 applied to that problem shows that with a constant C depending on j but independent of t and ϵ ,*

$$(3.2) \quad \|\partial_t^j \nabla_{t,x}[u^\epsilon(t, x) - U^k(\epsilon, t, x, x/\epsilon)]\|_{L^2(\mathbb{R}^d)} \leq C \min \{ \epsilon^{2k+1} \langle t \rangle^{k+1}, \epsilon^{2k+2} \langle t \rangle^{k+2} \}.$$

v. *Estimate (3.1) does not give any convergence result for times of the order or larger than ϵ^{-2} . One could wonder if it could not be improved by using exponential error estimate as in [17] for the elliptic setting with analyticity assumptions on the right hand side f . The 1-d example in Appendix A shows that it is not possible to improve (3.1) for larger times. The reason for this limitation is the secular growth in Theorem 2.15. Smoother source terms do not mitigate the polynomial time growth of the ansatz.*

The proof of Theorem 3.1 has three main ingredients. The first in §3.1 relies on all the work done so far. It is a precise formula for the difference between f and $\rho(x/\epsilon)\partial_t^2 U^k - \operatorname{div} a(x/\epsilon)\operatorname{grad} U^k$. That difference has terms no more regular than H^{-1} . They are estimated in §3.2. The error in energy with such singular source terms is bounded using Proposition C.1.

3.1. Formula for the residual. Take U^k as the finite power series from Theorem 3.1. Then

$$(3.3) \quad \left[\rho(y)\partial_t^2 - \mathcal{A}_{xx} - \frac{1}{\epsilon}\mathcal{A}_{xy} - \frac{1}{\epsilon^2}\mathcal{A}_{yy} \right] U^k(\epsilon, t, x, y) = W^k(\epsilon, t, x, y) = \sum_{n=-2}^{2k+2} \epsilon^n w_n(t, x, y).$$

The profiles w_j are the same as those in (2.1), except for the last two, $j = 2k+1, 2k+2$. For $j < 2k+1$,

$$w_0 = f, \quad \text{and for } j \neq 0, \quad -2 \leq j \leq 2k, \quad w_j = 0.$$

Indeed, for $j \leq 2k-2$ one has

$$w_j = (\rho(y)\partial_t^2 - \mathcal{A}_{xx})u_j - (\mathcal{A}_{yy}u_{j+2} + \mathcal{A}_{xy}u_{j+1}) = 0$$

by construction of the profiles u_j . For $j = 2k-1$ use the fact that $\pi u_{2k+1} = 0$ because $2k+1$ is odd to find that

$$\begin{aligned} w_{2k-1} &= (\rho(y)\partial_t^2 - \mathcal{A}_{xx})u_{2k-1} - (\mathcal{A}_{yy}(I - \pi)u_{2k+1} + \mathcal{A}_{xy}u_{2k}) \\ &= (\rho(y)\partial_t^2 - \mathcal{A}_{xx})u_{2k-1} - (\mathcal{A}_{yy}u_{2k+1} + \mathcal{A}_{xy}u_{2k}) = 0. \end{aligned}$$

For $j = 2k$ use $\pi u_{2k+1} = 0$ and $\mathcal{A}_{yy}\pi = 0$ to find

$$\begin{aligned} w_{2k} &= (\rho(y)\partial_t^2 - \mathcal{A}_{xx})u_{2k} - (\mathcal{A}_{yy}(I - \pi)u_{2k+2} + \mathcal{A}_{xy}(I - \pi)u_{2k+1}) \\ &= (\rho(y)\partial_t^2 - \mathcal{A}_{xx})u_{2k} - (\mathcal{A}_{yy}u_{2k+2} + \mathcal{A}_{xy}u_{2k+1}) = 0. \end{aligned}$$

Therefore

$$\begin{aligned} &\left[\rho(y)\partial_t^2 - \mathcal{A}_{xx} - \frac{1}{\epsilon}\mathcal{A}_{xy} - \frac{1}{\epsilon^2}\mathcal{A}_{yy} \right] U^k(\epsilon, t, x, y) - f \\ &= \epsilon^{2k+1}w_{2k+1}(t, x, y) + \epsilon^{2k+2}w_{2k+2}(t, x, y) =: r(\epsilon, t, x, y) \end{aligned}$$

and thus

$$\left[\rho(x/\epsilon)\partial_t^2 - \operatorname{div} a(x/\epsilon)\operatorname{grad} \right] U^k(\epsilon, t, x, x/\epsilon) - f = r(\epsilon, t, x, x/\epsilon).$$

Equation (3.3) shows that

$$\begin{aligned} w_{2k+1} &= (\rho(y)\partial_t^2 - \mathcal{A}_{xx})(I - \pi)u_{2k+1} - \mathcal{A}_{xy}(I - \pi)u_{2k+2}, \\ w_{2k+2} &= (\rho(y)\partial_t^2 - \mathcal{A}_{xx})(I - \pi)u_{2k+2}. \end{aligned}$$

Using $\pi u_{2k+1} = 0$ the term w_{2k+1} can be rewritten as

$$\begin{aligned} w_{2k+1} &= (\rho(y)\partial_t^2 - \mathcal{A}_{xx})(I - \pi)u_{2k+1} - \mathcal{A}_{xy}(I - \pi)u_{2k+2} \\ &= ((\rho(y)\partial_t^2 - \mathcal{A}_{xx})u_{2k+1} - \mathcal{A}_{xy}u_{2k+2} - \mathcal{A}_{yy}u_{2k+3}) + \mathcal{A}_{xy}\pi u_{2k+2} + \mathcal{A}_{yy}u_{2k+3} \\ &= \mathcal{A}_{xy}\pi u_{2k+2} + \mathcal{A}_{yy}u_{2k+3}. \end{aligned}$$

Therefore

$$\begin{aligned} (3.4) \quad r &= \left[\rho(y)\partial_t^2 - \mathcal{A}_{xx} \right] (I - \pi) \left(\epsilon^{2k+1}u_{2k+1} + \epsilon^{2k+2}u_{2k+2} \right) - \epsilon^{2k+1}\mathcal{A}_{xy}(I - \pi)u_{2k+2} \\ &= \epsilon^{2k+2} \left[\rho(y)\partial_t^2 - \mathcal{A}_{xx} \right] (I - \pi)u_{2k+2} + \epsilon^{2k+1} [\mathcal{A}_{xy}\pi u_{2k+2} + \mathcal{A}_{yy}u_{2k+3}]. \end{aligned}$$

The definitions of the operators \mathcal{A} yield for the first line of (3.4)

$$\begin{aligned} (3.5) \quad r &= \epsilon^{2k+1} \left[(\rho(y)\partial_t^2 - \operatorname{div}_x a(y)\operatorname{grad}_x) (I - \pi)u_{2k+1} + \epsilon\rho(y)\partial_t^2(I - \pi)u_{2k+2} \right] \\ &\quad - \epsilon^{2k+1} (\operatorname{div}_x a(y)\operatorname{grad}_y + \operatorname{div}_y a(y)\operatorname{grad}_x + \epsilon\operatorname{div}_x a(y)\operatorname{grad}_x) (I - \pi)u_{2k+2} \\ &:= \epsilon^{2k+1} (I(\epsilon, t, x, y) + II(\epsilon, t, x, y)) \end{aligned}$$

and for the second line, using that $\operatorname{grad}_y \pi u_{2k+2} = 0$,

$$\begin{aligned} (3.6) \quad r &= \epsilon^{2k+2} (\rho(y)\partial_t^2 - \mathcal{A}_{xx})(I - \pi)u_{2k+2} + \epsilon^{2k+1} [\mathcal{A}_{xy}\pi u_{2k+2} + \mathcal{A}_{yy}u_{2k+3}] \\ &= \epsilon^{2k+2} (\rho(y)\partial_t^2(I - \pi)u_{2k+2} - \mathcal{A}_{xx}u_{2k+2}) + \epsilon^{2k+2} \left[\left(\frac{1}{\epsilon}\mathcal{A}_{xy} + \mathcal{A}_{xx} \right) \pi u_{2k+2} + \frac{1}{\epsilon}\mathcal{A}_{yy}u_{2k+3} \right] \\ &= \epsilon^{2k+2} \left[\rho(y)\partial_t^2(I - \pi)u_{2k+2} - \operatorname{div}_x a(y)\operatorname{grad}_x u_{2k+2} \right] \\ &\quad + \epsilon^{2k+2} \left[\left(\frac{1}{\epsilon}\operatorname{div}_y a(y)\operatorname{grad}_x + \operatorname{div}_x a(y)\operatorname{grad}_x \right) \pi u_{2k+2} + \frac{1}{\epsilon}\operatorname{div}_y a(y)\operatorname{grad}_y u_{2k+3} \right] \\ &=: \epsilon^{2k+2} \left(\tilde{I}(\epsilon, t, x, y) + \widetilde{II}(\epsilon, t, x, y) \right). \end{aligned}$$

In addition, for $\ell = 2k + 1, 2k + 2, 2k + 3$, with χ_n given by (2.8),

$$(3.7) \quad (I - \pi)u_\ell = \sum_{n=1, \ell-n \text{ even}}^{\ell} \chi_n \pi u_{\ell-n} = \sum_{n=1, \ell-n \text{ even}}^{\ell} \sum_{|\beta|=n} c_{\beta,n}(y) \partial_{t,x}^\beta \pi u_{\ell-n}.$$

3.2. Estimates for the residual. In view of (3.7) and Theorem 2.15, $I(\epsilon, t, x, y)$ in (3.5) is a sum of terms of the form $\epsilon^p c(y)v(t, x)$ with $p \geq 0$, $c \in L^2(\mathbb{T}^d)$ and

$$\|\partial_{t,x}^\alpha v\|_{L^2(\mathbb{R}^d)} \lesssim \langle t \rangle^k$$

for $\alpha \in \mathbb{N}^{1+d}$, including $\alpha = 0$. Analogously, $\tilde{I}(\epsilon, t, x, y)$ in (3.6) is a sum of terms $\epsilon^p c(y)v(t, x)$ with

$$\|\partial_{t,x}^\alpha v\|_{L^2(\mathbb{R}^d)} \lesssim \langle t \rangle^{k+1}.$$

Proposition B.1 implies that the L^2 -norms of $I(\epsilon, t, x, x/\epsilon)$ and $\tilde{I}(\epsilon, t, x, x/\epsilon)$ satisfy

$$(3.8) \quad \begin{aligned} \|I(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} &\lesssim \langle t \rangle^k, \\ \|\tilde{I}(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} &\lesssim \langle t \rangle^{k+1}. \end{aligned}$$

The terms $II(\epsilon, t, x, y)$ in (3.5) and $\widetilde{II}(\epsilon, t, x, y)$ in (3.6) involve derivatives of $a(\cdot)$, so are not square integrable. Equation (3.7) shows that $II(\epsilon, t, x, y)$ is equal to

$$- (\operatorname{div}_x a(y) \operatorname{grad}_y + \operatorname{div}_y a(y) \operatorname{grad}_x + \epsilon \operatorname{div}_x a(y) \operatorname{grad}_x) \left[\sum_{n=1}^{2k+2} \sum_{|\beta|=n} c_{\beta,n}(y) \partial_{t,x}^\beta \pi u_{2k+2-n}(t, x) \right].$$

Evaluate at $y = x/\epsilon$ to show that, with div acting on functions depending on x as well as on functions depending on x/ϵ ,

$$\begin{aligned} (3.9) \quad II(\epsilon, t, x, x/\epsilon) &= -\epsilon \sum_{n=1}^{2k+2} \sum_{|\beta|=n} \operatorname{div} \left[a(x/\epsilon) c_{\beta,n}(x/\epsilon) \operatorname{grad} \partial_{t,x}^\beta \pi u_{2k+2-n}(t, x) \right] \\ &+ \sum_{n=1}^{2k+2} \sum_{|\beta|=n} \left(a(x/\epsilon) (\operatorname{grad}_y c_{\beta,n})(x/\epsilon) \cdot \operatorname{grad} \partial_{t,x}^\beta \pi u_{2k+2-n}(t, x) \right) \\ &:= \operatorname{div} (II^{(1)}(\epsilon, t, x, x/\epsilon)) + II^{(2)}(\epsilon, t, x, x/\epsilon). \end{aligned}$$

Arguing exactly as for I , using that each $c_{\beta,n} \in H^1(\mathbb{T}^d)$, it follows that $II^{(1)}(\epsilon, t, x, x/\epsilon)$, $\partial_t II^{(1)}(\epsilon, t, x, x/\epsilon)$, and, $II^{(2)}(\epsilon, t, x, x/\epsilon)$ are in $L^2(\mathbb{R}^d)$ with

$$(3.10) \quad \begin{aligned} \|\partial_t II^{(1)}(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} + \|II^{(1)}(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} &\lesssim \epsilon \langle t \rangle^k \\ \|II^{(2)}(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} &\lesssim \langle t \rangle^k. \end{aligned}$$

For $\widetilde{II}(\epsilon, t, x, y)$ one finds, exploiting that $\pi u_{2k+3} = 0$,

$$\begin{aligned} \widetilde{II}(\epsilon, t, x, y) &= \left(\frac{1}{\epsilon} \operatorname{div}_y a(y) \operatorname{grad}_x + \operatorname{div}_x a(y) \operatorname{grad}_y \right) \pi u_{2k+2} \\ &+ \frac{1}{\epsilon} \operatorname{div}_y a(y) \operatorname{grad}_y \left[\sum_{n=1}^{2k+3} \sum_{|\beta|=n} c_{\beta,n}(y) \partial_{t,x}^\beta \pi u_{2k+3-n}(t, x) \right]. \end{aligned}$$

Evaluate at $y = x/\epsilon$ to obtain

$$\begin{aligned} \widetilde{II}(\epsilon, t, x, x/\epsilon) &= \operatorname{div} [a(x/\epsilon) \operatorname{grad} \pi u_{2k+2}(t, x)] \\ &+ \sum_{n=1}^{2k+3} \sum_{|\beta|=n} \operatorname{div} \left[a(x/\epsilon) (\operatorname{grad}_y c_{\beta,n})(x/\epsilon) \partial_{t,x}^\beta \pi u_{2k+3-n}(t, x) \right] \\ &- \sum_{n=1}^{2k+3} \sum_{|\beta|=n} a(x/\epsilon) (\operatorname{grad}_y c_{\beta,n})(x/\epsilon) \cdot \operatorname{grad} \partial_{t,x}^\beta \pi u_{2k+3-n}(t, x) \\ &=: \operatorname{div} \left(\widetilde{II}^{(1)}(\epsilon, t, x, x/\epsilon) \right) + \widetilde{II}^{(2)}(\epsilon, t, x, x/\epsilon) \end{aligned}$$

with

$$(3.11) \quad \begin{aligned} \|\partial_t \widetilde{II}^{(1)}(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} + \|\widetilde{II}^{(1)}(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} &\lesssim \langle t \rangle^{k+1} \\ \|\widetilde{II}^{(2)}(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} &\lesssim \langle t \rangle^{k+1}. \end{aligned}$$

Combining estimates (3.10), (3.11) with (3.8) yields the residual

$$(3.12) \quad \begin{aligned} r(\epsilon, t, x, x/\epsilon) &= f(\epsilon, t, x) + \operatorname{div} g(\epsilon, t, x) \\ &= \tilde{f}(\epsilon, t, x) + \operatorname{div} \tilde{g}(\epsilon, t, x), \end{aligned}$$

with

$$\begin{aligned} f(\epsilon, t, x) &= \epsilon^{2k+1}(I + II^{(2)})(\epsilon, t, x, x/\epsilon), & \tilde{f}(\epsilon, t, x) &= \epsilon^{2k+2}(\tilde{I} + \widetilde{II}^{(2)})(\epsilon, t, x, x/\epsilon) \\ g(\epsilon, t, x) &= \epsilon^{2k+1}II^{(1)}(\epsilon, t, x, x/\epsilon), & \tilde{g}(\epsilon, t, x) &= \epsilon^{2k+2}\widetilde{II}^{(1)}(\epsilon, t, x, x/\epsilon) \end{aligned}$$

and the two estimates,

$$(3.13) \quad \begin{aligned} \|f(\epsilon, t, \cdot)\|_{L^2(\mathbb{R}^d)} &\lesssim \epsilon^{2k+1}\langle t \rangle^k, & \|\partial_t g(\epsilon, t, \cdot)\|_{L^2(\mathbb{R}^d)} + \|g(\epsilon, t, \cdot)\|_{L^2(\mathbb{R}^d)} &\lesssim \epsilon^{2k+2}\langle t \rangle^k \\ \|\tilde{f}(\epsilon, t, \cdot)\|_{L^2(\mathbb{R}^d)} &\lesssim \epsilon^{2k+2}\langle t \rangle^{k+1}, & \|\partial_t \tilde{g}(\epsilon, t, \cdot)\|_{L^2(\mathbb{R}^d)} + \|\tilde{g}(\epsilon, t, \cdot)\|_{L^2(\mathbb{R}^d)} &\lesssim \epsilon^{2k+2}\langle t \rangle^{k+1}. \end{aligned}$$

3.3. End of proof of Theorem 3.1. Denote by u^ϵ the exact solution and U^k the approximation from the statement of Theorem 3.1. We have proved that

$$\left[\rho(x/\epsilon) \partial_t^2 - \operatorname{div} a(x/\epsilon) \operatorname{grad} \right] (u^\epsilon(t, x) - U^k(\epsilon, t, x, x/\epsilon)) := r(\epsilon, t, x, x/\epsilon)$$

with $r(\epsilon, t, x, x/\epsilon)$ satisfying (3.12) and (3.13). Apply Proposition C.1. In the estimate of that Proposition, the $L^1([0, t]; L^2(\mathbb{R}^d))$ norms of f, \tilde{f} and $\partial_t g, \partial_t \tilde{g}$ are estimated by t times the $L^\infty([0, t]; L^2(\mathbb{R}^d))$ norms, which are controlled by (3.13). This yields the desired result (3.1). \square

Remark 3.3. *The results concerning the two scale expansions extend with only minor changes in the proofs to the case of coefficients $a(x, x/\epsilon)$ provided that for all β , $\partial_x^\beta a(x, y) \in L^\infty(\mathbb{R}^d \times \mathbb{T}^d)$. In this case, the operators χ_k are of order k but are no longer homogeneous. The coefficients $c_\alpha(x, y)$ satisfy $\partial_x^\beta c_\alpha \in L^\infty(\mathbb{R}_x^d, H^1(\mathbb{T}^d))$ by Lemma 2.3. The proofs of the leap frog structure and slow secular growth are unchanged. The residual estimate for the term II in (3.9) has a few additional terms treated using this regularity of c . For the criminal path, the case $a(x, x/\epsilon)$ is work in progress.*

4. THE CRIMINAL PATH

The criminal path, briefly presented in Section 1.2, yields approximations valid for times of order ϵ^{-N} for arbitrary N .

Main idea. *The criminal path changes the choice of the nonoscillatory parts πu_n . The oscillatory parts $(1 - \pi)u_n$ are given in terms of the nonoscillatory parts by (2.11) as in classical homogenization.*

We replace the traditional ansatz (1.2) for U by the criminal ansatz (1.10) for V . Since the terms v_n in (1.10) depend on ϵ , we commit the asymptotic crime of mixing different

orders in ϵ . The terms of order ϵ^2 in the criminal path are introduced in different but related ways in the seminal articles [23], [18].

4.1. Derivation of criminal equations. According to Definition 1.2 the criminal ansatz satisfies $v_0 = \pi v_0$, $\pi v_n = 0$ for $n \geq 1$ and $(I - \pi)v_n = \chi_n v_0$. The leading term $v_0 = \pi v_0$ and profile $V(\epsilon, t, x, y)$ are constructed so that the two formal identities

$$(4.1) \quad \left(a_2^*(\epsilon \partial_{t,x}) + \cdots + a_{2n}^*(\epsilon \partial_{t,x}) + \cdots \right) v_0(\epsilon, t, x) = \epsilon^2 f, \quad v_0 = 0 \text{ for } t < 0,$$

$$(4.2) \quad V(\epsilon, t, x, y) = \left(1 + \sum_{l=1}^{\infty} \epsilon^l \chi_l(y, \partial_{t,x}) \right) v_0(\epsilon, t, x),$$

are satisfied up to an acceptable error. Even if (4.1) is truncated to be a finite sum, it is high order in t . For each ϵ it usually defines an ill posed time evolution. In spite of this, the next sections construct functions v_0^k that satisfy (4.1) with small error.

Equation (4.1) can be understood in another way. Theorems 2.5 and 2.10 together with their remarks show that the standard homogenization hierarchy is equivalent to the pair of identities in the sense of formal power series,

$$\left(\sum_{n=1}^{\infty} a_{2n}^*(\epsilon \partial_{t,x}) \right) \left(\sum_{m=0}^{\infty} \epsilon^m \pi u_m \right) = \epsilon^2 f, \quad U = \left(1 + \sum_{l=1}^{\infty} \epsilon^l \chi_l \right) \left(\sum_{n=0}^{\infty} \epsilon^n \pi u_n \right).$$

Equivalently

$$(4.3) \quad \left(\sum_{n=1}^{\infty} a_{2n}^*(\epsilon \partial_{t,x}) \right) \pi U = \epsilon^2 f, \quad U = \left(1 + \sum_{l=1}^{\infty} \epsilon^l \chi_l \right) \pi U.$$

If one does **not** insist that πU be a formal power series in ϵ , this suggests the criminal equation (4.1) for $v_0 = \pi U$ and the criminal ansatz $V = U = \left(1 + \sum_{n=1}^{\infty} \epsilon^n \chi_n \right) v_0$.

In the next sections, equation (4.1) is converted to a normal form, truncated at order k and filtered, leading to the solutions v_0^k from Definition 1.2 with small enough error so that the approximation is very accurate.

4.2. Elimination algorithms. The algorithms of this section eliminate the time derivatives in (4.1), other than those in a_2^* , while changing the x -derivatives in the high order terms.

Proposition 4.1. *There are uniquely determined homogeneous operators $R_{2j}(\partial_{t,x})$ and $\tilde{a}_{2j}(\partial_x)$ of degree $2j$, the latter involving only ∂_x , so that (1.11) holds as an identity in the sense of formal power series.*

The heart of the proof is the following Lemma.

Lemma 4.2. *Suppose that $m \geq 2$ and $S_{2m}(\partial_{t,x})$ is homogeneous of degree $2m$ and contains only even powers of ∂_t . Then there exists a unique $r_{2m-2}(\partial_{t,x})$, homogeneous of degree $2m-2$, so that $r_{2m-2}a_2^* + S_{2m}$ is a differential operator in ∂_x only.*

Proof. Write

$$r_{2m-2}(\partial_{t,x}) = q_0 \partial_t^{2m-2} + q_2(\partial_x) \partial_t^{2m-4} + \cdots + q_{2m-4}(\partial_x) \partial_t^2 + q_{2m-2}(\partial_x).$$

The goal is to determine q_0, \dots, q_{2m-2} in such a way that $r_{2m-2}a_2^* + S_{2m}$ is a differential operator in ∂_x only. Order the terms in S_{2m} according to the order of the time derivative

$$S_{2m} = s_0(\partial_x) \partial_t^{2m} + s_2(\partial_x) \partial_t^{2m-2} + \cdots + s_{2m-2}(\partial_x) \partial_t^2.$$

Define $\underline{\rho} := \pi\rho$ and $a_2(\partial_x)$ so that a_2^* from (2.19) satisfies

$$(4.4) \quad a_2^*(\partial_{t,x}) = \underline{\rho} \partial_t^2 + a_2(\partial_x).$$

In particular $a_2(\partial_x)$ is second order in ∂_x . Then the terms containing time derivatives in $r_{2m-2}a_2^*$ are equal to

$$(4.5) \quad \underline{\rho} \left(q_0 \partial_t^{2m} + q_2(\partial_x) \partial_t^{2m-2} + \cdots + q_{2m-4}(\partial_x) \partial_t^4 + q_{2m-2}(\partial_x) \partial_t^2 \right) \\ + \left((q_0 a_2)(\partial_x) \partial_t^{2m-2} + (q_2 a_2)(\partial_x) \partial_t^{2m-4} + \cdots + (q_{2m-6} a_2)(\partial_x) \partial_t^4 + (q_{2m-4} a_2)(\partial_x) \partial_t^2 \right).$$

Regrouping in order of decreasing powers of ∂_t yields that (4.5) equals

$$\underline{\rho} q_0 \partial_t^{2m} + \left(\underline{\rho} q_2 + q_0 a_2 \right) (\partial_x) \partial_t^{2m-2} + \cdots \\ + \left(\underline{\rho} q_{2m-4} + q_{2m-6} a_2 \right) (\partial_x) \partial_t^4 + \left(\underline{\rho} q_{2m-2} + q_{2m-4} a_2 \right) (\partial_x) \partial_t^2.$$

The unique choice eliminating the time derivatives in $r_{2m-2}a_2^* + S_{2m}$ is given by

$$q_0 = -\underline{\rho}^{-1} s_0, \quad \text{and for } 1 \leq j \leq m-1, \quad q_{2j}(\partial_x) = -\underline{\rho}^{-1} \left(s_{2j}(\partial_x) + (q_{2j-2} a_2)(\partial_x) \right).$$

This completes the proof of Lemma 4.2. \square

Definition 4.3. *Denote by \mathcal{O}_N the set of constant coefficient partial differential operators in $\partial_{t,x}$ that are sums of terms homogeneous of degree at least N .*

The operators in \mathcal{O}_N are those whose symbols vanish to order N at the origin.

Proof of Proposition 4.1. In the next expressions \mathcal{O}_N represents an element of \mathcal{O}_N . The identity (1.11) holds if and only if for all $k \geq 2$

$$(4.6) \quad \left(1 + R_2 + R_4 + \cdots + R_{2k-2} \right) \left(a_2^* + a_4^* + \cdots + a_{2k}^* \right) = a_2^* + \tilde{a}_4 + \cdots + \tilde{a}_{2k} + \mathcal{O}_{2k+2}.$$

The goal is to find R_{2j} such that (4.6) holds. For $k=2$ expanding yields

$$(4.7) \quad \left(1 + R_2 \right) \left(a_2^* + a_4^* \right) = a_2^* + R_2 a_2^* + a_4^* + \mathcal{O}_6.$$

The term of order 4 is $R_2 a_2^* + a_4^*$. Choose R_2 using Lemma 4.2 as the unique homogeneous order 2 operator so that this fourth order term is independent of ∂_t . Denote by \tilde{a}_4 that differential operator.

The construction is recursive. Suppose that the R_2, \dots, R_{2k-2} and $\tilde{a}_4, \dots, \tilde{a}_{2k}$ have been uniquely determined so that (4.6) holds. We show that R_{2k} and \tilde{a}_{2k+2} are uniquely determined so that (4.6) is satisfied for $k+1$.

For $k+1$ the terms of order $\leq 2k$ on the right of (4.6) are only influenced by R_2, \dots, R_{2k-2} . Separating the lowest order term in \mathcal{O}_{2k+2} the right hand side of (4.6) can be written as

$$(4.8) \quad a_2^* + \tilde{a}_4 + \dots + \tilde{a}_{2k} + p_{2k+2} + \mathcal{O}_{2k+4},$$

where $p_{2k+2}(\partial_{t,x})$ is homogeneous of degree $2k+2$. To prove (4.6) for $k+1$ one must determine R_{2k} and \tilde{a}_{2k+2} such that

$$(4.9) \quad \left(1 + R_2 + R_4 + \dots + R_{2k-2} + R_{2k}\right) \left(a_2^* + a_4^* + \dots + a_{2k+2}^*\right) = a_2^* + \tilde{a}_4 + \dots + \tilde{a}_{2k+2} + \mathcal{O}_{2k+4}.$$

The term of order $2k+2$ is $R_{2k} a_2^* + p_{2k+2}$, where p_{2k+2} is given by (4.8), in terms of the R_{2j}, \tilde{a}_{2j} that are known from the inductive step. Lemma 4.2 shows that there is a unique R_{2k} so that this term of order $2k+2$ is independent of ∂_t . That is the uniquely determined R_{2k} and the operator in ∂_x is \tilde{a}_{2k+2} . The recursive construction is complete. \square

Remark 4.4. *The proof yields a recursive algorithm to compute R_{2j}, \tilde{a}_{2j} from the a_{2j}^* . The computation of the coefficients of a_{2j}^* requires the solution of $\sim d^{2j}$ cell problems.*

Remark 4.5. *Proposition 4.1 implies that if*

$$a_2^*(\partial_{t,x}) + a_4^*(\partial_{t,x}) + \dots + a_{2k}^*(\partial_{t,x}) = E(\partial_{t,x})$$

with $E(\partial_{t,x}) \in \mathcal{O}_{2k+2}$, then there is a $\tilde{E}(\partial_{t,x}) \in \mathcal{O}_{2k+2}$ so that

$$a_2^*(\partial_{t,x}) + \tilde{a}_4(\partial_x) + \dots + \tilde{a}_{2k}(\partial_x) = \tilde{E}(\partial_{t,x}).$$

The converse of Remark 4.5 is also true. In the ring of formal power series $1 + \sum_{j=1}^{\infty} R_{2j}(\partial_{t,x})$ has a unique multiplicative inverse

$$1 + \sum_{j=1}^{\infty} \tilde{R}_{2j}(\partial_{t,x}) = 1 + \sum_{k=1}^{\infty} \left(- \sum_{j=1}^{\infty} R_{2j}(\partial_{t,x}) \right)^k,$$

satisfying

$$\left(1 + \sum_{j=1}^{\infty} \tilde{R}_{2j}(\partial_{t,x})\right) \left(1 + \sum_{j=1}^{\infty} R_{2j}(\partial_{t,x})\right) = \left(1 + \sum_{j=1}^{\infty} R_{2j}(\partial_{t,x})\right) \left(1 + \sum_{j=1}^{\infty} \tilde{R}_{2j}(\partial_{t,x})\right) = 1.$$

The next Corollary is an immediate consequence.

Corollary 4.6. *Define $\tilde{R}^k(\partial_{t,x}) := \sum_{j=1}^k \tilde{R}_{2j}(\partial_{t,x})$. Then*

$$(4.10) \quad \begin{aligned} (1 + \tilde{R}^k(\partial_{t,x})) \left[a_2^*(\partial_{t,x}) + \tilde{a}_4(\partial_x) + \cdots + \tilde{a}_{2k+2}(\partial_x) \right] \\ = \left[a_2^*(\partial_{t,x}) + a_4^*(\partial_{t,x}) + \cdots + a_{2k+2}^*(\partial_{t,x}) \right] + \mathcal{O}_{2k+4}. \end{aligned}$$

4.3. Criminal equation with time derivatives eliminated. Having constructed the operators R_{2j} , the elimination algorithm is used to transform equation (4.1) to the normal form (1.12). Then equation (1.12) is truncated at order k yielding equation (1.13) repeated here,

$$\left[a_2^*(\partial_t, \partial_x) + \sum_{j=2}^{k+1} \epsilon^{2j-2} \tilde{a}_{2j}(\partial_x) \right] \underline{v}_0^k(\epsilon, t, x) = \left[1 + R^k(\epsilon \partial_{t,x}) \right] f, \quad \underline{v}_0^k = 0 \text{ for } t < 0.$$

The operator in brackets on the left may define an ill posed time evolution because the sign of the coefficients of the higher order space derivatives may be wrong (for example, it is known [13], [18], [4] that the operator \tilde{a}_4 has the wrong sign). This instability does not hinder the construction of good approximations. Committing an error, which is high order in ϵ , we filter the source term $[1 + R^k(\epsilon \partial_{t,x})]f$.

Choose cutoff functions $\psi_j \in C_0^\infty(\mathbb{R}^d)$ for $j = 1, 2$ with $\psi_1 = 1$ on a neighborhood of the origin and $\psi_2 = 1$ on a neighborhood of $\text{supp } \psi_1$. Choose $0 < \alpha < 1$. We compute a profile v_0^k that satisfies with $D := (1/i)\partial_x$,

$$(4.11) \quad \left[a_2^*(\partial_t, \partial_x) + \epsilon^2 \tilde{a}_4(\partial_x) + \cdots + \epsilon^{2k} \tilde{a}_{2k+2}(\partial_x) \right] v_0^k = \psi_1(\epsilon^\alpha D) (1 + R^k(\epsilon \partial_{t,x})) f.$$

The ill posed evolutions remain. However, for the filtered sources on the right in (4.11), there exist nice solutions. Fourier transformation in x yields ordinary differential equations in time parametrized by ξ for any tempered solution of (4.11). It shows that a tempered solution must have transform with support in $\epsilon^{-\alpha} \{\text{supp } \psi_1\}$. Such a solution satisfies $\psi_2(\epsilon^\alpha D) v_0^k = v_0^k$. Therefore it also satisfies

$$(4.12) \quad \left[a_2^*(\partial_t, \partial_x) + [\epsilon^2 \tilde{a}_4(\partial_x) + \cdots + \epsilon^{2k} \tilde{a}_{2k+2}(\partial_x)] \psi_2(\epsilon^\alpha D) \right] v_0^k = \psi_1(\epsilon^\alpha D) (1 + R^k(\epsilon \partial_{t,x})) f.$$

4.4. Stability Theorem. The operator applied to v_0^k in (4.12) is,

$$a_2^*(\partial_t, \partial_x) + M_\alpha(\epsilon, k, \partial_x), \quad M_\alpha(\epsilon, k, \partial_x) := \sum_{j=2}^{k+1} \epsilon^{2j-2} \tilde{a}_{2j}(\partial_x) \psi_2(\epsilon^\alpha D), \quad 0 < \alpha < 1.$$

The operator M_α depends on the choice of ψ_2 and α .

Theorem 4.7. *For any $k \in \mathbb{N}$ and any $\alpha_0 < 1$ there is an $\epsilon_0 > 0$ so that for each $\epsilon \leq \epsilon_0$ and any $0 < \alpha < \alpha_0$ the following holds.*

1. For any $g_0, g_1 \in H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ there is a unique solution v with $\partial_t^j v \in C(\mathbb{R}; H^{1-j}(\mathbb{R}^d))$ for $j \geq 0$ to

$$\left[a_2^*(\partial_{t,x}) + M_\alpha(\epsilon, k, \partial_x) \right] v = 0, \quad v(0, \cdot) = g_0, \quad \partial_t v(0, \cdot) = g_1.$$

This solution satisfies with a constant $C = C(\alpha_0, \epsilon_0, k)$ independent of $0 < \epsilon \leq \epsilon_0$ and $0 < \alpha \leq \alpha_0$,

$$\sup_{t \in \mathbb{R}} \left(\|\nabla_x v(t)\|_{L^2(\mathbb{R}^d)} + \|\partial_t v(t)\|_{L^2(\mathbb{R}^d)} \right) \leq C \left(\|\nabla_x v(0)\|_{L^2(\mathbb{R}^d)} + \|\partial_t v(0)\|_{L^2(\mathbb{R}^d)} \right).$$

2. For any $f \in L^1_{loc}([0, \infty[; L^2(\mathbb{R}^d))$ there is a unique w with $\{w, \nabla_{t,x} w\} \in C([0, \infty[, L^2(\mathbb{R}^d))$ satisfying

$$\left[a_2^*(\partial_{t,x}) + M_\alpha(\epsilon, k, \partial_x) \right] w = f, \quad w(0, \cdot) = \partial_t w(0, \cdot) = 0.$$

This solution satisfies with the same constant C from 1 and for all $t > 0$,

$$\|\nabla_{t,x} w(t)\|_{L^2(\mathbb{R}^d)} \leq C \|f\|_{L^1([0,t]; L^2(\mathbb{R}^d))}.$$

Proof. 1. M_α is bounded from $H^s \rightarrow H^{s+\sigma}$ for all s, σ with bound independent of s . The bound tends to infinity as $\epsilon \rightarrow 0$. The boundedness implies the existence statement of the Theorem. The uniform bounds lie deeper.

With the notation from (4.4), the equation $(a_2^* + M_\alpha)v = 0$ has the form $\underline{\rho}v_{tt} = -\mu(D)v$ with

$$(4.13) \quad -\mu(\xi) := a_2(i\xi) + \psi(\epsilon^\alpha \xi) \sum_{2 \leq j \leq k+1} \epsilon^{2j-2} \tilde{a}_{2j}(i\xi).$$

Each summand on the right is real valued. Statement 1. follows from the following estimate. For each $\alpha_0 \in]0, 1[$ there is an $\epsilon_0 > 0$ and constants $0 < c < C$ so that for all $0 < \epsilon \leq \epsilon_0$, $0 < \alpha \leq \alpha_0$ and all $\xi \in \mathbb{R}^d$,

$$(4.14) \quad c |\xi|^2 \leq -\mu(\xi) \leq C |\xi|^2.$$

By ellipticity of a_2 , there exists $0 < c_1 < C_1$ such that, for all $\xi \in \mathbb{R}^d$, $c_1 |\xi|^2 \leq a_2(i\xi) \leq C_1 |\xi|^2$. Therefore, to prove (4.14) it suffices to show that the modulus of the second summand on the right hand side of (4.13) is $o(|\xi|^2)$. In the support of the second summand $|\xi| \lesssim \epsilon^{-\alpha}$. Then

$$|\epsilon^{2j-2} \tilde{a}_{2j}(i\xi)| \lesssim \epsilon^{2j-2} |\xi|^{2j} = (\epsilon^\alpha |\xi|)^{2j-2} \epsilon^{(2j-2)(1-\alpha)} |\xi|^2$$

The first factor on the right is bounded and the second tends to zero as $\epsilon \rightarrow 0$. This proves the desired inequality. The Fourier transform of the solution satisfies

$$\underline{\rho} \frac{\partial^2 \widehat{v}}{\partial^2 t} + \mu(\xi) \widehat{v} = 0.$$

Multiplying by the complex conjugate of $\partial_t \widehat{v}$ and taking the real part proves the conservation laws

$$\forall \xi \in \mathbb{R}^d, \quad \frac{\partial}{\partial t} \left(\frac{\rho |\partial_t \widehat{v}(t, \xi)|^2}{2} + \frac{\mu(\xi) |\widehat{v}(t, \xi)|^2}{2} \right) = 0.$$

The estimate (4.14) implies that the conserved quantity

$$\int_{\mathbb{R}^d} \left(\rho |\partial_t \widehat{v}(t, \xi)|^2 + \mu(\xi) |\widehat{v}(t, \xi)|^2 \right) d\xi$$

is uniformly equivalent to $\|\partial_t v(t)\|_{L^2(\mathbb{R}^d)}^2 + \|\nabla_x v(t)\|_{L^2(\mathbb{R}^d)}^2$. This completes the proof of statement 1.

2. Uniqueness follows from 1. Multiplying by a constant it suffices to consider the case of a_2^* for which the coefficient of ∂_t^2 is equal to 1. Define $S(t, g_0, g_1) := v(t, \cdot)$, where v is provided by 1. A solution with the desired estimates is given by Duhamel's formula

$$w(t) := \int_0^t S(t-s, 0, f(s)) ds.$$

□

Corollary 4.8. *Let $0 < \alpha_0 < 1, k \in \mathbb{N}$ and $f \in H^\infty(\mathbb{R} \times \mathbb{R}^d)$ supported in $[0, 1] \times \mathbb{R}^d$. Then for and all $0 < \alpha \leq \alpha_0$ and $0 < \epsilon < \epsilon_0$ there is a unique solution $\zeta \in C^\infty(\mathbb{R}; H^\infty(\mathbb{R}^d))$ of (4.11). It satisfies $\text{supp } \widehat{\zeta} \subset \epsilon^{-\alpha} \text{supp } \psi_1$ and*

$$(4.15) \quad \sup_{t \in [0, \infty[} \left\| \partial_{t,x}^\beta \zeta \right\|_{L^2(\mathbb{R}^d)} \leq C(k, f, \beta) < \infty, \quad \beta \neq 0,$$

$$(4.16) \quad \left\| \zeta(t) \right\|_{L^2(\mathbb{R}^d)} \leq C(k, f) \langle t \rangle.$$

Proof. Uniqueness. Taking the Fourier transform shows that any solution $\zeta \in C^\infty(\mathbb{R}; H^\infty(\mathbb{R}^d))$ to (4.11) satisfies $\text{supp } \widehat{\zeta} \subset \epsilon^{-\alpha} \text{supp } \psi_1$. Therefore ζ also satisfies (4.12). The solutions of those equations are uniquely determined thanks to Theorem 4.7.

Existence. Define ζ as the solution to (4.12). The same Fourier transform argument as for uniqueness shows that this function satisfies $\text{supp } \widehat{\zeta} \subset \epsilon^{-\alpha} \text{supp } \psi_1$. In particular $\psi_2(\epsilon^\alpha D)\zeta = \zeta$. This implies that ζ satisfies (4.11). Theorem 4.7 implies that it has the additional properties claimed in Corollary 4.8 establishing existence. □

4.5. Criminal approximation error. This section performs the computations that are the main ingredients in the proof of Theorem 1.3.

Remark 4.9. *The profile $v_0^k = \pi v_0^k$ from Definition 1.2 satisfies $\psi_2(\epsilon^\alpha D)v_0^k = v_0^k$ so is the unique tempered solution supported in $t \geq 0$ of*

$$(4.17) \quad \left[a_2^*(\partial_{t,x}) + \left(\sum_{j=2}^{k+1} \epsilon^{2j-2} \widetilde{a}_{2j}(\partial_x) \right) \psi_2(\epsilon^\alpha D) \right] \pi v_0^k = [1 + R^k(\epsilon \partial_{t,x})] \psi_1(\epsilon^\alpha D) f.$$

§4.5.1 computes a precise formula for the residual.

4.5.1. *Formula for the residual.* Recall the criminal approximation V^k from (1.15),

$$V^k(\epsilon, t, x, y) = \sum_{n=0}^{2k+2} \epsilon^n v_n^k(\epsilon, t, x, y) = \left(I + \sum_{n=1}^{2k+2} \epsilon^n \chi_n(y, \partial_t, \partial_x) \right) v_0^k(\epsilon, t, x).$$

Define

$$(4.18) \quad Z(\epsilon, t, x, y) := \left[\rho(y) \partial_t^2 - \frac{1}{\epsilon^2} \mathcal{A}_{yy} - \frac{1}{\epsilon} \mathcal{A}_{xy} - \mathcal{A}_{xx} \right] V^k(\epsilon, t, x, y).$$

This is regrouped in powers of ϵ as if the v_n^k did not depend on ϵ .³ This yields

$$(4.19) \quad Z = \sum_{j=-2}^{2k+2} \epsilon^j Z_j = \sum_{j=-2}^{2k} \epsilon^j Z_j + \sum_{j=2k+1}^{2k+2} \epsilon^j Z_j =: \sum_{j=-2}^{2k} \epsilon^j Z_j + \mathcal{E}_1.$$

The term \mathcal{E}_1 is the first error term. It is estimated in Lemma 4.10. Theorem 2.5 shows that the definition of the nonoscillatory part of v_n^k is equivalent to

$$(4.20) \quad (I - \pi) Z_n = 0, \quad -2 \leq n \leq 2k.$$

Since $\pi v_n^k = 0$ for $1 \leq n \leq 2k+2$ and $v_0^k = \pi v_0^k$, Theorem 2.10 yields

$$\pi \sum_{j=-2}^{2k} \epsilon^j Z_j = \sum_{j=0}^{2k} \epsilon^j a_{j+2}^*(\partial_{t,x}) v_0^k = \epsilon^{-2} \sum_{j=0}^{2k} \epsilon^j a_{j+2}^*(\epsilon \partial_{t,x}) v_0^k.$$

Equation (4.10) of Corollary 4.6 implies

$$\begin{aligned} \pi \sum_{j=-2}^{2k} \epsilon^j Z_j &= \epsilon^{-2} (1 + \tilde{R}^k(\epsilon \partial_{t,x})) \left[a_2^*(\epsilon \partial_{t,x}) + \sum_{n=2}^{k+1} \tilde{a}_{2n}(\epsilon \partial_x) \right] v_0^k + \epsilon^{-2} \mathcal{O}_{2k+4}(\epsilon \partial_{t,x}) v_0^k, \\ &= (1 + \tilde{R}^k(\epsilon \partial_{t,x})) \left[a_2^*(\partial_{t,x}) + \sum_{n=2}^{k+1} \epsilon^{2n-2} \tilde{a}_{2n}(\partial_x) \right] v_0^k + \epsilon^{-2} \mathcal{O}_{2k+4}(\epsilon \partial_{t,x}) v_0^k. \end{aligned}$$

Since v_0^k satisfies equation (4.11),

$$\begin{aligned} \pi \sum_{j=-2}^{2k} \epsilon^j Z_j &= (1 + \tilde{R}^k(\epsilon \partial_{t,x})) \psi_1(\epsilon^\alpha D) (1 + R^k(\epsilon \partial_{t,x})) f + \epsilon^{-2} \mathcal{O}_{2k+4}(\epsilon \partial_{t,x}) v_0^k \\ &=: (1 + \tilde{R}^k(\epsilon \partial_{t,x})) (1 + R^k(\epsilon \partial_{t,x})) \psi_1(\epsilon^\alpha D) f + \mathcal{E}_2. \end{aligned}$$

³That is, the computations are made in the ring of Laurent expansions in ϵ whose coefficients are functions of ϵ, t, x, y . In $\epsilon^n v_n^k$, the function v_n^k is a coefficient of ϵ^n . If for instance $v_n^k = \epsilon^2$ the power from v_n^k must not be combined with the ϵ^n , the expression $\epsilon^n v_n^k$ is still a term in ϵ^n .

Use $(1 + \tilde{R}^k(\partial_{t,x}))(1 + R^k(\partial_{t,x})) = 1 + \mathcal{O}_{2k+2}$ to continue the computation,

$$\begin{aligned} (1 + \tilde{R}^k(\epsilon\partial_{t,x}))(1 + R^k(\epsilon\partial_{t,x}))\psi_1(\epsilon^\alpha D)f &= (1 + \mathcal{O}_{2k+2}(\epsilon\partial_{t,x}))\psi_1(\epsilon^\alpha D)f \\ &=: \psi_1(\epsilon^\alpha D)f + \mathcal{E}_3 \\ &= f + (\psi_1(\epsilon^\alpha D) - 1)f + \mathcal{E}_3 \\ &=: f + \mathcal{E}_4 + \mathcal{E}_3. \end{aligned}$$

Therefore,

$$(4.21) \quad Z(\epsilon, t, x, y) - f(t, x) = \mathcal{E}_1(\epsilon, t, x, y) + \sum_{j=2}^4 \mathcal{E}_j(\epsilon, t, x).$$

with

$$(4.22) \quad \begin{aligned} \mathcal{E}_1 &= \epsilon^{2k+1}Z_{2k+1} + \epsilon^{2k+2}Z_{2k+2}, & \mathcal{E}_2 &= \epsilon^{-2}\mathcal{O}_{2k+4}(\epsilon\partial_{t,x})v_0^k, \\ \mathcal{E}_3 &= \mathcal{O}_{2k+2}(\epsilon\partial_{t,x})\psi_1(\epsilon^\alpha D)f, & \mathcal{E}_4 &= (\psi_1(\epsilon^\alpha D) - 1)f. \end{aligned}$$

4.6. Residual estimates and proof of Theorem 1.3. The error $u^\epsilon(t, x) - V^k(\epsilon, t, x, x/\epsilon)$ satisfies

$$(4.23) \quad \begin{aligned} &\left[\rho(x/\epsilon)\partial_t^2 - \operatorname{div} a(x/\epsilon) \operatorname{grad} \right] \left(u^\epsilon(t, x) - V^k(\epsilon, t, x, x/\epsilon) \right) \\ &= \mathcal{E}_1(\epsilon, t, x, x/\epsilon) + \sum_{j=2}^4 \mathcal{E}_j(\epsilon, t, x). \end{aligned}$$

Lemma 4.10. *The error term $\mathcal{E}_1(\epsilon, t, x, x/\epsilon)$ from (4.22) is of the form*

$$\mathcal{E}_1(\epsilon, t, x, x/\epsilon) = f(\epsilon, t, x) + \operatorname{div} g(\epsilon, t, x)$$

with f, g satisfying uniformly in $t \geq 0$,

$$(4.24) \quad \|f(\epsilon, t, \cdot)\|_{L^2(\mathbb{R}^d)} + \|\partial_t g(\epsilon, t, \cdot)\|_{L^2(\mathbb{R}^d)} + \|g(\epsilon, t, \cdot)\|_{L^2(\mathbb{R}^d)} \lesssim \epsilon^{2k+2}.$$

Proof. As for the residual in the non criminal approximation, using $\pi v_{2k+1}^k = \pi v_{2k+2}^k$, write

$$\begin{aligned} \mathcal{E}_1(\epsilon, t, x, y) &= \epsilon^{2k+1}Z_{2k+1}(t, x, y) + \epsilon^{2k+2}Z_{2k+2}(t, x, y) \\ &= \left[\rho(y)\partial_t^2 - \mathcal{A}_{xx} \right] (I - \pi) \left(\epsilon^{2k+1}v_{2k+1}^k + \epsilon^{2k+2}v_{2k+2}^k \right) - \epsilon^{2k+1}\mathcal{A}_{xy}(I - \pi)v_{2k+2}^k \\ &= \epsilon^{2k+1} \left((\rho(y)\partial_t^2 - \mathcal{A}_{xx})v_{2k+1}^k - \mathcal{A}_{xy}v_{2k+2}^k - \mathcal{A}_{yy}v_{2k+3}^k \right) \\ &\quad + \epsilon^{2k+1}\mathcal{A}_{yy}v_{2k+3}^k + \epsilon^{2k+2}(\rho(y)\partial_t^2 - \mathcal{A}_{xx})v_{2k+2}^k. \end{aligned}$$

Moreover, by construction

$$(\rho(y)\partial_t^2 - \mathcal{A}_{xx})v_{2k+1}^k - \mathcal{A}_{xy}v_{2k+2}^k - \mathcal{A}_{yy}v_{2k+3}^k = a_{2k+3}^*\pi v_0^k = 0,$$

since the odd order operators have been shown to vanish. As a consequence,

$$\begin{aligned}\mathcal{E}_1(\epsilon, t, x, y) &= \epsilon^{2k+2}(\rho(y)\partial_t^2 - \mathcal{A}_{xx})v_{2k+2}^k + \epsilon^{2k+2}\left(\frac{1}{\epsilon}\mathcal{A}_{yy}v_{2k+3}^k\right) \\ &=: \epsilon^{2k+2}(I(\epsilon, t, x, y) + II(\epsilon, t, x, y)).\end{aligned}$$

In addition, for $l = 2k + 2, 2k + 3$,

$$v_l^k = \chi_l \pi v_0^k = \sum_{|\beta|=l} c_{\beta,l}(y) \partial_{t,x}^\beta \pi v_0^k$$

Use that $c_{\beta,2k+2}, c_{\beta,2k+3} \in H^1(\mathbb{T}_y^d)$ to show that $I(\epsilon, t, x, y)$ is a sum of terms of the form $c(y)v(t, x)$, $c \in L^2(\mathbb{T}^d)$. Corollary 4.8 implies that each v satisfies, uniformly in $t \geq 0$,

$$\|\partial_{t,x}^\alpha v(t, x)\|_{L^2(\mathbb{R}^d)} \lesssim 1$$

for $\alpha \in \mathbb{N}^{1+d}$, including $\alpha = 0$. Proposition B.1 in the appendix implies that uniformly for $t \geq 0$,

$$\|I(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} \lesssim 1.$$

The second term, $II(\epsilon, t, x, y)$, involves derivatives of $a(\cdot)$. Write

$$II(\epsilon, t, x, y) = \frac{1}{\epsilon} \operatorname{div}_y a(y) \operatorname{grad}_y \left[\sum_{|\beta|=2k+3} c_{\beta,2k+3}(y) \partial_{t,x}^\beta \pi v_0^k(\epsilon, t, x) \right].$$

Evaluate at $y = x/\epsilon$ to show that, with div acting on functions depending on x as well as on x/ϵ ,

$$\begin{aligned}II(\epsilon, t, x, x/\epsilon) &= \sum_{|\beta|=2k+3} \operatorname{div} \left[a(x/\epsilon) (\operatorname{grad}_y c_{\beta,2k+3})(x/\epsilon) \partial_{t,x}^\beta \pi v_0^k(\epsilon, t, x) \right] \\ &\quad - \sum_{|\beta|=2k+3} a(x/\epsilon) (\operatorname{grad}_y c_{\beta,2k+3})(x/\epsilon) \operatorname{grad} \partial_{t,x}^\beta \pi v_0^k(\epsilon, t, x) \\ &=: \operatorname{div} (II^{(1)}(\epsilon, t, x, x/\epsilon)) + II^{(2)}(\epsilon, t, x, x/\epsilon)\end{aligned}$$

As for I it follows that, uniformly in $t \geq 0$,

$$(4.25) \quad \begin{aligned}\|\partial_t II^{(1)}(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} + \|II^{(1)}(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} &\lesssim 1 \\ \|II^{(2)}(\epsilon, t, x, x/\epsilon)\|_{L^2(\mathbb{R}_x^d)} &\lesssim 1\end{aligned}$$

Defining $f(\epsilon, t, x) := \epsilon^{2k+2}(I + II^{(2)})(\epsilon, t, x, x/\epsilon)$ and $g(\epsilon, t, x) := \epsilon^{2k+2}II^{(1)}(\epsilon, t, x, x/\epsilon)$ completes the proof of Lemma 4.10. \square

End of proof of Theorem 1.3. Estimate the error terms $\mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4$ from (4.22). Since v_0^k and all of its derivatives are uniformly bounded in $L^2(\mathbb{R}_x^d)$, one has

$$(4.26) \quad \|\mathcal{E}_2(\epsilon, t, x)\|_{L^2(\mathbb{R}^d)} + \|\mathcal{E}_3(\epsilon, t, x)\|_{L^2(\mathbb{R}^d)} \lesssim \epsilon^{2k+2}.$$

The error from \mathcal{E}_4 is smaller. For any N one has

$$(4.27) \quad \|\mathcal{E}_4(\epsilon, t, x)\|_{L^2(\mathbb{R}^d)} \lesssim \epsilon^N.$$

Theorem 1.3 is a consequence of Lemma 4.10, (4.26), (4.27), and Proposition B.1 in the appendix. \square

Remark 4.11. *In the same way as in part iv of Remark 3.2, one finds a constant C depending on j but independent of $\epsilon \leq 1, t \geq 0$ so that*

$$(4.28) \quad \|\partial_t^j \nabla_{t,x}(u^\epsilon(t) - V^k(\epsilon, t, x, x/\epsilon))\|_{L^2(\mathbb{R}_x^d)} \leq C \epsilon^{2k+2} \langle t \rangle.$$

5. SOURCES GROWING POLYNOMIALLY IN TIME

The error estimates for sources with compact support in $0 < t < 1$ (Theorem 3.1 for the classical case and Theorem 1.3 in the criminal case) easily imply similar estimates (5.1) and (5.2) for sources that grow at most polynomially in time. Suppose that $f = 0$ for $t \leq 0$ and

$$\exists m, \forall \alpha, \exists C, \forall t, \quad \|\partial_{t,x}^\alpha f(t)\|_{L^2(\mathbb{R}_x^d)} \leq C t^m.$$

Use a partition of unity to write $f = \sum_{j=1}^\infty f_j$ with f_j supported in $[j-1, j]$ and such that

$$\exists C, \forall \alpha, j, \forall t, \quad \sup_{t \in \mathbb{R}} \|\partial_{t,x}^\alpha f_j(t)\|_{L^2(\mathbb{R}_x^d)} \leq C \sup_{j-1 \leq t \leq j} \|\partial_{t,x}^\alpha f(t)\|_{L^2(\mathbb{R}_x^d)}.$$

Denote by u_j^ϵ the exact solution with right hand side f_j and by $u_{j,\text{approx}}^\epsilon$ either the classical or the criminal approximation. Then $u^\epsilon = \sum u_j^\epsilon$ and $u_{\text{approx}}^\epsilon = \sum u_{j,\text{approx}}^\epsilon$. With $e_j^\epsilon(t) := \|\nabla_{t,x}[u_j^\epsilon(t) - u_{j,\text{approx}}^\epsilon(t)]\|_{L^2(\mathbb{R}_x^d)}$, the triangle inequality implies that the error in energy satisfies

$$\|\nabla_{t,x}[u^\epsilon(t) - u_{\text{approx}}^\epsilon(t)]\|_{L^2(\mathbb{R}_x^d)} \leq \sum_{j=1 \leq t} e_j^\epsilon(t).$$

5.1. Classical homogenization. Fix the index k in the classical approximation. Apply Theorem 3.1 with initial time shifted to $j-1$. Since f_j has size $O(j^m)$, and the error in Theorem 3.1 depends linearly on f ,

$$\exists C, \forall j, t, \quad e_j^\epsilon(t) \leq C \min \left\{ \epsilon^{2k+1} \langle (t - (j-1))_+ \rangle^{k+1}, \epsilon^{2k+2} \langle (t - (j-1))_+ \rangle^{k+2} \right\} j^m.$$

This implies

$$(5.1) \quad \begin{aligned} & \|\nabla_{t,x}[u^\epsilon(t) - u_{\text{approx}}^\epsilon(t)]\|_{L^2(\mathbb{R}_x^d)} \\ & \lesssim \min \left\{ \epsilon^{2k+1} \int_0^t \langle (t-s)_+ \rangle^{k+1} s^m ds, \epsilon^{2k+2} \int_0^t \langle (t-s)_+ \rangle^{k+2} s^m ds \right\} \\ & \lesssim \min \left\{ \epsilon^{2k+1} \langle t \rangle^{k+2+m}, \epsilon^{2k+2} \langle t \rangle^{k+3+m} \right\}. \end{aligned}$$

Given $N > 0$ and $0 < \delta < 2$, choosing k so large that $2k+1 \geq N + (k+2+m)(2-\delta)$ guarantees that the total error is $O(\epsilon^N)$ for times $t \leq 1/\epsilon^{2-\delta}$.

5.2. Criminal path. Applying Theorem 1.3 with a time shift yields $e_j^\epsilon(t) \lesssim \epsilon^{2k+2} \langle (t - (j - 1))_+ \rangle j^m$, so

$$(5.2) \quad \|\nabla_{t,x} [u^\epsilon(t) - u_{\text{approx}}^\epsilon(t)]\|_{L^2(\mathbb{R}_x^d)} \lesssim \epsilon^{2k+2} \int_0^t \langle (t-s)_+ \rangle s^m ds \lesssim \epsilon^{2k+2} \langle t \rangle^{m+2}.$$

Choosing k so large that $2k + 2 \geq N + N(m + 2)$ shows that the total error is $O(\epsilon^N)$ on intervals $t \leq \epsilon^{-N}$.

6. SKETCH OF SOME GENERALIZATIONS

6.1. Schrödinger equation. Consider the homogenization of Schrödinger's equation

$$i \rho(x/\epsilon) \partial_t u^\epsilon - \operatorname{div}(a(x/\epsilon) \operatorname{grad} u^\epsilon) = f(t, x).$$

With only the most minor modifications of the proof one finds an analogue of Theorem 2.13. This yields a leap frog structure and slow secular growth as in Theorem 2.15. The elimination is as for the wave equation case with ∂_t^2 replaced by $i\partial_t$. The result is

$$\left(I + \sum_{j \geq 1} R_{2j}(\partial_t, \partial_x) \right) \left(\sum_{n \geq 1} a_{2n}^*(\partial_t, \partial_x) \right) = a_2^*(\partial_t, \partial_x) + \tilde{a}_4(\partial_x) + \dots$$

where the a_{2n}^* denote the high order homogenizations of the Schrödinger operator. The R_{2j} , a_{2j}^* , and \tilde{a}_{2n} are equal to the polynomials from the wave equation case bearing the same name, but with ∂_t^2 replaced by $i\partial_t$. The polynomials R_{2j} and a_{2j}^* are homogeneous in the sense that they contain only monomials $\partial_t^m \partial_x^\alpha$ with $2m + |\alpha| = 2j$. The analogue of Proposition C.1 is that if $g \in L^\infty(\mathbb{R}; L^2(\mathbb{R}^d)) \cap L^1(\mathbb{R}; L^2(\mathbb{R}^d))$ vanishes for $t \leq 0$ and $g_t \in L^1(\mathbb{R}; L^2(\mathbb{R}^d))$, then the solution of $i \rho(x/\epsilon) \partial_t u = \operatorname{div}(a(x/\epsilon) \operatorname{grad} u) + \operatorname{div} g$ vanishing for $t \leq 0$ is uniformly bounded in $L^\infty(\mathbb{R}; H^1(\mathbb{R}^d))$. The classical approximation is accurate for times $\sim 1/\epsilon^{-2+\delta}$ for any small $\delta > 0$. The criminal approach yields approximations with error $\leq \epsilon^N$ for times $1/\epsilon^N$.

6.2. Systems of wave equations. For systems of wave equations the leap frog structure can fail. There are convincing numerical examples of systems of linear elastodynamics for which the third order homogenized operator does not vanish. The secular growth of u_n is as t^n instead of $t^{n/2}$. The standard approximation loses its validity at time scale $1/\epsilon$. A result that is true and reduces to the leap frog structure in the scalar case is that the operators $a_n^*(i\partial_t, i\partial_x)$ are all self-adjoint for both, odd and even n .

The elimination strategy encounters a difficulty. The term $\psi(\epsilon^\alpha D) \sum_{k \geq 3} \epsilon^k \tilde{a}_k(D_x)$ is a small real perturbation of the spatial part of a_2^* . However it need not be self-adjoint. In the scalar case self-adjointness is a consequence of reality. The lack of self-adjointness can lead to exponential growth in time destroying the stability proof. The system analogue of stability is work in progress.

APPENDIX A. EXAMPLES WITH MAXIMAL SECULAR GROWTH

A.1. Saturated secular growth. Theorem 2.15 gives an upper bound on the growth in time of the profiles. This appendix shows that the upper bound is attained for generic problems in dimension $d = 1$. Consider

$$\partial_t^2 u - \partial_x (a(x/\epsilon) \partial_x u) = f(t, x) \in C_0^\infty([0, 1] \times \mathbb{R}).$$

with $a(y)$ 1-periodic and not identically constant. Denote

$$a_2^*(\partial_t, \partial_x) := \partial_t^2 - c^2 \partial_x^2 = (\partial_x - c\partial_t)(\partial_x + c\partial_t),$$

the homogenized operator. Then $u_0 = u_0(t, x)$ is the solution of $a_2^*(\partial_t, \partial_x)u_0 = f$ that vanishes for $t < 0$. Therefore for $t > 1$ there are uniquely determined $g_0, h_0 \in C^\infty(\mathbb{R})$ with $g_0(s) = 0$ for $s \gg 1$ and $h_0(s) = 0$ for $s \ll -1$ so that for $t > 1$ one has

$$u_0 = g_0(x - ct) + h_0(x + ct).$$

For most f , both g_0 and h_0 are not identically equal to zero. This is true in particular if $f \geq 0$ and not identically equal to zero. In that case both g_0 and h_0 are non negative and not identically zero. The profile πu_2 satisfies of

$$a_2^*(\partial_t, \partial_x)\pi u_2 = -a_4^*(\partial_t, \partial_x)u_0.$$

The fourth order homogeneous polynomial $-a_4^*(\tau, \xi)$ contains no odd powers of τ . For $t > 1$, u_0 satisfies $a_2^*(\partial_t, \partial_x)u_0 = 0$. In this domain, replacing systematically $\partial_t^2 u_0$ by $c^2 \partial_x^2 u_0$ and equivalently τ^2 by $c^2 \xi^2$ yields a new polynomial $q(\xi)$ so that $-a_4^*(\partial_t, \partial_x)u_0 = q(\partial_x)u_0$. As soon as $a(y)$ is not constant one has $q(\partial_x) = \gamma \partial_x^4$ with $\gamma \neq 0$ (see [13]). This shows that

$$-a_4^*(\pm c, 1) = \gamma \neq 0.$$

The equation for πu_2 for $t > 1$ is

$$(A.1) \quad a_2^*(\partial_t, \partial_x)\pi u_2 = (\gamma \partial^4 g_0)(x - ct) + (\gamma \partial^4 h_0)(x + ct).$$

Writing $a_2^*(\partial_t, \partial_x) = (\partial_t - c\partial_x)(\partial_t + c\partial_x)$ one verifies that the function

$$z_2(t, x) = \frac{1}{4c^2} \left[(ct + x) (\gamma \partial^3 g_0)(x - ct) + (ct - x) (\gamma \partial^3 h_0)(x + ct) \right].$$

satisfies (A.1). Choose a cutoff function $\chi \in C^\infty(\mathbb{R}_t)$ equal to zero for $t < 1/2$ and equal to 1 for $t \geq 1$. Then

$$a_2^*(\partial_t, \partial_x) \left(\pi u_2 - \chi(t) z_2 \right)$$

is compactly supported in $0 \leq t \leq 1$. Therefore

$$\pi u_2 = \chi(t) z_2(t, x) + r_2(t, x), \quad \text{and,} \quad \forall \alpha, \partial_{t,x}^\alpha r_2 \in L^\infty(\mathbb{R}^{1+1}).$$

The equation for πu_4 is

$$a_2^*(\partial_t, \partial_x)\pi u_4 = -a_4^*(\partial_t, \partial_x)\pi u_2 - a_6^*(\partial_t, \partial_x)\pi u_0.$$

Only the πu_2 term on the right is unbounded. Furthermore, if any of the derivatives in $a_4^*(\partial_t, \partial_x)$ fall on the factors $ct \pm x$ in πu_2 the resulting function is bounded. Therefore

$$\begin{aligned} a_4^*(\partial_t, \partial_x) & \left[(ct+x)(\gamma\partial^3 g_0)(x-ct) + (ct-x)(\gamma\partial^3 h_0)(x+ct) \right] \\ & = (ct+x)(\gamma^2\partial^7 g_0)(x-ct) + (ct-x)(\gamma^2\partial^7 h_0)(x+ct) + L^\infty(\mathbb{R}^{1+1}). \end{aligned}$$

Reasoning as above yields with $\partial_{t,x}^\alpha r_3 \in L^\infty(\mathbb{R}^{1+1})$,

$$\pi u_4 = \frac{1}{(4c^2)^2} \left[\frac{(ct+x)^2}{2} ((\gamma\partial^3)^2 g_0)(x-ct) + \frac{(ct-x)^2}{2} ((\gamma\partial^3)^2 h_0)(x+ct) \right] + \langle t \rangle r_3.$$

An induction yields

$$\begin{aligned} \pi u_{2n} & = \frac{(ct+x)^n}{n!} g_n(x-ct) + \frac{(ct-x)^n}{n!} h_n(x+ct) + \langle t \rangle^{n-1} r_n, \\ g_n & := \left(\frac{\gamma\partial^3}{4c^2} \right)^n g_0, \quad h_n := \left(\frac{\gamma\partial^3}{4c^2} \right)^n h_0, \quad \partial_{t,x}^\alpha r_n \in L^\infty(\mathbb{R}^{1+1}). \end{aligned}$$

The u_n saturate the upper bounds of Theorem 2.15.

In addition note that u_{2n} grows with n as the $(3n)^{\text{th}}$ derivative of g_0, h_0 . This implies that for generic real analytic g_0, h_0 , the series $\sum \epsilon^n u_n$ is divergent.

A.2. Classic homogenization is inaccurate beyond $t = \epsilon^{-2}$. One could imagine that by including many correctors, the classical algorithm might be accurate for times beyond ϵ^{-2} . The example of the preceding subsection shows that that is not the case. For that example the exact solution satisfies $\sup_{0 < \epsilon < 1} \sup_{\mathbb{R}^{1+1}} |u| < \infty$. For $c > 0, \delta > 0$ as small as one likes define $t_\epsilon = c/\epsilon^{2+\delta}$. To show the inaccuracy of the classical approximation it suffices to show that for any $0 < N \in \mathbb{Z}$,

$$(A.2) \quad \lim_{\epsilon \rightarrow 0} \left\| \sum_{j=0}^{2N} \epsilon^j u_j(t_\epsilon) \right\|_{L^\infty(\mathbb{R}_x)} = \infty.$$

For the example one has for t large

$$(A.3) \quad \epsilon^{2j} t^j \lesssim \|\epsilon^{2j} \pi u_{2j}(t)\|_{L^\infty(\mathbb{R}_x)} \lesssim \epsilon^{2j} t^j.$$

For N fixed, formula (1.6) implies that

$$(A.4) \quad \|\epsilon^k (I - \pi) u_k(t)\|_{L^\infty(\mathbb{R}_x)} \lesssim \epsilon^{k-1} t^{k-1}.$$

Therefore, with $\lambda_\epsilon := \epsilon^{-\delta}$ one has

$$\left\| \epsilon^{2N} (I - \pi) u_{2N}(t_\epsilon) + \sum_{k=0}^{2N-1} \epsilon^k u_k(t_\epsilon) \right\|_{L^\infty(\mathbb{R}_x)} \lesssim \lambda_\epsilon^{2N-1} \quad \text{and} \quad \lambda_\epsilon^{2N} \lesssim \|\epsilon^{2N} \pi u_{2N}(t_\epsilon)\|_{L^\infty(\mathbb{R}_x)}.$$

It follows that with $C_j > 0$,

$$\left\| \sum_{j=0}^{2N} \epsilon^j u_j(t_\epsilon) \right\|_{L^\infty(\mathbb{R}_x)} \geq C_1 \lambda_\epsilon^{2N} - C_2 \lambda_\epsilon^{2N-1}.$$

The limit $\epsilon \rightarrow 0$ yields (A.2).

APPENDIX B. TWO SCALE L^2 ESTIMATE

This appendix contains a proof of a classical estimate for oscillating two scale functions. It is used in the error estimates in Sections 3.2 and 4.5.

Proposition B.1. *For each integer $s > d/2$, there is a constant C so that for all $v \in H^s(\mathbb{R}^d)$ and $c \in L^2(\mathbb{T}^d)$,*

$$(B.1) \quad \int_{\mathbb{R}^d} |v(x) c(x/\epsilon)|^2 dx \leq C \|c\|_{L^2(\mathbb{T}^d)}^2 \sum_{|\alpha| \leq s} \int_{\mathbb{R}^d} |(\epsilon \partial_x)^\alpha v(x)|^2 dx.$$

Proof. Denote by $Y := [0, 1)^d$ the unit box. Then \mathbb{R}^d is a disjoint union of boxes $Y_k := k + Y$, $\mathbb{R}^d = \bigcup_{k \in \mathbb{Z}^n} Y_k$. Scaling by ϵ yields $\mathbb{R}^d = \bigcup_{k \in \mathbb{Z}^n} \epsilon Y_k$. Sobolev's inequality for Y_k reads

$$(B.2) \quad \|w\|_{L^\infty(Y_k)}^2 \leq C(s, d) \sum_{|\alpha| \leq s} \int_{Y_k} |\partial_y^\alpha w(y)|^2 dy, \quad w \in H^s(Y_k).$$

When $y \in Y_k$, $x := \epsilon y \in \epsilon Y_k$. For $v \in H^s(\epsilon Y_k)$, apply (B.2) to $w(y) := v(\epsilon y) \in H^s(Y_k)$ to find

$$(B.3) \quad \|v\|_{L^\infty(\epsilon Y_k)}^2 \leq C(s, d) \epsilon^{-d} \sum_{|\alpha| \leq s} \int_{\epsilon Y_k} |(\epsilon \partial_x)^\alpha v(x)|^2 dx.$$

Estimate, using (B.3) in the last line,

$$\begin{aligned} \int_{\epsilon Y_k} |v(x) c(x/\epsilon)|^2 dx &\leq \|v\|_{L^\infty(\epsilon Y_k)}^2 \int_{\epsilon Y_k} |c(x/\epsilon)|^2 dx \\ &\leq C(s, d) \epsilon^{-d} \sum_{|\alpha| \leq s} \int_{\epsilon Y_k} |(\epsilon \partial_x)^\alpha v(x)|^2 dx \epsilon^d \|c\|_{L^2(\mathbb{T}^d)}^2. \end{aligned}$$

Summing over k yields (B.1). □

APPENDIX C. STABILITY ESTIMATE FOR THE WAVE EQUATION

This appendix contains an estimate for wave equations with sources in $L_{\text{loc}}^\infty(\mathbb{R}; H^{-1}(\mathbb{R}^d))$. The weak regularity in x is compensated by additional regularity in time. The residuals in the criminal and the non criminal approximation are of that form. For completeness the proof is included. The systems case is exactly analogous.

Proposition C.1. *Suppose that $0 < m_1 < m_2 < \infty$ are real numbers and $a, \rho \in L^\infty(\mathbb{R}^d)$ satisfy $m_1 \leq a, \rho \leq m_2$. There is a constant $C > 0$ depending only on m_1, m_2 so that for all $f \in L^1_{loc}(\mathbb{R}; L^2(\mathbb{R}^d))$ and $g \in L^\infty_{loc}(\mathbb{R}; L^2(\mathbb{R}^d; \mathbb{R}^d))$ with $\partial_t g \in L^1_{loc}(\mathbb{R}; L^2(\mathbb{R}^d; \mathbb{R}^d))$ and $f = g = 0$ for $t \leq 0$ the solution of*

$$(\rho \partial_t^2 - \operatorname{div}(a \operatorname{grad})) u = f + \operatorname{div} g, \quad u = 0 \quad \text{for } t \leq 0,$$

satisfies for all $t > 0$,

$$(C.1) \quad \|\nabla_{t,x} u(t)\|_{L^2(\mathbb{R}^d)} \leq C \left[\|f\|_{L^1([0,t]; L^2(\mathbb{R}^d))} + \|\partial_t g\|_{L^1([0,t]; L^2(\mathbb{R}^d))} + \|g\|_{L^\infty([0,t]; L^2(\mathbb{R}^d))} \right].$$

Proof. Approximating a, ρ, f, g by smooth functions it suffices to prove the estimate for solutions and right hand side that belong to $H^s([0, t] \times \mathbb{R}^d)$ for all $s, t > 0$ with a constant that depends only on the m_j . Introduce continuous functions

$$E(t) := \frac{1}{2} \int_{\mathbb{R}^d} \left(\rho(x) |\partial_t u(t, x)|^2 + a(x) |\nabla u(t, x)|^2 \right) dx, \quad \text{and} \quad M(t) := \sup_{0 \leq \underline{t} \leq t} E(\underline{t}).$$

Testing the equation with $\partial_t u$ yields the standard energy identity

$$E(t) = \int_0^t \int_{\mathbb{R}^d} \partial_t u (f + \operatorname{div} g) dx dt.$$

Estimate the first of the two summands on the right as

$$\left| \int_0^t \int_{\mathbb{R}^d} f \partial_t u dx dt \right| \leq \|f\|_{L^1([0,t]; L^2(\mathbb{R}^d))} \|\partial_t u\|_{L^\infty([0,t]; L^2(\mathbb{R}^d))} \leq C \|f\|_{L^1([0,t]; L^2(\mathbb{R}^d))} M(t)^{1/2}$$

with a constant depending only on the m_j . For the second summand two integrations by parts yield

$$\begin{aligned} \int_0^t \int_{\mathbb{R}^d} \operatorname{div} g \partial_t u dx dt &= - \int_0^t \int_{\mathbb{R}^d} g \cdot \operatorname{grad} \partial_t u dx dt \\ &= \int_0^t \int_{\mathbb{R}^d} \partial_t g \cdot \operatorname{grad} u dx dt - \int_{\mathbb{R}^d} g(t, x) \cdot \operatorname{grad} u(t, x) dx. \end{aligned}$$

Therefore

$$\begin{aligned} \left| \int_0^t \int_{\mathbb{R}^d} \operatorname{div} g \partial_t u dx dt \right| &\leq \\ &\|\partial_t g\|_{L^1([0,t]; L^2(\mathbb{R}^d))} \|\operatorname{grad} u\|_{L^\infty([0,t]; L^2(\mathbb{R}^d))} + \|g\|_{L^\infty([0,t]; L^2(\mathbb{R}^d))} \|\operatorname{grad} u\|_{L^\infty([0,t]; L^2(\mathbb{R}^d))} \\ &\leq C \left(\|\partial_t g\|_{L^1([0,t]; L^2(\mathbb{R}^d))} + \|g\|_{L^\infty([0,t]; L^2(\mathbb{R}^d))} \right) M(t)^{1/2}. \end{aligned}$$

Combining yields

$$(C.2) \quad E(t) \leq C \left(\|f\|_{L^1([0,t]; L^2(\mathbb{R}^d))} + \|\partial_t g\|_{L^1([0,t]; L^2(\mathbb{R}^d))} + \|g\|_{L^\infty([0,t]; L^2(\mathbb{R}^d))} \right) M(t)^{1/2}.$$

For each $t > 0$, choose $0 < \underline{t} \leq t$ so that $E(\underline{t}) = M(t)$. Estimate (C.2) at time \underline{t} yields

$$\begin{aligned} M(t) &\leq C \left(\|f\|_{L^1([0,\underline{t}]; L^2(\mathbb{R}^d))} + \|\partial_t g\|_{L^1([0,\underline{t}]; L^2(\mathbb{R}^d))} + \|g\|_{L^\infty([0,\underline{t}]; L^2(\mathbb{R}^d))} \right) M(\underline{t})^{1/2} \\ &\leq C \left(\|f\|_{L^1([0,t]; L^2(\mathbb{R}^d))} + \|\partial_t g\|_{L^1([0,t]; L^2(\mathbb{R}^d))} + \|g\|_{L^\infty([0,t]; L^2(\mathbb{R}^d))} \right) M(t)^{1/2}. \end{aligned}$$

If $M(t) \neq 0$, dividing by $M(t)^{1/2}$ yields (C.1). If $M(t) = 0$, (C.1) holds with $C = 0$. \square

Acknowledgments. G.A. is a member of the DEFI project at INRIA Saclay Ile-de-France.

REFERENCES

- [1] A. ABDULLE, M. GROTE, CH. STOHRER, *Finite element heterogeneous multiscale method for the wave equation: long-time effects*, Multiscale Model. Simul. **12** (2014), pp.1230-1257.
- [2] A. ABDULLE & T. N. POUCHON, *A priori error analysis of the finite element heterogeneous multiscale method for the wave equation in heterogeneous media over long time*, SIAM J. Numer. Anal., **54**(3), 1507-1534 (2016).
- [3] A. ABDULLE & T. N. POUCHON, *Effective models for the multidimensional wave equation in heterogeneous media over long time and numerical homogenization*, Mathematical Models & Methods In Applied Sciences, vol. 26, num. 14 (2016).
- [4] G. ALLAIRE, M. BRIANE & M. VANNINATHAN, *A comparison between two scale asymptotic expansions and Bloch wave expansions for the homogenization of periodic structures*, SEMA Journal **73**(3), 237-259 (2016).
- [5] G. ALLAIRE, M. PALOMBARO, J. RAUCH, *Diffractive Geometric Optics for Bloch Wave Packets*, Archive Rat. Mech. Anal., **202**, pp.373-426 (2011).
- [6] G. ALLAIRE, T. YAMADA, *Optimization of dispersive coefficients in the homogenization of the wave equation in periodic structures*, Numerische Mathematik, Volume 140, Issue 2, pp 265-326 (2018).
- [7] I. ANDRIANOV, V. BOLSHAKOV, V. DANISHEVSKYY, D. WEICHERT, *Higher order asymptotic homogenization and wave propagation in periodic composite materials*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **464**, no. 2093, 1181-1201 (2008).
- [8] N. BAKHVALOV, G. PANASENKO, *Homogenization: Averaging Processes in Periodic Media*, Kluwer, Dordrecht (1989).
- [9] A. BENOIT, A. GLORIA, *Long-time homogenization and asymptotic ballistic transport of classical waves*, Annales Scientifiques de l'École Normale Supérieure, **52**(3), 703-759 (2019).
- [10] A. BENSOUSSAN, J.-L. LIONS & G. PAPANICOLAOU, *Asymptotic analysis for periodic structures*, corrected reprint of the 1978 original, AMS Chelsea Publishing, Providence, RI, (2011).
- [11] S. BRAHIM-OTSMANE, G. FRANCFORT & F. MURAT, *Correctors for the homogenization of the wave and heat equations*, J. Math. Pures Appl. (9) **71**:197-231 (1992).
- [12] C. CONCA, R. ORIVE & M. VANNINATHAN, *Bloch approximation in homogenization and applications*, SIAM J. Math. Anal., **33** (5) (2002), 1166-1198.
- [13] C. CONCA, R. ORIVE & M. VANNINATHAN, *On Burnett coefficients in periodic media*, J. Math. Phys., **47** (3) (2006), 11 pp.
- [14] T. DOHNAL, A. LAMACZ & B. SCHWEIZER, *Bloch-wave homogenization on large time scales and dispersive effective wave equations*, Multiscale Model. Simul., **12** (2) (2014), 488-513.
- [15] J. FISH, W. CHEN, & G. NAGAI, *Non-local dispersive model for wave propagation in heterogeneous media: multi-dimensional case*, Internat. J. Numer. Methods Engrg., **54**(3):347-363, (2002).
- [16] V. JIKOV, S. KOZLOV, O. OLEINIK, *Homogenization of differential operators and integral functionals*, Springer, Berlin, (1995).

- [17] V. KAMOTSKI, K. MATTHIES, V. P. SMYSHLYAEV, *Exponential homogenization of linear second order elliptic PDEs with periodic coefficients*, SIAM J. Math. Anal. , **38** (5) (2006/07), 1565-1587
- [18] A. LAMACZ, *Dispersive effective models for waves in heterogeneous media*, Math. Models Methods Appl. Sci., **21** (9) (2011), 1871-1899.
- [19] D. LANNES, *High-frequency nonlinear optics: from the nonlinear Schrödinger approximation to ultrashort-pulses equations*, Proc. Roy. Soc. Edinburgh Sect. A **141** (2011), no. 2, 253-286.
- [20] S. E. PASTUKHOVA, *Approximations of the exponential of an operator with periodic coefficients. Problems in mathematical analysis*, J. Math. Sci. (N.Y.) **181** (2012), no. 5, 668-700.
- [21] J. RAUCH, *Hyperbolic partial differential equations and geometrics optics*, volume 133 of *Graduate Studies in Mathematics*, American Mathematical Society (2012).
- [22] E. SANCHEZ-PALENCIA, *Non-Homogeneous Media and Vibration Theory*, Springer Lecture Notes in Physics **129** (1980).
- [23] F. SANTOSA, W. SYMES, *A dispersive effective medium for wave propagation in periodic composites*, SIAM J. Appl. Math., **51** (1991), 984-1005.
- [24] V. P. SMYSHLYAEV, K. D. CHEREDNICHENKO, *On rigorous derivation of strain gradient effects in the overall behaviour of periodic heterogeneous media*, The J. R. Willis 60th anniversary volume. J. Mech. Phys. Solids, **48** (6-7) (2000), 1325-1357.
- [25] G. STRANG, *Variational crimes in the finite element method*, The mathematical foundations of the finite element method with applications to partial differential equations (Proc. Sympos., Univ. Maryland, Baltimore, Md., 1972), pp.689–710, Academic Press, New York (1972).