Commun. Math. Phys. 225, 1 – 32 (2002)

Communications in
**Mathematical
Physics**

On Convergence to Equilibrium Distribution, I. The Klein–Gordon Equation with Mixing

T. V. Dudnikova[1,⋆], A. I. Komech[2,⋆⋆], E. A. Kopylova[3,⋆⋆⋆], Yu. M. Suhov[4]

[1] Mathematics Department, Elektrostal Polytechnical Institute, Elektrostal, 144000 Russia.
E-mail: misis@elsite.ru
[2] Mechanics and Mathematics Department, Moscow State University, Moscow, 119899 Russia.
E-mail: komech@mech.math.msu.ru
[3] Physics and Applied Mathematics Department, Vladimir State University, Vladimir, Russia.
E-mail: ks@vpti.vladimir.ru
[4] Statistical Laboratory, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, UK. E-mail: Y.M.Suhov@statslab.cam.ac.uk

Received: 4 January 2001 / Accepted: 2 July 2001

Dedicated to M. I. Vishik on the occasion of his 80th anniversary

Abstract: Consider the Klein–Gordon equation (KGE) in $\mathbb{R}^n$, $n \geq 2$, with constant or variable coefficients. We study the distribution μ_t of the *random solution* at time $t \in \mathbb{R}$. We assume that the initial probability measure μ_0 has zero mean, a translation-invariant covariance, and a finite mean energy density. We also assume that μ_0 satisfies a Rosenblatt- or Ibragimov–Linnik-type mixing condition. The main result is the convergence of μ_t to a Gaussian probability measure as $t \to \infty$ which gives a Central Limit Theorem for the KGE. The proof for the case of constant coefficients is based on an analysis of long time asymptotics of the solution in the Fourier representation and Bernstein's "room-corridor" argument. The case of variable coefficients is treated by using an "averaged" version of the scattering theory for infinite energy solutions, based on Vainberg's results on local energy decay.

1. Introduction

The aim of this paper is to underline a special role of equilibrium distributions in statistical mechanics of systems governed by hyperbolic partial differential equations (for parabolic equations see [6, 27]). Important examples arise when one discusses the role of a *canonical* Gibbs distribution (CGD) in the Planck theory of spectral density of the black-body emission and in the Einstein–Debye quantum theory of solid state (see, e.g. [31]). [The word "canonical" is used in this paper to emphasize the fact that the probability distribution under consideration is formally related to the "Hamiltonian", or the energy functional, of the corresponding equation by the Gibbs exponential formula. Owing to the linearity of our equations, there are plenty of other first integrals which lead to other stationary measures.] Historically, the emission law was established

⋆ Supported partly by research grants of DFG (436 RUS 113/615/0-1) and RFBR (01-01-04002)
⋆⋆ Supported partly by the Institute of Physics and Mathematics of Michoacan in Morelia, the Max-Planck Institute for the Mathematics in Sciences (Leipzig) and by research grant of DFG (436 RUS 113/615/0-1)
⋆⋆⋆ Supported partly by research grant of RFBR (01-01-04002)

at a heuristic level by Kirchhoff in 1859 (see [34]) and stated formally by Planck in 1900 (see [25]). The law concerns the correspondence between the temperature and the colour of an emitting body (e.g., a burning carbon, or an incandescent wire in an electric bulb). Furthermore, it provides fundamental information on an interaction between the Maxwell field and "matter". Planck's formula specifies a "radiation intensity" $I_T(\omega)$ of the electromagnetic field at a fixed temperature $T > 0$, as a function of the frequency $\omega > 0$. It is convenient to treat $I_T(\cdot)$ as the spectral correlation function of a stationary random process. Then if g_T denotes an equilibrium distribution of this process, the Kirchhoff–Planck law suggests the long-time convergence

$$\mu_t \rightharpoondown g_T, \quad t \to \infty. \tag{1.1}$$

Here μ_t is the distribution at time t of a nonstationary random solution. The resulting equilibrium temperature T is determined by an initial distribution μ_0. Convergence to equilibrium (1.1) is also expected in a system of Maxwell's equations coupled to an equation of evolution of "matter". For example, both (1.1) and the Kirchhoff–Planck law should hold for the coupled Maxwell–Dirac equations [5], or for their second-quantised modifications. However, the rigorous proof here is still an open problem.

Previously, the convergence of type (1.1) to a CGD g_T has been established for an ideal gas with infinitely many particles by Sinai (see, e.g., [7]). Similar results were later obtained for other infinite-dimensional systems (see [2,10] and a survey [9]). For nonlinear wave problems, the first such result has been established by Jaksic and Pillet in [18]: they consider a system of a classical particle coupled to a wave field in a smooth nonlocal fashion. For all these models, the CGD g_T is well-defined, although the convergence is highly non-trivial. On the other hand, for the local coupling such as in the Maxwell–Dirac equations, the problem of "ultraviolet divergence" arises: the CGDs cannot be defined directly as the local energy is formally infinite almost surely. This is a serious technical difficulty that suggests that, to begin with, one should analyse convergence to *non-canonical* stationary measures μ_∞, with finite mean local energy:

$$\mu_t \rightharpoondown \mu_\infty, \quad t \to \infty. \tag{1.2}$$

In fact, most of the above-mentioned papers establish the convergence to both CGDs and non-canonical stationary measures, by using the same methods. In our situation, the aforementioned ultraviolet divergence makes the difference between (1.1) and (1.2).

In this paper we prove convergence (1.2) for the Klein–Gordon equation (KGE) in $\mathbb{R}^n$, $n \geq 2$:

$$\begin{cases} \ddot{u}(x,\ t) = \sum_{j=1}^n (\partial_j - iA_j(x))^2 u(x,t) - m^2 u(x,t), \quad x \in \mathbb{R}^n, \\ u|_{t=0} = u_0(x), \quad \dot{u}|_{t=0} = v_0(x). \end{cases} \tag{1.3}$$

Here $\partial_j \equiv \dfrac{\partial}{\partial x_j}$, $x \in \mathbb{R}^n$, $t \in \mathbb{R}$, $m > 0$ is a fixed constant and $(A_1(x), \ldots, A_n(x))$ a vector potential of an external magnetic field; we assume that functions $A_j(x)$ vanish outside a bounded domain. The solution $u(x,t)$ is considered as a complex-valued *classical* function.

It is important to identify a natural property of the initial measure μ_0 guaranteeing convergence (1.2). We follow an idea of Dobrushin and Suhov [10] and use a "space"-mixing condition of Rosenblatt- or Ibragimov–Linnik-type. Such a condition is natural from physical point of view. It replaces a "quasiergodic hypothesis" and allows us to avoid introducing a "thermostat" with a prescribed time-behaviour. Similar conditions have been used in [2,3,33,32]. In this paper, mixing is defined and applied in the context of the KGE.

Thus we prove convergence (1.2) for a class of initial measures μ_0 on a classical function space, with a finite mean local energy and satisfying a mixing condition. The limiting measure μ_∞ is stationary and turns out to be a Gaussian probability measure (GPM). Hence, this result is a form of the Central Limit Theorem for the KGE.

Another important question we discuss below is the relation of the limiting measure μ_∞ to the CGD g_T. The (formal) Klein–Gordon Hamiltonian is given by a quadratic form and so the CGDs g_T are also GPMs, albeit generalised (i.e. living in generalised function spaces). As our limiting measures μ_∞ are "classical" GPMs, they do not include CGDs. However, in the case of constant coefficients, a CGD can be obtained as a limit of measures μ_∞ as the "correlation radius" figuring in the mixing conditions imposed on μ_0 tends to zero. More precisely, we assume that for a fixed $T > 0$,

$$\frac{1}{2}E\Big(v_0(x)v_0(y) + \nabla u_0(x)\cdot\nabla u_0(y) + m^2u_0(x)u_0(y)\Big) \to T\delta(x-y), \quad r \to 0, \tag{1.4}$$

where E denotes the expectation. Then the covariance functions (CFs) of the corresponding limit GPM μ_∞ converge to the covariance functions of the CGD g_T. In turn, this implies the convergence

$$\mu_t \rightharpoondown \mu_\infty \sim g_T, \quad r \to 1. \tag{1.5}$$

See Sect. 4.

It should be noted that the existence of a "massive" (in a sense, infinite-dimensional) set of the limiting measures μ_∞ that are different from CGD's is related to the fact that KGE (1.3) is degenerate and admits infinitely many "additive" first integrals. Like the Klein–Gordon Hamiltonian, these integrals are quadratic forms; hence they generate GPMs via Gibbs exponential formulas.

Convergence (1.2) has been obtained in [19–21] for translation-invariant initial measures μ_0. However, the original proofs were too long and used a specific apparatus of Bessel's functions applicable exclusively in the case of the KGE. They have not been published in detail because of the lack of a unifying argument that could show the limits of the method and its forthcoming developments. To clarify the mechanism behind the results, one needed some new and robust ideas. The current work provides a modern approach applicable to a wide class of linear hyperbolic equations with a nondegenerate "dispersion relation", see Eq. (7.20) below. We also weaken considerably the mixing condition on measure μ_0. Moreover, our approach yields much shorter proofs and is applicable to non-translation invariant initial measures. The last fact is important in relation to the two-temperature problem [3,12,33] and the hydrodynamic limit [8]. Such progress became possible in large part owing to the systematic use of a Fourier transform (FT) and a duality argument of Lemma 7.1. [The importance of the Fourier transform was demonstrated in earlier works [3,32,33].]

Similar results, for the wave equation (WE) in $\mathbb{R}^n$ with odd $n \geq 3$, are established in [11] which develops the results [26]. The KGE shares some common features with the WE (which is formally obtained by setting $m = 0$ in (1.3)), and the exposition in [20, 21] followed the structure of the earlier work [26]. On the other hand, the KGE and WE also have serious differences, see below.

It is worth mentioning that possible extensions of our methods include, on the one hand, Dirac's and other relativistic-invariant linear hyperbolic equations and on the other hand harmonic lattices, as well as "coupled" systems of both types. We intend to return to these problems elsewhere.

We now pass to a detailed description of the results. Formal definitions and statements are given in Sect. 2. Set: $Y(t) = (Y^0(t), Y^1(t)) \equiv (u(\cdot,t), \dot u(\cdot,t))$, $Y_0 = (Y_0^0, Y_0^1) \equiv (u_0, v_0)$. Then (1.3) takes the form of an evolution equation

$$\dot Y(t) = \mathcal{A} Y(t), \quad t \in \mathbb{R}; \quad Y(0) = Y_0. \tag{1.6}$$

Here,

$$\mathcal{A} = \begin{pmatrix} 0 & 1 \\ A & 0 \end{pmatrix}, \tag{1.7}$$

where $A = \sum_{j=1}^n (\partial_j - iA_j(x))^2 - m^2$. We assume that the initial date Y_0 is a random element of a complex functional space $\mathcal{H}$ corresponding to states with a finite local energy, see Definition 2.1 below. The distribution of Y_0 is a probability measure μ_0 of mean zero satisfying some additional assumptions, see Conditions **S1**–**S3** below. Given $t \in \mathbb{R}$, denote by μ_t the measure that gives the distribution of $Y(t)$, the random solution to (1.6). We study the asymptotics of μ_t as $t \to \pm\infty$.

We identify $\mathbb{C} \equiv \mathbb{R}^2$ and denote by $\otimes$ the tensor product of real vectors. The CFs of the initial measure are supposed to be translation-invariant:

$$\begin{aligned} Q_0^{ij}(x,y) &:= E\Big(Y_0^i(x) \otimes Y_0^j(y)\Big) \\ &= q_0^{ij}(x-y), \quad x,y \in \mathbb{R}^n, \quad i,j = 0,1 \end{aligned} \tag{1.8}$$

(in fact our methods require a weaker assumption, but to simplify the exposition, we will not discuss it here). We also assume that the initial mean energy density is finite:

$$\begin{aligned} e_0 &:= E\Big(|v_0(x)|^2 + |\nabla u_0(x)|^2 + m^2|u_0(x)|^2\Big) \\ &= q_0^{11}(0) - \Delta q_0^{00}(0) + m^2 q_0^{00}(0) < \infty, \quad x \in \mathbb{R}^n. \end{aligned} \tag{1.9}$$

Finally, we assume that measure μ_0 satisfies a mixing condition of a Rosenblatt- or Ibragimov–Linnik type, which means that

$$Y_0(x) \quad \text{and} \quad Y_0(y) \ \text{ are } \textbf{asymptotically independent} \text{ as } \quad |x-y| \to \infty. \tag{1.10}$$

As was said before, our main result gives the (weak) convergence (1.2) of μ_t to a limiting measure μ_∞ which is a stationary GPM on $\mathcal{H}$. A similar convergence holds for $t \to -\infty$. Explicit formulas are then given for the CFs of μ_∞.

The strategy of the proof is as follows. First, we prove (1.2) for the equation with constant coefficients ($A_k(x) \equiv 0$), in three steps.

I. We check that the family of measures μ_t, $t \ge 0$, is weakly compact.

II. We check that the CFs converge to a limit: for $i, j = 0, 1$,

$$Q_t^{ij}(x,y) = \int Y^i(x) \otimes Y^j(y) \mu_t(dY) \to Q_\infty^{ij}(x,y), \quad t \to \infty. \tag{1.11}$$

III. Finally, we check that the characteristic functionals converge to a Gaussian one:

$$\hat{\mu}_t(\Psi) := \int \exp\{i\langle Y, \Psi\rangle\}\mu_t(dY) \to \exp\{-\frac{1}{2}\mathcal{Q}_\infty(\Psi, \Psi)\}, \quad t \to \infty. \tag{1.12}$$

Here Ψ is an arbitrary element of the dual space and $\mathcal{Q}_\infty$ the quadratic form with the integral kernel $(Q_\infty^{ij}(x, y))_{i,j=0,1}$; $\langle Y, \Psi\rangle$ denotes the scalar product in a real Hilbert space $L^2(\mathbb{R}^n) \otimes \mathbb{R}^n$.

Property I follows from the Prokhorov Theorem by a method used in [37]. First, we prove a uniform bound for the mean local energy in μ_t, using the conservation of mean energy density. The conditions of the Prokhorov Theorem are then checked by using Sobolev's embedding Theorem in conjunction with Chebyshev's inequality. Next, we deduce Property II from an analysis of oscillatory integrals arising in the FT. An important role is attributed to Proposition 6.1 reflecting the properties of the CFs in the FT deduced from the mixing condition.

On the other hand, the FT approach alone is not sufficient for proving Property III even in the case of constant coefficients. The reason is that a function of infinite energy corresponds to a singular generalised function in the FT, and the exact interpretation of the mixing condition (1.10) for such generalised functions is unclear. We deduce Property III from a representation of the solution in terms of the initial date in coordinate space. This is a modification of the approach adopted in [19–21]. It allows us to combine the mixing condition with the fact that waves in the coordinate space disperse to infinity. This leads to a representation of the solution as a sum of weakly dependent random variables. Then (1.12) follows from a Central Limit Theorem (CLT) under a Lindeberg-type condition. Checking such a condition is an important part of the proof.

It is useful to discuss the dispersive mechanism that is behind (1.12) and compare the KGE ($m > 0$) and WE ($m = 0$). Take, for simplicity, $n = 3$ and $u_0 \equiv 0$. The solution to (1.3) (with $A_k(x) \equiv 0$) is given by

$$u(x, t) = \int \mathcal{E}(x - y, t)v_0(y)\, dy, \quad t > 0, \tag{1.13}$$

where $\mathcal{E}$ is the "retarded" fundamental solution

$$\mathcal{E}(x, t) = \frac{1}{4\pi t}\delta(|x| - t) - \frac{m\theta(t - |x|)}{4\pi}\frac{J_1(m\sqrt{t^2 - x^2})}{\sqrt{t^2 - x^2}}, \tag{1.14}$$

J_1 is the Bessel function of the first order. For $m = 0$ the function $\mathcal{E}(\cdot, t)$ is supported by the sphere $|x| = t$ of area $\sim t^2$, and (1.13) becomes the Kirchhoff formula

$$u(x, t) = \frac{1}{4\pi t}\int\limits_{|x-y|=t} v_0(y)\, dS(y), \tag{1.15}$$

which manifests the dispersion of waves in the 3D space. Dividing the sphere $\{y \in \mathbb{R}^3 : |x - y| = t\}$ into $N \sim t^2$ "rooms" of a fixed width $d \gg 1$, we rewrite (1.15) as

$$u(x, t) \sim \frac{\sum\limits_{k=1}^{N} r_k}{\sqrt{N}}, \tag{1.16}$$

where r_k are nearly independent owing to the mixing condition. Then (1.2) follows by the well-known Bernstein "room-corridor" arguments.

For $m>0$ function $\mathcal{E}(\cdot,t)$ is supported by the ball $|x|\le t$ which means the absence of a *strong* Huyghen's principle for the KGE. The volume of the ball is $\sim t^3$, hence rewriting (1.13) in the form (1.16) would need asymptotics of the type

$$\mathcal{E}(x,t)=\mathcal{O}(t^{-3/2}),\quad |x|\le t \tag{1.17}$$

as $t\to\infty$. As $J_1(r)\sim\cos(r-3\pi/4)/\sqrt{r}$, asymptotics (1.17) only holds in the region $|x|\le vt$ with $v<1$. For instance,

$$\mathcal{E}\Big|_{|x|=vt}\sim\frac{\cos(m\gamma t-3\pi/4)}{(\gamma t)^{3/2}},$$

where $\gamma=\sqrt{1-v^2}$. However, the degree of the decay is different near the light cone $|x|=t$ corresponding to $v=1$ and $\gamma=0$. For example, for a fixed $r>0$,

$$\mathcal{E}\Big|_{|x|=t-r}\sim\frac{\cos(m\sqrt{2rt}-3\pi/4)}{(2rt)^{3/4}}=\mathcal{O}(t^{-3/4}), \tag{1.18}$$

where $r=t-|x|$ is the "distance" from the light cone. This illustrates that an application of Bernstein's method in the case of the KGE requires a new idea.

The key observation is that the asymptotics (1.18) displays oscillations $\sim\cos m\sqrt{2rt}$ of $\mathcal{E}$ near the light cone as $t\to\infty$. The solution becomes an oscillatory integral, and one is able to compensate the weak decay $\sim t^{-3/4}$ by a partial integration with Bessel functions, by a method following an argument from [23, Appendix B]. Such an approach was used in [21] and was accompanied by tedious computations in a combined "coordinate-momentum" representation. The approach adopted in this paper allows us to avoid this part of the argument. An important role is played by a duality argument of Lemma 7.1 leading to an analysis of an oscillatory integral with a phase function (="dispersion relation") with a nondegenerate Hessian, see (7.20).

Simple examples show that the convergence may fail when the mixing condition does not hold. For instance, take $u_0(x)\equiv\pm1$ and $v_0(x)\equiv0$ with probability $p_\pm=0.5$. Then the mean value is zero and (1.9) holds, but (1.10) does not. The solution $u(x,\ t)\equiv\pm\cos\ (mt)$ a.s., hence μ_t is periodic in time, and (1.2) fails.

Finally, a comment on the case of variable coefficients $A_k(x)$. In this case explicit formulas for the solution are unavailable. Here we construct a scattering theory for solutions of infinite global energy. This version of the scattering theory allows us to reduce the proof of (1.2) to the case of constant coefficients (this strategy is similar to [4,11,12]). In particular, in [11] one establishes, in the case of a WE, a long-time asymptotics

$$U(t)Y_0=\Theta U_0(t)Y_0+\rho(t)Y_0,\quad t>0. \tag{1.19}$$

Here $U(t)$ is the dynamical group of the WE with variable coefficients, $U_0(t)$ corresponds to the "free" equation with constant coefficients, and Θ is a "scattering operator". In this paper, instead of (1.19), we use a dual representation:

$$U'(t)\Psi=U_0'(t)W\Psi+r(t)\Psi,\quad t\ge0. \tag{1.20}$$

Here $U'(t)$ is a "formal adjoint" to the dynamical group of Eq. (1.3), while $U_0'(t)$ corresponds to the "free" equation, with $A_k(x) \equiv 0$. The remainder $r(t)$ is small in mean:

$$E|\langle Y_0, r(t)\Psi\rangle|^2 \to 0, \quad t \to \infty. \tag{1.21}$$

This version of scattering theory is essentially based on Vainberg's bounds for the local energy decay (see [35,36]).

Remark 1.1. (i) In [11] we deduce asymptotics (1.19) from its primal counterpart (1.20). In this paper we do not analyse connections between (1.20) and (1.19).

(ii) It is useful to comment on the difference between two versions of scattering theory produced for the WE and KGE. In the first theory, the remainders $\rho(t)$ and $r(t)$ are small a.s., while in the second theory, developed in this paper, $r(t)$ is small in mean (see (1.21)). Such a difference is related to a slow (power) decay of solutions to the KGE.

The main result of the paper is stated in Sect. 2 (see Theorem A). Sections 3–8 deal with the case of constant coefficients: the main statement is given in Sect. 3 (see Theorem B), the relation to CGDs is discussed in Sect. 4, the compactness (Property I) is established in Sect. 5, convergence (1.11) in Sect. 6, and convergence (1.12) in Sects. 7, 8. In Sect. 9 we check the Lindeberg condition needed for convergence to a Gaussian limit. In Sect. 10 we discuss the infinite energy version of the scattering theory, and in Sect. 11 convergence (1.2). In Appendix A we collected FT-type calculations. Appendix B is concerned with a formula on generalised GPMs on Sobolev spaces.

2. Main Results

2.1. Notation. We assume that functions $A_k(x)$ in (1.3) satisfy the following conditions:

E1. $A_j(x)$ are real C^∞-functions.
E2. $A_j(x) = 0$ for $|x| > R_0$, where $R_0 < \infty$.

E3. $\dfrac{\partial A_1}{\partial x_2} \not\equiv \dfrac{\partial A_2}{\partial x_1}$ if $n = 2$.

Assume that the initial state Y_0 belongs to the phase space $\mathcal{H}$ defined below.

Definition 2.1. *$\mathcal{H} \equiv H^1_{\rm loc}(\mathbb{R}^n) \oplus H^0_{\rm loc}(\mathbb{R}^n)$ is the Fréchet space of pairs $Y(x) \equiv (u(x), v(x))$ of complex functions $u(x)$, $v(x)$, endowed with local energy seminorms*

$$\|Y\|_R^2 = \int\limits_{|x|<R} \Big(|v(x)|^2 + |\nabla u(x)|^2 + m^2|u(x)|^2\Big)dx < \infty, \quad \forall R > 0. \tag{2.1}$$

Proposition 2.2 follows from [22, Thms. V.3.1, V.3.2]) as the speed of propagation for Eq. (1.3) is finite.

Proposition 2.2. (i) *For any $Y_0 \in \mathcal{H}$ there exists a unique (generalised) solution $Y(t) \in C(\mathbb{R}, \mathcal{H})$ to (1.6).*

(ii) *For any $t \in \mathbb{R}$ the operator $U(t) : Y_0 \mapsto Y(t)$ is continuous in $\mathcal{H}$.*

Let us choose a function $\zeta(x) \in C_0^\infty(\mathbb{R}^n)$ with $\zeta(0) \neq 0$. Denote by $H^s_{\rm loc}(\mathbb{R}^n)$, $s \in \mathbb{R}$, the local Sobolev spaces, i.e. the Fréchet spaces of distributions $u \in D'(\mathbb{R}^n)$ with finite seminorms

$$\|u\|_{s,R} := \|\Lambda^s\Big(\zeta(x/R)u\Big)\|_{L^2(\mathbb{R}^n)}, \tag{2.2}$$

where $\Lambda^s v := F^{-1}_{k\to x}(\langle k\rangle^s \hat{v}(k))$, $\langle k\rangle := \sqrt{|k|^2+1}$, and $\hat{v} := Fv$ is the FT of a tempered distribution v. For $\psi \in D$ define $F\psi(k) = \displaystyle\int e^{ik\cdot x}\psi(x)dx$.

Definition 2.3. *For $s \in \mathbb{R}$ denote $\mathcal{H}^s \equiv H^{1+s}_{\rm loc}(\mathbb{R}^n) \oplus H^s_{\rm loc}(\mathbb{R}^n)$.*

Using standard techniques of pseudodifferential operators (see, e.g. [16]) and Sobolev's Theorem, it is possible to prove that $\mathcal{H}^0 = \mathcal{H} \subset \mathcal{H}^{-\varepsilon}$ for every $\varepsilon > 0$, and the embedding is compact.

2.2. Random solution. Convergence to equilibrium. Let (Ω, Σ, P) be a probability space with expectation E and $\mathcal{B}(\mathcal{H})$ denote the Borel σ-algebra in $\mathcal{H}$. We assume that $Y_0 = Y_0(\omega, \cdot)$ in (1.6) is a measurable random function with values in $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$. In other words, $(\omega, x) \mapsto Y_0(\omega, x)$ is a measurable map $\Omega \times \mathbb{R}^n \to \mathbb{C}^2$ with respect to the (completed) σ-algebras $\Sigma \times \mathcal{B}(\mathbb{R}^n)$ and $\mathcal{B}(\mathbb{C}^2)$. Then, owing to Proposition 2.2, $Y(t) = U(t)Y_0$ is again a measurable random function with values in $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$. We denote by $\mu_0(dY_0)$ a probability measure on $\mathcal{H}$ giving the distribution of the Y_0. Without loss of generality, we assume $(\Omega, \Sigma, P) = (\mathcal{H}, \mathcal{B}(\mathcal{H}), \mu_0)$ and $Y_0(\omega, x) = \omega(x)$ for $\mu_0(d\omega) \times dx$-almost all $(\omega, x) \in \mathcal{H} \times \mathbb{R}^n$.

Definition 2.4. *μ_t is a probability measure on $\mathcal{H}$ which gives the distribution of $Y(t)$:*

$$\mu_t(B) = \mu_0(U(-t)B), \quad B \in \mathcal{B}(\mathcal{H}), \quad t \in \mathbb{R}. \tag{2.3}$$

Our main goal is to derive the weak convergence of the measures μ_t in the Fréchet space $\mathcal{H}^{-\varepsilon}$ for each $\varepsilon > 0$,

$$\mu_t \overset{\mathcal{H}^{-\varepsilon}}{\longrightarrow} \mu_\infty, \quad t \to \infty, \tag{2.4}$$

where μ_∞ is a limiting measure on the space $\mathcal{H}$. This means the convergence

$$\int f(Y)\mu_t(dY) \to \int f(Y)\mu_\infty(dY), \quad t \to \infty, \tag{2.5}$$

for any bounded continuous functional f on $\mathcal{H}^{-\varepsilon}$. Recall that we identify $\mathbb{C} \equiv \mathbb{R}^2$ and $\otimes$ stands for the tensor product of real vectors. Denote $M^2 = \mathbb{R}^2 \otimes \mathbb{R}^2$.

Definition 2.5. *The CFs of the measure μ_t are defined by*

$$Q_t^{ij}(x, y) \equiv E\Big(Y^i(x,t) \otimes Y^j(y,t)\Big), \quad i, j = 0, 1, \quad \text{for almost all } x, y \in \mathbb{R}^n \times \mathbb{R}^n, \tag{2.6}$$

assuming that the expectations in the RHS are finite.

We set $\mathcal{D} = D \oplus D$, and $\langle Y, \Psi\rangle = \langle Y^0, \Psi^0\rangle + \langle Y^1, \Psi^1\rangle$ for $Y = (Y^0, Y^1) \in \mathcal{H}$ and $\Psi = (\Psi^0, \Psi^1) \in \mathcal{D}$. For a probability measure μ on $\mathcal{H}$, denote by $\hat{\mu}$ the characteristic functional (FT)

$$\hat{\mu}(\Psi) \equiv \int \exp(i\langle Y, \Psi\rangle)\,\mu(dY), \quad \Psi \in \mathcal{D}.$$

A probability measure μ is called a GPM (of mean zero) if its characteristic functional has the form

$$\hat{\mu}(\Psi) = \exp\{-\frac{1}{2}\mathcal{Q}(\Psi, \Psi)\}, \quad \Psi \in \mathcal{D},$$

where $\mathcal{Q}$ is a real nonnegative quadratic form in $\mathcal{D}$. A measure μ is called translation-invariant if

$$\mu(T_h B) = \mu(B), \quad B \in \mathcal{B}(\mathcal{H}), \quad h \in \mathbb{R}^n,$$

where $T_h Y(x) = Y(x - h)$, $x \in \mathbb{R}^n$.

2.3. Mixing condition. Let $O(r)$ denote the set of all pairs of open bounded subsets $\mathcal{A}, \mathcal{B} \subset \mathbb{R}^n$ at distance $\text{dist}(\mathcal{A}, \mathcal{B}) \geq r$ and $\sigma(\mathcal{A})$ the σ-algebra in $\mathcal{H}$ generated by the linear functionals $Y \mapsto \langle Y, \Psi\rangle$, where $\Psi \in \mathcal{D}$ with $\text{supp}\,\Psi \subset \mathcal{A}$. Define the Ibragimov-Linnik mixing coefficient of a probability measure μ_0 on $\mathcal{H}$ by (cf. [17, Def. 17.2.2])

$$\varphi(r) \equiv \sup_{(\mathcal{A},\mathcal{B})\in O(r)} \sup_{\substack{A \in \sigma(\mathcal{A}),\, B \in \sigma(\mathcal{B}) \\ \mu_0(B) > 0}} \frac{|\mu_0(A \cap B) - \mu_0(A)\mu_0(B)|}{\mu_0(B)}. \tag{2.7}$$

Definition 2.6. *The measure μ_0 satisfies the strong, uniform Ibragimov–Linnik mixing condition if*

$$\varphi(r) \to 0, \quad r \to \infty. \tag{2.8}$$

Below, we specify the rate of decay of φ (see Condition S3).

2.4. Main assumptions and results. We assume that measure μ_0 has the following properties **S0–S3**:

S0. μ_0 has zero expectation value,

$$EY_0(x) \equiv 0, \quad x \in \mathbb{R}^n. \tag{2.9}$$

S1. μ_0 has translation-invariant CFs, i.e. Eq. (1.8) holds for almost all $x, y \in \mathbb{R}^n$.

S2. μ_0 has a finite mean energy density, i.e. Eq. (1.9) holds.

S3. μ_0 satisfies the strong uniform Ibragimov–Linnik mixing condition, with

$$\overline{\varphi} \equiv \int_0^\infty r^{n-1}\varphi^{1/2}(r)dr < \infty. \tag{2.10}$$

Define, for almost all $x, y \in \mathbb{R}^n$, the matrix $Q_\infty(x, y) \equiv \Big(Q_\infty^{ij}(x, y)\Big)_{i,j=0,1}$ by

$$Q_\infty(x, y) \equiv \frac{1}{2}\begin{pmatrix} (q_0^{00} + \mathcal{P} * q_0^{11})(x - y) & (q_0^{01} - q_0^{10})(x - y) \\ (q_0^{10} - q_0^{01})(x - y) & (q_0^{11} - (\Delta - m^2)q_0^{00})(x - y) \end{pmatrix}. \quad (2.11)$$

Here $\mathcal{P}(z)$ is the fundamental solution for the operator $-\Delta + m^2$, and $*$ stands for the convolution of generalized functions. We show below that $q_0^{11} \in L^2(\mathbb{R}^n)$ (see (6.1)). Then the convolution $\mathcal{P} * q_0^{11}$ in (2.11) also belongs to $L^2(\mathbb{R}^n)$.

Let $H = L^2(\mathbb{R}^n) \oplus H^1(\mathbb{R}^n)$ denote the space of complex valued functions $\Psi = (\Psi_0, \Psi_1)$ with a finite norm

$$\|\Psi\|_H^2 = \int\limits_{\mathbb{R}^n} (|\Psi_0(x)|^2 + |\nabla\Psi_1(x)|^2 + |\Psi_1(x)|^2)\, dx < \infty. \quad (2.12)$$

Denote by $\mathcal{Q}_\infty$ a real quadratic form in H defined by

$$\mathcal{Q}_\infty(\Psi, \Psi) = \sum_{i,j=0,1} \int\limits_{\mathbb{R}^n\times\mathbb{R}^n} \Big(Q_\infty^{ij}(x, y)\Psi^i(x), \Psi^j(y)\Big)dx\, dy, \quad (2.13)$$

where $\Big(\cdot, \cdot\Big)$ stands for the real scalar product in $\mathbb{C}^2 \equiv \mathbb{R}^4$. The form $\mathcal{Q}_\infty$ is continuous in H as the functions $Q_\infty^{ij}(x, y)$ are bounded.

Theorem A. *Let* $n \geq 2$, $m > 0$, *and assume that* E1–E3, S0–S3 *hold. Then*

(i) *The convergence in (2.4) holds for any* $\varepsilon > 0$.

(ii) *The limiting measure* μ_∞ *is a* GPM *on* $\mathcal{H}$.

(iii) *The characteristic functional of* μ_∞ *has the form*

$$\hat{\mu}_\infty(\Psi) = \exp\{-\frac{1}{2}\mathcal{Q}_\infty(W\Psi, W\Psi)\}, \quad \Psi \in \mathcal{D},$$

where $W : \mathcal{D} \to H$ *is a linear continuous operator.*

2.5. Remarks on conditions on the initial measure. (i) The (rather strong) form of mixing in Definition 2.6 is motivated by two facts: (a) it greatly simplifies the forthcoming arguments, (b) it allows us to produce an "optimal" (most slow) decay of φ indicating natural limits of Bernstein's room-corridor method. Condition (2.7) can be easily verified for GPMs with finite-range dependence and their images under "local" maps $\mathcal{H} \to \mathcal{H}$. See the examples in Sect. 2.6 below. (ii) The *uniform* Rosenblatt mixing condition [30] also suffices, together with a higher power > 2 in the bound (1.9): there exists $\delta > 0$ such that

$$E\Big(|v_0(x)|^{2+\delta} + |\nabla u_0(x)|^{2+\delta} + m^2|u_0(x)|^{2+\delta}\Big) < \infty. \quad (1.4')$$

Then (2.10) requires a modification:

$$\int_0^\infty r^{n-1}\alpha^p(r)dr < \infty, \quad \text{where} \quad p = \min\Big(\frac{\delta}{2+\delta}, \frac{1}{2}\Big), \quad (2.10')$$

where $\alpha(r)$ is the Rosenblatt mixing coefficient defined as in (2.7) but without $\mu(B)$ in the denominator. The statements of Theorem A and their proofs remain essentially unchanged, only Lemma 8.2 requires a suitable modification [17].

2.6. Examples of initial measures with mixing condition.

2.6.1. Gaussian measures. In this section we construct initial GPMs μ_0 satisfying **S0–S3**. Let μ_0 be a GPM on $\mathcal{H}$ with the characteristic functional

$$\hat{\mu}_0(\Psi) \equiv E \exp(i\langle Y, \Psi\rangle) = \exp\Big\{ -\frac{1}{2}\mathcal{Q}_0(\Psi, \Psi)\Big\}, \quad \Psi \in \mathcal{D}. \tag{2.14}$$

Here $\mathcal{Q}_0$ is a real nonnegative quadratic form with an integral kernel $(Q_0^{ij}(x, y))_{i,j=0,1}$. Let

$$Q_0^{ij}(x, y) \equiv q_0^{ij}(x - y), \tag{2.15}$$

for any $i,\ j$, where the function $q_0^{ij} \in C^2(\mathbb{R}^n) \otimes M^2$ has compact support. Then **S0**, **S1** and **S2** are satisfied; **S3** holds with $\varphi(r) \equiv 0$ for $r \geq r_0$ if $q_0^{ij}(z) \equiv 0$ for $|z| \geq r_0$. For a given matrix function $\left(q_0^{ij}(z)\right)$ such a measure exists on space $\mathcal{H}$ iff the corresponding FT is a nonnegative matrix-valued measure: $\left(\hat{q}_0^{ij}(k)\right) \geq 0$, $k \in \mathbb{R}^n$, [15, Thm V.5.1]. For example, all these conditions hold if $\hat{q}_0^{ij}(k) = D_i\delta^{ij} f(k_1)\cdot\cdots\cdot f(k_n)$ with $D_i \geq 0$ and

$$f(z) = \left(\frac{1-\cos(r_0 z/\sqrt{n})}{z^2}\right)^2, \quad z \in \mathbb{R}.$$

2.6.2. Non-Gaussian measures. Now choose a pair of odd functions $f^0,\ f^1 \in C^1(\mathbb{R})$, with bounded first derivatives. Define μ_0^* as the distribution of the random function $(f^0(Y^0(x)), f^1(Y^1(x)))$, where (Y^0, Y^1) is a random function with a Gaussian distribution μ_0 from the previous example. Then **S0–S3** hold for μ_0^* with a mixing coefficient $\varphi^*(r) \equiv 0$ for $r \geq r_0$. Measure μ_0^* is not Gaussian if $D_i > 0$ and the functions f^i are bounded and nonconstant.

3. Equations with Constant Coefficients

In Sects. 3–9 we assume that coefficients $A_k(x) \equiv 0$. Problem (1.3) then becomes

$$\begin{cases} \ddot{u}(x,t) = \Delta u(x,t) - m^2 u(x,t), & t \in \mathbb{R}, \\ u|_{t=0} = u_0(x), \ \ \dot{u}|_{t=0} = v_0(x). \end{cases} \tag{3.1}$$

As in (1.6), we rewrite (3.1) in the form

$$\dot{Y}(t) = \mathcal{A}_0 Y(t), \ \ t \in \mathbb{R}; \quad Y(0) = Y_0. \tag{3.2}$$

Here we denote

$$\mathcal{A}_0 = \begin{pmatrix} 0 & 1 \\ A_0 & 0 \end{pmatrix}, \tag{3.3}$$

where $A_0 = \Delta - m^2$. Denote by $U_0(t)$, $t \in \mathbb{R}$, the dynamical group for problem (3.2), then $Y(t) = U_0(t)Y_0$. The following proposition is well-known and is proved by a standard integration by parts.

Proposition 3.1. *Let $Y_0 = (u_0, v_0) \in \mathcal{H}$, and $Y(\cdot, t) = (u(\cdot, t), \dot{u}(\cdot, t)) \in C(\mathbb{R}, \mathcal{H})$ is the solution to (3.1). Then the following energy bound holds: for $R > 0$ and $t \in \mathbb{R}$,*

$$\int\limits_{|x|<R} \Big(|\dot{u}(x,t)|^2 + |\nabla u(x,t)|^2 + m^2|u(x,t)|^2\Big)dx \leq \int\limits_{|x|<R+|t|} \Big(|v_0(x)|^2 + |\nabla u_0(x)|^2 + m^2|u_0(x)|^2\Big)\,dx. \quad (3.4)$$

Set $\mu_t(B) = \mu_0(U_0(-t)B)$, $B \in \mathcal{B}(\mathcal{H})$, $t \in \mathbb{R}$. Then our main result for problem (3.2) is

Theorem B. *Let $n \geq 1$, $m > 0$, and Conditions* **S0–S3** *hold. Then the conclusions of Theorem A hold with $W = I$, and the limiting measure μ_∞ is translation-invariant.*

Theorem B can be deduced from Propositions 3.2 and 3.3 below, by the same arguments as in [37, Thm XII.5.2].

Proposition 3.2. *The family of measures $\{\mu_t, t \in \mathbb{R}\}$, is weakly compact in $\mathcal{H}^{-\varepsilon}$ with any $\varepsilon > 0$, and the bounds hold:*

$$\sup_{t\geq 0} E\|U_0(t)Y_0\|_R^2 < \infty, \quad R > 0. \quad (3.5)$$

Proposition 3.3. *For every $\Psi \in \mathcal{D}$,*

$$\hat{\mu}_t(\Psi) \equiv \int \exp(i\langle Y, \Psi\rangle)\mu_t(dY) \to \exp\Big\{-\frac{1}{2}\mathcal{Q}_\infty(\Psi,\Psi)\Big\}, \quad t \to \infty. \quad (3.6)$$

Propositions 3.2 and 3.3 are proved in Sects. 5 and 7–9, respectively. We will use repeatedly the FT (12.2) and (12.3) from Appendix A.

4. Relation to CGDs

In this section we discuss how our results are related to CGDs. We restrict consideration to the case of Eq. (1.3) with constant coefficients and to the translation-invariant isotropic case. The CGD g_T with the absolute temperature $T \geq 0$ is defined formally by

$$g_T(du \times dv) = \frac{1}{Z}\, e^{-\frac{H}{T}} \prod_x du(x)dv(x), \quad (4.1)$$

where $H := \frac{1}{2}\int \Big(|v(x)|^2 + |\nabla u(x)|^2 + m^2|u(x)|^2\Big)dx$, and Z is a normalisation constant.

To make the definition rigorous, let us introduce a scale of weighted Sobolev spaces $H^{s,\alpha}(\mathbb{R}^n)$ with arbitrary $s, \alpha \in \mathbb{R}$. We use notation (2.2).

Definition 4.1. (i) *$H^{s,\alpha}(\mathbb{R}^n)$ is the complex Hilbert space of the distributions $w \in S'(\mathbb{R}^n)$ with the finite norm*

$$\|w\|_{s,\alpha} \equiv \|\langle x\rangle^\alpha \Lambda^s w\|_{L_2(\mathbb{R}^n)} < \infty. \tag{4.2}$$

(ii) *$\mathcal{H}^{s,\alpha}$ is the Hilbert space of the pairs $Y = (u, v) \in H^{1+s,\alpha}(\mathbb{R}^n) \oplus H^{s,\alpha}(\mathbb{R}^n)$ with the norm*

$$|||Y|||_{s,\alpha} \equiv \|u\|_{1+s,\alpha} + \|v\|_{s,\alpha}. \tag{4.3}$$

Note that $\mathcal{H}^{\overline{s},\overline{\alpha}} \subset \mathcal{H}^{s,\alpha}$ if $\overline{s} < s$ and $\overline{\alpha} < \alpha$, and this embedding is compact. These facts follow by standard methods of pseudodifferential operators and Sobolev's Theorem (see, e.g. [16]).

Now we can define the CGDs rigorously: g_T is a GPM on a space $\mathcal{H}^{s,\alpha}$, $s, \alpha < -n/2$, with the CFs

$$\begin{aligned} &g_T^{00}(x-y) = T\mathcal{P}(x-y), \quad g_T^{11}(x-y) = T\delta(x-y),\\ &g_T^{01}(x-y) = g_T^{10}(x-y) = 0. \end{aligned} \tag{4.4}$$

By Minlos Theorem [15, Thm. V.5.1], such a measure exists on $\mathcal{H}^{s,\alpha}$ with $s, \alpha < -n/2$ as, *formally* (see Appendix B),

$$\int |||Y|||^2_{s,\alpha}\, g_T(dY) < \infty. \tag{4.5}$$

Measure g_T is stationary for the KGE, as its CFs are stationary; the last fact follows from formulas (12.6), (12.2). Also, g_T is translation invariant, so **S1** holds. Condition **S2** fails since the "mean energy density" $g_T^{11}(0) - \Delta g_T^{00}(0) + m^2 g_T^{00}(0)$ is infinite; this gives an "ultraviolet divergence". Mixing condition **S3** holds due to an exponential decay of the $\mathcal{P}(z)$. The convergence of type (1.1) holds for initial measures μ_0 that are absolutely continuous with respect to the CGD g_T, and the limit measure coincides with g_T. This mixing property (and even the K-property) can be proved by using well-known methods developed for Gaussian processes [7], and we do not discuss it here.

Remark. Assumption **S2** implies that $\mu_0(\mathcal{H}) = 1$ and hence $\mu_\infty(\mathcal{H}) = 1$. This excludes the case of a limiting CGD as it is a generalised GPM not supported by $\mathcal{H}$. However, it is possible to extend our results to a class of generalised initial measures converging to CGDs. For the case of constant coefficients such an extension could be done by smearing the initial generalised field as the dynamics commutes with the averaging (cf. [12]). For variable coefficients such an extension requires a further work.

To demonstrate the special role of the CGDs we consider a family of initial GPMs $\mu_{0,r}$, $r \in (0, 1]$, satisfying **S0–S3**, with the radius of correlation r. More precisely, suppose that the corresponding CFs $q_{0,r}^{ij}$ have the following properties **G0–G3**:

G0. $q_{0,r}^{01}(z) = q_{0,r}^{01}(-z), \quad z \in \mathbb{R}^n.$

G1. $q_{0,r}^{11}(z) - \Delta q_{0,r}^{00}(z) + m^2 q_{0,r}^{00}(z) = 0, \quad |z| \geq r.$

G2. For some $T>0$, $\displaystyle\frac{1}{2}\int\Big(q_{0,r}^{11}(z)-\Delta q_{0,r}^{00}(z)+m^2q_{0,r}^{00}(z)\Big)dz\to T,\quad r\to 0.$

G3. $\displaystyle\sup_{r\in(0,1]}\int\Big(|q_{0,r}^{11}(z)|+|\Delta q_{0,r}^{00}(z)|+m^2|q_{0,r}^{00}(z)|\Big)dz<\infty.$

Note that **G0** means a symmetry relation $Eu_0(x)v_0(y)=Eu_0(y)v_0(x)$ that holds for an isotropic measure where the CFs depend only on $|x-y|$. Examples of such a family will be provided later.

Properties **G0–G3** imply conditions **S0–S3** for the initial measures $\mu_{0,r}$. Therefore, Theorem B implies the convergence $\mu_{t,r}\overset{\mathcal{H}^{-\varepsilon}}{\longrightarrow}\mu_{\infty,r}$, $t\to\infty$, of type (2.4). The following proposition means that the limiting measure $\mu_{\infty,r}$ is close to CGD g_T on the Sobolev space of distributions $\mathcal{H}^{s,\alpha}$ with $s,\alpha<-n/2$.

Proposition 4.2. *Let Conditions* **G0–G3** *hold. Then corresponding limiting measures* $\mu_{\infty,r}$ *are concentrated on any space* $\mathcal{H}^{s,\alpha}$ *with* $s,\alpha<-n/2$ *and weakly converge to CGD* g_T *on the space* $\mathcal{H}^{s,\alpha}$*:*

$$\mu_{\infty,r}\overset{\mathcal{H}^{s,\alpha}}{\longrightarrow}g_T,\quad r\to 0. \tag{4.6}$$

Proof. The convergence follows by the same arguments as in [37] from two facts (cf. Propositions 3.2, 3.3): for any $\overline{s},\overline{\alpha}$ with $s<\overline{s}<-n/2$ and $\alpha<\overline{\alpha}<-n/2$,

(I) $$\sup_{r\in(0,1]}\int |\!|\!|Y|\!|\!|^2_{\overline{s},\overline{\alpha}}\mu_{\infty,r}(dY)<\infty.$$

(II) For $\Psi\in\mathcal{D}$, $\quad\mathcal{Q}_{\infty,r}(\Psi,\Psi)\to\mathcal{G}_T(\Psi,\Psi),\ r\to 0,$

where $\mathcal{Q}_{\infty,r}$ is the quadratic form with the integral kernel $\Big(q_\infty^{ij}(x-y)\Big)$, and $\mathcal{G}_T$ corresponds to $\Big(g_T^{ij}(x-y)\Big)$. It is important that the embedding $\mathcal{H}^{\overline{s},\overline{\alpha}}\subset\mathcal{H}^{s,\alpha}$ is compact. Property (I) can be checked with the help of formula (13.3) and by using the Parseval identity:

$$\int|\!|\!|Y|\!|\!|^2_{\overline{s},\overline{\alpha}}\mu_{\infty,r}(dY)=\frac{C(\overline{\alpha})}{(2\pi)^n}\int\Big(\langle k\rangle^{2\overline{s}}\,\mathrm{tr}\,\hat{q}_{\infty,r}^{11}(k)+\langle k\rangle^{2(1+\overline{s})}\,\mathrm{tr}\,\hat{q}_{\infty,r}^{00}(k)\Big)dk$$
$$=C(\overline{\alpha})\int\Big(f_1(z)\,\mathrm{tr}\,q_{\infty,r}^{11}(z)+f_2(z)(-\Delta+m^2)\mathrm{tr}\,q_{\infty,r}^{00}(z)\Big)dz,$$

where $f_1(z)=\displaystyle\frac{1}{(2\pi)^n}\int e^{-ikz}\langle k\rangle^{2\overline{s}}dk$ and $f_2(z)=\displaystyle\frac{1}{(2\pi)^n}\int e^{-ikz}\frac{\langle k\rangle^{2(1+\overline{s})}}{k^2+m^2}dk$. More precisely, Property (I) follows from **G3** as both functions $f_j(x)$ are bounded and continuous for $\overline{s}<-n/2$. Furthermore, **G0** and (2.11) imply that $q_{\infty,r}^{01}=q_{\infty,r}^{10}=0$, hence

$$\mathcal{Q}_{\infty,r}(\Psi,\Psi)=\int\Big(q_{\infty,r}^{00}(x-y)\Psi^0(x),\,\Psi^0(y)\Big)dxdy$$
$$+\int\Big(q_{\infty,r}^{11}(x-y)\Psi^1(x),\,\Psi^1(y)\Big)dxdy. \tag{4.7}$$

G1 and **G2** together imply that

$$q^{11}_{\infty,r}(x-y) \to T\delta(x-y), \quad r \to 0.$$

Then (2.11) implies

$$q^{00}_{\infty,r} = \mathcal{P} * q^{11}_{\infty,r} \to T\mathcal{P}, \quad r \to 0.$$

Therefore, Property (II) follows from (4.7): the justification follows easily in the FT space. The convergence of the covariance (II) provides the convergence of the measures (4.6) as all the measures are Gaussian. □

Example. Consider an initial measure μ_0 constructed in Example in Sect. 2.6.1. It satisfies Assumptions **S0–S3** and **G0**. Furthermore, $q_0^{00}(z) = q_0^{11}(z) = 0, \ |z| \geq 1$ if we choose $r_0 = 1$. Denote by $Y_0(x) = (Y_0^0(x), Y_0^1(x))$ a random function with distribution μ_0. Denote by $\mu_{0,r}$, $r > 0$, the distribution of the random function $Y_{0,r}(x) = (r^{1-\nu}Y_0^0(r^{-1}x), r^{-\nu}Y_0^1(r^{-1}x))$, where $\nu = n/2$. The corresponding CFs are $q^{ij}_{0,r}(z) = r^{2-n-i-j}q_0^{ij}(r^{-1}z)$. Then all Conditions G0–G3 hold with $T := \dfrac{1}{4}\displaystyle\int \operatorname{tr} q_0^{11}(z)dz$.

5. Compactness of the Family of Measures μ_t

This section gives the proof of bound (3.5). Proposition 3.2 will follow then with the help of the Prokhorov Theorem [37, Lemma II.3.1] as in the proof of [37, Thm. XII.5.2]. It is important that the embedding $\mathcal{H} \subset \mathcal{H}^{-\varepsilon}$ is compact, by virtue of Sobolev's Theorem, if $\varepsilon > 0$. Set:

$$e_t \equiv E\Big(|\dot{u}(x,t)|^2 + |\nabla u(x,t)|^2 + m^2|u(x,t)|^2\Big), \quad x \in \mathbb{R}^n. \tag{5.1}$$

The CFs of measure μ_t are translation invariant due to condition S1. Hence, taking expectation in (3.4), we get by S2,

$$e_t|B_R| \leq e_0|B_{R+t}| < \infty. \tag{5.2}$$

Here B_R is the ball $|x| \leq R$ in $\mathbb{R}^n$, and $|B_R|$ is its volume. Taking $R \to \infty$ we derive from (5.2) that $e_t \leq e_0$: in fact, the reversibility implies then $e_t = e_0$ (the mean energy density conservation). Hence, taking the expectation in (1.5), we get (3.5):

$$E\|U_0(t)Y_0\|^2_R = e_0|B_R| < \infty.$$

Corollary 5.1. *Bound (3.5) implies the convergence of the integrals in (2.6).*

6. Convergence of the Covariance Functions

In this section we check the convergence of the CFs of measures μ_t with the help of the FT. This convergence is used in Sect. 8.

6.1. Mixing in terms of the spectral density. The next proposition gives the mixing property in terms of the FT $\hat q_0^{ij}$ of the initial CFs q_0^{ij}. Assumption **S2** implies that $q_0^{ij}(z)$ is a measurable bounded function. Therefore, it belongs to the Schwartz space of tempered distributions as well as its FT.

Proposition 6.1. *Let the assumptions of Theorem B hold. Then* $\hat q_0^{ij} \in L^1(\mathbb{R}^n)\otimes M^2$, $\forall i, j$.

Proof. Step 1. First, let us prove that

$$\partial^\gamma q_0^{ij}(z) \in L^p(\mathbb{R}^n) \otimes M^2, \qquad p \ge 1, \ \ |\gamma| \le 2 - i - j. \tag{6.1}$$

Conditions **S0**, **S2** and **S3** imply, by [17, Lemma 17.2.3] (see Lemma 8.2 (i) below), that

$$|\partial^\gamma q_0^{ij}(z)| \le C e_0 \varphi^{1/2}(|z|), \ \ z \in \mathbb{R}^n. \tag{6.2}$$

The mixing coefficient φ is bounded, hence (6.2) and (2.10) imply (6.1):

$$\int |\partial^\gamma q_0^{ij}(z)|^p\, dz \le C e_0^p \int_{\mathbb{R}^n} \varphi^{p/2}(|z|)\, dz \le C_1 e_0^p \int_0^\infty r^{n-1}\varphi^{1/2}(r) dr < \infty. \tag{6.3}$$

Step 2. By Bohner's theorem, $\hat q_0 \equiv (\hat q_0^{ij}(k))dk$ is a complex positive-definite matrix-valued measure on $\mathbb{R}^n$, and **S2** implies that the total measure $\hat q_0(\mathbb{R}^n)$ is finite. On the other hand, (6.1) with $p = 2$ implies that $\hat q_0^{ij} \in L^2(\mathbb{R}^n) \otimes M^2$. □

6.2. Proof of convergence of covariance functions. Formulas (12.3), (12.2) and Proposition 6.1 imply for example,

$$q_t^{00}(x-y) := E\Big(u(x,t) \otimes u(y,t)\Big) \tag{6.4}$$

$$= \frac{1}{(2\pi)^n} \int e^{-ik(x-y)} \Big[\frac{1+\cos 2\omega t}{2} \hat q_0^{00}(k) + \frac{\sin 2\omega t}{2\omega} (\hat q_0^{01}(k) + \hat q_0^{10}(k)) + \frac{1-\cos 2\omega t}{2\omega^2} \hat q_0^{11}(k)\Big] dk,$$

where the integral converges and defines a continuous function determined for *all* $x, y \in \mathbb{R}^n$. Similar integrals give a convenient modification for all functions $q_t^{ij}(x-y)$, which we will work with.

Proposition 6.2. *Covariance functions* $q_t^{ij}(z)$, $i, j = 0, 1$, *converge for all* $z \in \mathbb{R}^n$:

$$q_t^{ij}(z) \to q_\infty^{ij}(z), \quad t \to \infty, \tag{6.5}$$

where functions $q_\infty^{ij}(z)$ *are defined in (2.11).*

Proof. Equation (6.4) and Proposition 6.1 imply,

$$q_t^{00}(z) \to \frac{1}{2}\Big(q_0^{00}(z) + \mathcal{P} * q_0^{11}(z)\Big) = q_\infty^{00}(z), \ \ t \to \infty, \tag{6.6}$$

as the oscillatory integrals tend to zero by the Lebesgue–Riemann Lemma. For other i, j the proof is similar. □

Note that $\mathcal{P}(z) \in L^1(\mathbb{R}^n)$. Therefore, (6.1) with $p = 1$ and explicit formulas (2.11) imply the following

Corollary 6.3. *Functions* q_∞^{ij} *belong to* $L^1(\mathbb{R}^n) \otimes M^2$, $i, j = 0, 1$.

Remark 6.4. A similar argument in the FT representation implies compactness in Proposition 3.2. We provided an independent proof of the compactness in Sect. 5 to show the relation with energy conservation.

7. Bernstein's Argument for the Klein–Gordon Equation

In this and the subsequent section we develop a version of Bernstein's "room-corridor" method. We use the standard integral representation for solutions, divide the domain of integration into "rooms" and "corridors" and evaluate their contribution. As a result, the value $\langle U_0(t)Y_0, \Psi\rangle$ for $\Psi \in \mathcal{D}$ is represented as the sum of weakly dependent random variables. We evaluate the variances of these random variables which will be important in next section.

First, we evaluate $\langle Y(t), \Psi\rangle$ in (3.6) by using duality arguments. For $t \in \mathbb{R}$, introduce "formal adjoint" operators $U_0'(t), U'(t)$ from space $\mathcal{D}$ to a suitable space of distributions. For example,

$$\langle Y, U_0'(t)\Psi\rangle = \langle U_0(t)Y, \Psi\rangle, \quad \Psi \in \mathcal{D}, \quad Y \in \mathcal{H}. \tag{7.1}$$

Denote $\Phi(\cdot, t) = U_0'(t)\Psi$. Then (7.1) can be rewritten as

$$\langle Y(t), \Psi\rangle = \langle Y_0, \Phi(\cdot, t)\rangle, \quad t \in \mathbb{R}. \tag{7.2}$$

The adjoint groups admit a convenient description. Lemma 7.1 below displays that the action of groups $U_0'(t)$, $U'(t)$ coincides, respectively, with the action of $U_0(t)$, $U(t)$, up to the order of the components. In particular, $U_0'(t)$, $U'(t)$ are continuous groups of operators $\mathcal{D} \to \mathcal{D}$.

Lemma 7.1. *For* $\Psi = (\Psi^0, \Psi^1) \in \mathcal{D}$,

$$U_0'(t)\Psi = (\dot\phi(\cdot, t), \phi(\cdot, t)), \quad U'(t)\Psi = (\dot\psi(\cdot, t), \psi(\cdot, t)), \tag{7.3}$$

where $\phi(x, t)$ *is the solution of Eq. (3.1) with the initial date* $(u_0, v_0) = (\Psi^1, \Psi^0)$ *and* $\psi(x, t)$ *is the solution of Eq. (1.3) with the initial state* $(u_0, v_0) = (\Psi^1, \Psi^0)$.

Proof. Differentiating (7.1) in t with $Y, \Psi \in \mathcal{D}$, we obtain

$$\langle Y, \dot U_0'(t)\Psi\rangle = \langle \dot U_0(t)Y, \Psi\rangle. \tag{7.4}$$

Group $U_0(t)$ has the generator (3.3). The generator of $U_0'(t)$ is the conjugate operator

$$\mathcal{A}_0' = \begin{pmatrix} 0 & A_0 \\ 1 & 0 \end{pmatrix}. \tag{7.5}$$

Hence, Eq. (7.3) holds with $\ddot\psi = A_0\psi$. For group $U'(t)$ the proof is similar. $\square$

Next we introduce a "room-corridor" partition of $\mathbb{R}^n$. Given $t > 0$, choose $d \equiv d_t \geq 1$ and $\rho \equiv \rho_t > 0$. Asymptotic relations between t, d_t and ρ_t are specified below. Set $h = d + \rho$ and

$$a^j = jh, \quad b^j = a^j + d, \quad j \in \mathbb{Z}. \tag{7.6}$$

We call the slabs $R_t^j = \{x \in \mathbb{R}^n : \ a^j \leq x^n \leq b^j\}$ "rooms" and $C_t^j = \{x \in \mathbb{R}^n : \ b^j \leq x^n \leq a^{j+1}\}$ "corridors". Here $x = (x^1, \dots, x^n)$, d is the width of a room, and ρ of a corridor.

Denote by χ_r the indicator of the interval $[0,\ d]$ and χ_c that of $[d,\ h]$ so that $\sum_{j\in\mathbb{Z}}(\chi_r(s-jh)+\chi_c(s-jh)) = 1$ for (almost all) $s \in \mathbb{R}$. The following decomposition holds:

$$\langle Y_0, \Phi(\cdot, t)\rangle = \sum_{j\in\mathbb{Z}} (\langle Y_0, \chi_r^j \Phi(\cdot, t)\rangle + \langle Y_0, \chi_c^j \Phi(\cdot, t)\rangle), \tag{7.7}$$

where $\chi_r^j := \chi_r(x^n - jh)$ and $\chi_c^j := \chi_c(x^n - jh)$. Consider random variables r_t^j, c_t^j, where

$$r_t^j = \langle Y_0, \chi_r^j \Phi(\cdot, t)\rangle, \quad c_t^j = \langle Y_0, \chi_c^j \Phi(\cdot, t)\rangle, \qquad j \in \mathbb{Z}. \tag{7.8}$$

Then (7.7) and (7.2) imply

$$\langle U_0(t) Y_0, \Psi\rangle = \sum_{j\in\mathbb{Z}} (r_t^j + c_t^j). \tag{7.9}$$

The series in (7.9) is indeed a finite sum. In fact, (7.5) and (12.1) imply that in the FT representation, $\dot{\hat{\Phi}}(k, t) = \hat{\mathcal{A}}_0'(k)\hat{\Phi}(k, t)$ and $\hat{\Phi}(k, t) = \hat{\mathcal{G}}_t'(k)\hat{\Psi}(k)$. Therefore,

$$\Phi(x, t) = \frac{1}{(2\pi)^n} \int\limits_{\mathbb{R}^n} e^{-ikx} \hat{\mathcal{G}}_t'(k)\hat{\Psi}(k)\, dk. \tag{7.10}$$

This can be rewritten as a convolution

$$\Phi(\cdot, t) = \mathcal{R}_t * \Psi, \tag{7.11}$$

where $\mathcal{R}_t = F^{-1}\hat{\mathcal{G}}_t'$. The support $\operatorname{supp}\Psi \subset B_{\overline{r}}$ with an $\overline{r} > 0$. Then the convolution representation (7.11) implies that the support of the function Φ at $t > 0$ is a subset of an "inflated future cone"

$$\operatorname{supp}\Phi \subset \{(x, t) \in \mathbb{R}^n \times \mathbb{R}_+ : \ |x| \leq t + \overline{r}\}, \tag{7.12}$$

as $\mathcal{R}_t(x)$ is supported by the "future cone" $|x| \leq t$. The last fact follows from general formulas (see [13, (II.4.5.12)]), or from the Paley–Wiener Theorem (see, e.g. [13, Thm. II.2.5.1]), as $\hat{\mathcal{R}}_t(k)$ is an entire function of $k \in \mathbb{C}^n$ satisfying suitable bounds. Finally, (7.8) implies that

$$r_t^j = c_t^j = 0 \qquad \text{for} \qquad jh + t < -\overline{r} \qquad \text{or} \qquad jh - t > \overline{r}. \tag{7.13}$$

Therefore, series (7.9) becomes a sum

$$\langle U_0(t) Y_0, \Psi\rangle = \sum_{-N_t}^{N_t} (r_t^j + c_t^j), \quad N_t \sim \frac{t}{h} \tag{7.14}$$

as $h \geq 1$.

Lemma 7.2. *Let* $n \geq 1$*,* $m > 0$*, and* **S0–S3** *hold. The following bounds hold for* $t > 1$*:*

$$E|r_t^j|^2 \leq C(\Psi)\, d_t/t, \quad E|c_t^j|^2 \leq C(\Psi)\, \rho_t/t, \qquad j \in \mathbb{Z}. \tag{7.15}$$

Proof. We discuss the first bound in (7.15) only; the second is done in a similar way.

Step 1. Rewrite the left-hand side as the integral of CFs. Definition (7.8) and Corollary 5.1 imply by Fubini's Theorem that

$$E|r_t^j|^2 = \langle \chi_r^j(x^n)\chi_r^j(y^n)q_0(x-y), \Phi(x,t) \otimes \Phi(y,t) \rangle. \tag{7.16}$$

The following bound holds true (cf. [29, Thm. XI.17 (b)]):

$$\sup_{x\in\mathbb{R}^n} |\Phi(x,t)| = \mathcal{O}(t^{-n/2}), \quad t \to \infty. \tag{7.17}$$

In fact, (7.10) and (12.2) imply that Φ can be written as the sum

$$\Phi(x,t) = \frac{1}{(2\pi)^n} \sum_{\pm} \int_{\mathbb{R}^n} e^{-i(kx\mp\omega t)} a^{\pm}(\omega)\hat{\Psi}(k)\, dk, \tag{7.18}$$

where $a^{\pm}(\omega)$ is a matrix whose entries are linear functions of ω or $1/\omega$. Let us prove the asymptotics (7.17) along each ray $x = vt + x_0$ with $|v| \leq 1$, then it holds uniformly in $x \in \mathbb{R}^n$ owing to (7.12). We have by (7.18),

$$\Phi(vt + x_0, t) = \frac{1}{(2\pi)^n} \sum_{\pm} \int_{\mathbb{R}^n} e^{-i(kv\mp\omega)t - ikx_0} a^{\pm}(\omega)\hat{\Psi}(k)\, dk. \tag{7.19}$$

This is a sum of oscillatory integrals with the phase functions $\phi_{\pm}(k) = kv \pm \omega(k)$. Each function has two stationary points, solutions to the equation $v = \pm\nabla\omega(k)$ if $|v| < 1$, and has none if $|v| \geq 1$. The phase functions are nondegenerate, i.e.

$$\det \left(\frac{\partial^2 \phi_{\pm}(k)}{\partial k_i \partial k_j} \right)_{i,j=1}^n \neq 0, \quad k \in \mathbb{R}^n. \tag{7.20}$$

At last, $\hat{\Psi}(k)$ is smooth and decays rapidly at infinity. Therefore, $\Phi(vt + x_0, t) = \mathcal{O}(t^{-n/2})$ according to the standard method of stationary phase, [14].

Step 2. According to (7.12) and (7.17), Eq. (7.16) implies that

$$E|r_t^j|^2 \leq Ct^{-n} \int_{|x|\leq t+\bar{r}} \chi_r^j(x^n)\|q_0(x-y)\|\, dxdy = Ct^{-n} \int_{|x|\leq t+\bar{r}} \chi_r^j(x^n)dx \int_{\mathbb{R}^n} \|q_0(z)\|dz, \tag{7.21}$$

where $\|q_0(z)\|$ stands for the norm of a matrix $\left(q_0^{ij}(z)\right)$. Therefore, (7.15) follows as $\|q_0(\cdot)\| \in L^1(\mathbb{R}^n)$ by (6.1). □

8. Convergence of Characteristic Functionals

In this section we complete the proof of Proposition 3.3. We use a version of the CLT developed by Ibragimov and Linnik. If $\mathcal{Q}_\infty(\Psi,\Psi)=0$, Proposition 3.3 is obvious. Thus, we may assume that for a given $\Psi\in\mathcal{D}$,

$$\mathcal{Q}_\infty(\Psi,\Psi)\neq 0. \tag{8.1}$$

Choose $0<\delta<1$ and

$$\rho_t\sim t^{1-\delta},\quad d_t\sim\frac{t}{\ln t},\qquad t\to\infty. \tag{8.2}$$

Lemma 8.1. *The following limit holds true:*

$$N_t\Big(\varphi(\rho_t)+\Big(\frac{\rho_t}{t}\Big)^{1/2}\Big)+N_t^2\Big(\varphi^{1/2}(\rho_t)+\frac{\rho_t}{t}\Big)\to 0,\quad t\to\infty. \tag{8.3}$$

Proof. Function $\varphi(r)$ is nonincreasing, hence by (2.10),

$$r^n\varphi^{1/2}(r)=n\int\limits_0^r s^{n-1}\varphi^{1/2}(r)\,ds\le n\int\limits_0^r s^{n-1}\varphi^{1/2}(s)\,ds\le C\overline{\varphi}<\infty. \tag{8.4}$$

Then Eq. (8.3) follows as (8.2) and (7.14) imply that $N_t\sim\ln t$. □

By the triangle inequality,

$$\begin{aligned}|\hat\mu_t(\Psi)-\hat\mu_\infty(\Psi)|\le&\Big|E\exp\{i\langle U_0(t)Y_0,\Psi\rangle\}-E\exp\Big\{i\sum\nolimits_t r_t^j\Big\}\Big|\\&+\Big|\exp\Big\{-\frac12\sum\nolimits_t E|r_t^j|^2\Big\}-\exp\Big\{-\frac12\mathcal{Q}_\infty(\Psi,\Psi)\Big\}\Big|\\&+\Big|E\exp\Big\{i\sum\nolimits_t r_t^j\Big\}-\exp\Big\{-\frac12\sum\nolimits_t E|r_t^j|^2\Big\}\Big|\\\equiv& I_1+I_2+I_3,\end{aligned} \tag{8.5}$$

where the sum $\sum_t$ stands for $\sum\limits_{j=-N_t}^{N_t}$. We are going to show that all summands I_1, I_2, I_3 tend to zero as $t\to\infty$.

Step (i). Equation (7.14) implies

$$I_1=\Big|E\exp\Big\{i\sum\nolimits_t r_t^j\Big\}(\exp\Big\{i\sum\nolimits_t c_t^j\Big\}-1)\Big|\le\sum\nolimits_t E|c_t^j|\le\sum\nolimits_t(E|c_t^j|^2)^{1/2}. \tag{8.6}$$

From (8.6), (7.15) and (8.3) we obtain that

$$I_1\le CN_t(\rho_t/t)^{1/2}\to 0,\ \ t\to\infty. \tag{8.7}$$

Step (ii). By the triangle inequality,

$$\begin{aligned}I_2\le&\frac12\Big|\sum\nolimits_t E|r_t^j|^2-\mathcal{Q}_\infty(\Psi,\Psi)\Big|\le\frac12\,|\mathcal{Q}_t(\Psi,\Psi)-\mathcal{Q}_\infty(\Psi,\Psi)|\\&+\frac12\Big|E\Big(\sum\nolimits_t r_t^j\Big)^2-\sum\nolimits_t E|r_t^j|^2\Big|+\frac12\Big|E\Big(\sum\nolimits_t r_t^j\Big)^2-\mathcal{Q}_t(\Psi,\Psi)\Big|\\\equiv& I_{21}+I_{22}+I_{23},\end{aligned} \tag{8.8}$$

where $\mathcal{Q}_t$ is a quadratic form with the integral kernel $\left(Q_t^{ij}(x, y)\right)$. Equation (6.5) implies that $I_{21} \to 0$. As to I_{22}, we first have that

$$I_{22} \le \sum_{j<l} E|r_t^j r_t^l|. \tag{8.9}$$

The next lemma is a corollary of [17, Lemma 17.2.3].

Lemma 8.2. *Let ξ be a complex random variable measurable with respect to the σ-algebra $\sigma(\mathcal{A})$, η with respect to the σ-algebra $\sigma(\mathcal{B})$, and the distance* $\mathrm{dist}(\mathcal{A}, \mathcal{B}) \ge r > 0$.

(i) *Let $(E|\xi|^2)^{1/2} \le a$, $(E|\eta|^2)^{1/2} \le b$. Then*

$$|E\xi\eta - E\xi E\eta| \le Cab\, \varphi^{1/2}(r).$$

(ii) *Let $|\xi| \le a$, $|\eta| \le b$ a.s. Then*

$$|E\xi\eta - E\xi E\eta| \le Cab\, \varphi(r).$$

We apply Lemma 8.2 to deduce that $I_{22} \to 0$ as $t \to \infty$. Note that $r_t^j = \langle Y_0(x), \chi_r^j(x^n)(\mathcal{R}_t * \Psi)\rangle$ is measurable with respect to the σ-algebra $\sigma(R_t^j)$. The distance between the different rooms R_t^j is greater than or equal to ρ_t according to (7.6). Then (8.9) and **S1**, **S3** imply, together with Lemma 8.2 (i), that

$$I_{22} \le CN_t^2 \varphi^{1/2}(\rho_t), \tag{8.10}$$

which goes to 0 as $t \to \infty$ because of (7.15) and (8.3). Finally, it remains to check that $I_{23} \to 0$, $t \to \infty$. By the Cauchy–Schwartz inequality,

$$\begin{aligned} I_{23} &\le \left|E\Big(\sum_t r_t^j\Big)^2 - E\Big(\sum_t r_t^j + \sum_t c_t^j\Big)^2\right| \\ &\le CN_t \sum_t E|c_t^j|^2 + C\Big(E\Big(\sum_t r_t^j\Big)^2\Big)^{1/2}\Big(N_t \sum_t E|c_t^j|^2\Big)^{1/2}. \end{aligned} \tag{8.11}$$

Then (7.15), (8.9) and (8.10) imply

$$E\Big(\sum_t r_t^j\Big)^2 \le \sum_t E|r_t^j|^2 + 2\sum_{j<l} E|r_t^j r_t^l| \le CN_t d_t/t + C_1 N_t \varphi^{1/2}(\rho_t) \le C_2 < \infty.$$

Now (7.15), (8.11) and (8.3) yield

$$I_{23} \le C_1 N_t^2 \rho_t/t + C_2 N_t(\rho_t/t)^{1/2} \to 0, \quad t \to \infty. \tag{8.12}$$

So, all terms I_{21}, I_{22}, I_{23} in (8.8) tend to zero. Then (8.8) implies that

$$I_2 \le \frac{1}{2}\left|\sum_t E|r_t^j|^2 - \mathcal{Q}_\infty(\Psi, \Psi)\right| \to 0, \quad t \to \infty. \tag{8.13}$$

Step (iii). It remains to verify that

$$I_3 = \left|E \exp\Big\{i\sum_t r_t^j\Big\} - \exp\Big\{-\frac{1}{2}E\Big(\sum_t r_t^j\Big)^2\Big\}\right| \to 0, \quad t \to \infty. \tag{8.14}$$

Using Lemma 8.2 (ii) yields:

$$
\begin{aligned}
&\Big|E\exp\Big\{i\sum\nolimits_t r_t^j\Big\}-\prod_{-N_t}^{N_t}E\exp\{ir_t^j\}\Big|\\
&\le\Big|E\exp\{ir_t^{-N_t}\}\exp\Big\{i\sum_{-N_t+1}^{N_t}r_t^j\Big\}-E\exp\{ir_t^{-N_t}\}E\exp\Big\{i\sum_{-N_t+1}^{N_t}r_t^j\Big\}\Big|\\
&\quad+\Big|E\exp\{ir_t^{-N_t}\}E\exp\Big\{i\sum_{-N_t+1}^{N_t}r_t^j\Big\}-\prod_{-N_t}^{N_t}E\exp\{ir_t^j\}\Big|\\
&\le C\varphi(\rho_t)+\Big|E\exp\Big\{i\sum_{-N_t+1}^{N_t}r_t^j\Big\}-\prod_{-N_t+1}^{N_t}E\exp\{ir_t^j\}\Big|.
\end{aligned}
$$

We then apply Lemma 8.2 (ii) recursively and get, according to Lemma 8.1,

$$
\Big|E\exp\Big\{i\sum\nolimits_t r_t^j\Big\}-\prod_{-N_t}^{N_t}E\exp\{ir_t^j\}\Big|\le CN_t\varphi(\rho_t)\to 0,\quad t\to\infty. \tag{8.15}
$$

It remains to check that

$$
\Big|\prod_{-N_t}^{N_t}E\exp\{ir_t^j\}-\exp\Big\{-\frac{1}{2}\sum\nolimits_t E|r_t^j|^2\Big\}\Big|\to 0,\ \ t\to\infty. \tag{8.16}
$$

According to the standard statement of the CLT (see, e.g. [24, Thm. 4.7]), it suffices to verify the Lindeberg condition: $\forall\varepsilon>0$,

$$
\frac{1}{\sigma_t}\sum\nolimits_t E_{\varepsilon\sqrt{\sigma_t}}|r_t^j|^2\to 0,\ \ t\to\infty. \tag{8.17}
$$

Here $\sigma_t\equiv\sum_t E|r_t^j|^2$, and $E_\delta f\equiv EX_\delta f$, where X_δ is the indicator of the event $|f|>\delta^2$. Note that (8.13) and (8.1) imply that

$$
\sigma_t\to\mathcal{Q}_\infty(\Psi,\Psi)\neq 0,\ \ t\to\infty.
$$

Hence it remains to verify that $\forall\varepsilon>0$,

$$
\sum\nolimits_t E_\varepsilon|r_t^j|^2\to 0,\ \ t\to\infty. \tag{8.18}
$$

We check Eq. (8.18) in Sect. 9. This will complete the proof of Proposition 3.3. □

9. The Lindeberg Condition

The proof of (8.18) can be reduced to the case when for some $\Lambda \geq 0$ we have, almost surely, that

$$|u_0(x)| + |v_0(x)| \leq \Lambda < \infty, \quad x \in \mathbb{R}^n. \tag{9.1}$$

Then the proof of (8.18) is reduced to the convergence

$$\sum\nolimits_t E|r_t^j|^4 \to 0, \quad t \to \infty \tag{9.2}$$

by using Chebyshev's inequality. The general case can be covered by standard cutoff arguments by taking into account that the bound (7.15) for $E|r_t^j|^2$ depends only on e_0 and φ. The last fact is obvious from (7.21) and (6.3) with $p = 1$ and $\gamma = 0$.

We deduce (9.2) from

Theorem 9.1. *Let the conditions of Theorem B hold and assume that (9.1) is fulfilled. Then for any $\Psi \in \mathcal{D}$ there exists a constant $C(\Psi)$ such that*

$$E|r_t^j|^4 \leq C(\Psi)\Lambda^4 d_t^2/t^2, \quad t > 1. \tag{9.3}$$

Proof. *Step 1.* Given four points $x_1, x_2, x_3, x_4 \in \mathbb{R}^n$, set:

$$M_0^{(4)}(x_1, ..., x_4) = E\left(Y_0(x_1) \otimes ... \otimes Y_0(x_4)\right).$$

Then, similarly to (7.16), Eqs. (9.1) and (7.8) imply by the Fubini Theorem that

$$E|r_t^j|^4 = \langle \chi_r^j(x_1^n) \dots \chi_r^j(x_4^n) M_0^{(4)}(x_1, \dots, x_4), \Phi(x_1, t) \otimes \cdots \otimes \Phi(x_4, t) \rangle. \tag{9.4}$$

Let us analyse the domain of the integration $(\mathbb{R}^n)^4$ in the RHS of (9.4). We partition $(\mathbb{R}^n)^4$ into three parts, W_2, W_3 and W_4:

$$(\mathbb{R}^n)^4 = \bigcup_{i=2}^4 W_i, \quad W_i = \{\bar{x} = (x_1, x_2, x_3, x_4) \in (\mathbb{R}^n)^4 : |x_1 - x_i| = \max_{p=2,3,4} |x_1 - x_p|\}. \tag{9.5}$$

Furthermore, given $\bar{x} = (x_1, x_2, x_3, x_4) \in W_i$, divide $\mathbb{R}^n$ into three parts S_j, $j = 1, 2, 3$: $\mathbb{R}^n = S_1 \cup S_2 \cup S_3$, by two hyperplanes orthogonal to the segment $[x_1, x_i]$ and partitioning it into three equal segments, where $x_1 \in S_1$ and $x_i \in S_3$. Denote by x_p, x_q the two remaining points with $p, q \neq 1, i$. Set: $\mathcal{A}_i = \{\bar{x} \in W_i : x_p \in S_1, x_q \in S_3\}$, $\mathcal{B}_i = \{\bar{x} \in W_i : x_p, x_q \notin S_1\}$ and $\mathcal{C}_i = \{\bar{x} \in W_i : x_p, x_q \notin S_3\}$, $i = 2, 3, 4$. Then $W_i = \mathcal{A}_i \cup \mathcal{B}_i \cup \mathcal{C}_i$. Define the function $\mathrm{m}_0^{(4)}(\bar{x})$, $\bar{x} \in (\mathbb{R}^n)^4$, in the following way:

$$\mathrm{m}_0^{(4)}(\bar{x})\Big|_{W_i} = \begin{cases} M_0^{(4)}(\bar{x}) - q_0(x_1 - x_p) \otimes q_0(x_i - x_q), & \bar{x} \in \mathcal{A}_i, \\ M_0^{(4)}(\bar{x}), & \bar{x} \in \mathcal{B}_i \cup \mathcal{C}_i. \end{cases} \tag{9.6}$$

This determines $\mathrm{m}_0^{(4)}(\bar{x})$ correctly for almost all quadruples $\bar{x}$. Note that

$$\begin{aligned} &\left\langle \chi_r^j(x_1^n) \dots \chi_r^j(x_4^n) q_0(x_1 - x_p) \otimes q_0(x_i - x_q), \Phi(x_1, t) \otimes \cdots \otimes \Phi(x_4, t) \right\rangle \\ &\quad = \left\langle \chi_r^j(x_1^n) \chi_r^j(x_p^n) q_0(x_1 - x_p), \Phi(x_1, t) \right. \\ &\qquad \left. \otimes \Phi(x_p, t) \right\rangle \left\langle \chi_r^j(x_i^n) \chi_r^j(x_q^n) q_0(x_i - x_q), \Phi(x_i, t) \otimes \Phi(x_q, t) \right\rangle. \end{aligned}$$

Each factor here is bounded by $C(\Psi)\, d_t/t$. Similarly to (7.15), this can be deduced from an expression of type (7.16) for the factors. Therefore, the proof of (9.3) reduces to the proof of the bound

$$\begin{aligned} I_t &:= |\langle \chi_r^j(x_1^n)\dots\chi_r^j(x_4^n)\mathrm{m}_0^{(4)}(x_1,\dots,x_4),\, \Phi(x_1,t)\otimes\dots\otimes\Phi(x_4,t)\rangle| \\ &\le C(\Psi)\Lambda^4 d_t^2/t^2, \qquad t>1. \end{aligned} \tag{9.7}$$

Step 2. Similarly to (7.21), Eq. (7.17) implies,

$$I_t \le C(\Psi)\, t^{-2n} \int\limits_{(B_t^{\overline{r}})^4} \chi_r^j(x_1^n)\dots\chi_r^j(x_4^n)|\mathrm{m}_0^{(4)}(x_1,\dots,x_4)|dx_1\, dx_2\, dx_3\, dx_4, \tag{9.8}$$

where $B_t^{\overline{r}}$ is the ball $\{x\in\mathbb{R}^n : |x|\le t+\overline{r}\}$. We estimate $\mathrm{m}_0^{(4)}$ using Lemma 8.2 (ii).

Lemma 9.2. *For each* $i=2,3,4$ *and almost all* $\overline{x}\in W_i$ *the following bound holds:*

$$|\mathrm{m}_0^{(4)}(x_1,\dots,x_4)| \le C\Lambda^4\varphi(|x_1-x_i|/3). \tag{9.9}$$

Proof. For $\bar{x}\in\mathcal{A}_i$ we apply Lemma 8.2 (ii) to $\mathbb{C}^2\otimes\mathbb{C}^2\equiv\mathbb{R}^4\otimes\mathbb{R}^4$-valued random variables $\xi=Y_0(x_1)\otimes Y_0(x_p)$ and $\eta=Y_0(x_i)\otimes Y_0(x_q)$. Then (9.1) implies the bound for almost all $\bar{x}\in\mathcal{A}_i$,

$$|\mathrm{m}_0^{(4)}(\bar{x})| \le C\Lambda^4\varphi(|x_1-x_i|/3). \tag{9.10}$$

For $\bar{x}\in\mathcal{B}_i$, we apply Lemma 8.2 (ii) to $\xi=Y_0(x_1)$ and $\eta=Y_0(x_p)\otimes Y_0(x_q)\otimes Y_0(x_i)$. Then **S0** implies a similar bound for almost all $\bar{x}\in\mathcal{B}_i$,

$$\begin{aligned} |\mathrm{m}_0^{(4)}(\bar{x})| &= \Big|M_0^{(4)}(\bar{x}) - EY_0(x_1)\otimes E\Big(Y_0(x_p)\otimes Y_0(x_q)\otimes Y_0(x_i)\Big)\Big| \\ &\le C\Lambda^4\varphi(|x_1-x_i|/3), \end{aligned} \tag{9.11}$$

and the same for almost all $\bar{x}\in\mathcal{C}_i$. □

Step 3. It remains to prove the following bounds for each $i=2,3,4$:

$$V_i(t) := \int\limits_{(B_t^{\overline{r}})^4} \chi_r^j(x_1^n)\dots\chi_r^j(x_4^n)X_i(\overline{x})\varphi(|x_1-x_i|/3)dx_1\, dx_2\, dx_3\, dx_4 \le Cd_t^2 t^{2n-2}, \tag{9.12}$$

where X_i is an indicator of the set W_i. In fact, this integral does not depend on i, hence set $i=2$ in the integrand:

$$V_i(t) \le C \int\limits_{(B_t^{\overline{r}})^2} \chi_r^j(x_1^n)\varphi(|x_1-x_2|/3)\left[\int\limits_{B_t^{\overline{r}}} \chi_r^j(x_3^n)\left(\int\limits_{B_t^{\overline{r}}} X_2(\overline{x})\, dx_4\right) dx_3\right] dx_1\, dx_2. \tag{9.13}$$

Now a key observation is that the inner integral in dx_4 is $\mathcal{O}(|x_1 - x_2|^n)$ as $X_2(\overline{x}) = 0$ for $|x_4 - x_1| > |x_1 - x_2|$. This implies

$$V_i(t) \le C\bar{r}^4 \int\limits_{B_t^{\bar{r}}} \chi_r^j(x_1^n) \left(\int\limits_{B_t^{\bar{r}}} \varphi(|x_1 - x_2|/3)|x_1 - x_2|^n \, dx_2 \right) dx_1 \int\limits_{B_t^{\bar{r}}} \chi_r^j(x_3^n) \, dx_3. \tag{9.14}$$

The inner integral in dx_2 is bounded as

$$\begin{aligned} \int\limits_{B_t^{\bar{r}}} \varphi(|x_1 - x_2|/3)|x_1 - x_2|^n \, dx_2 &\le C(n) \int_0^{2(t+\bar{r})} r^{2n-1}\varphi(r/3) \, dr \\ &\le C_1(n) \sup_{r\in[0,2(t+\bar{r})]} r^n \varphi^{1/2}(r/3) \int_0^{2(t+\bar{r})} r^{n-1}\varphi^{1/2}(r/3) \, dr, \end{aligned} \tag{9.15}$$

where the "sup" and the last integral are bounded by (8.4) and (2.10), respectively. Therefore, (9.12) follows from (9.14). This completes the proof of Theorem 9.1. □

Proof of convergence (9.2). As $d_t \le h \sim t/N_t$, bound (9.3) implies,

$$\sum\nolimits_t E|r_t^j|^4 \le \frac{C\Lambda^4 d_t^2}{t^2} N_t \le \frac{C_1 \Lambda^4}{N_t} \to 0, \quad N_t \to \infty. \qquad \square$$

10. The Scattering Theory for Infinite Energy Solutions

In this section we develop a version of the scattering theory to deduce Theorem A from Theorem B. The main step is to establish an asymptotics of type (1.20) for adjoint groups by using results of Vainberg [35].

Consider operators $U'(t)$, $U_0'(t)$ in the complex space $H = L^2(\mathbb{R}^n) \oplus H^1(\mathbb{R}^n)$ (see (2.12)). The energy conservation for the KGE implies the following corollary:

Corollary 10.1. *There exists a constant $C > 0$ such that $\forall\Psi \in H$:*

$$\|U_0'(t)\Psi\|_H \le C\|\Psi\|_H, \quad \|U'(t)\Psi\|_H \le C\|\Psi\|_H, \quad t \in \mathbb{R}. \tag{10.1}$$

Lemma 10.3 below develops earlier results [35, Thms. 3, 4, 5]. Consider a family of finite seminorms in H,

$$\|\Psi\|_{(R)}^2 = \int\limits_{|x|\le R} (|\Psi_0(x)|^2 + |\Psi_1(x)|^2 + |\nabla\Psi_1(x)|^2) \, dx, \qquad R > 0.$$

Denote by $H_{(R)}$ the subspace of functions from H with a support in the ball B_R.

Definition 10.2. *H_c denotes the space $\cup_{R>0} H_{(R)}$ endowed with the following convergence: a sequence Ψ_n converges to Ψ in H_c iff $\exists R > 0$ such that all $\Psi_n \in H_{(R)}$, and Ψ_n converge to Ψ in the norm $\|\cdot\|_{(R)}$.*

Below, we speak of continuity of maps from H_c in the sense of sequential continuity. Given $t \geq 0$, denote

$$\varepsilon(t) = \begin{cases} (t+1)^{-3/2}, & n \geq 3, \\ (t+1)^{-1} \ln^{-2}(t+2), & n = 2. \end{cases} \tag{10.2}$$

Lemma 10.3. *Let Assumptions* **E1–E3** *hold, and $n \geq 2$. Then for any $R, R_0 > 0$ there exists a constant $C = C(R, R_0)$ such that for $\Psi \in H_{(R)}$,*

$$\|U'(t)\Psi\|_{(R_0)} \leq C\varepsilon(t)\|\Psi\|_{(R)}, \ \ t \geq 0. \tag{10.3}$$

This lemma has been proved in [21] by using Conditions **E1–E3** and a method developed in [36]. For the proof, the contour of integration in the k-plane from [35] had to be curved logarithmically at infinity as in [36], but should not be chosen parallel to the real axis.

The main result of this section is Theorem 10.4 below. Given $t \geq 0$, set

$$\varepsilon_1(t) = \begin{cases} (t+1)^{-1/2}, \ n \geq 3, \\ \ln^{-1}(t+2), \ n = 2. \end{cases} \tag{10.4}$$

Theorem 10.4. *Let Assumptions* **E1–E3** *and* **S0–S3** *hold, and $n \geq 2$. Then there exist linear continuous operators $W, r(t): \ H_c \to H$ such that for $\Psi \in H_c$,*

$$U'(t)\Psi = U_0'(t)W\Psi + r(t)\Psi, \ \ t \geq 0, \tag{10.5}$$

and the following bounds hold $\forall R > 0$ and $\Psi \in H_{(R)}$:

$$\|r(t)\Psi\|_H \leq C(R)\varepsilon_1(t)\|\Psi\|_{(R)}, \ \ t \geq 0, \tag{10.6}$$

$$E|\langle Y_0, r(t)\Psi\rangle|^2 \leq C(R)\varepsilon_1^2(t)\|\Psi\|_{(R)}^2, \ \ t \geq 0. \tag{10.7}$$

Proof. We apply the standard Cook method: see, e.g., [29, Thm. XI.4]. Fix $\Psi \in H_{(R)}$ and define $W\Psi$, formally, as

$$W\Psi = \lim_{t\to\infty} U_0'(-t)U'(t)\Psi = \Psi + \int_0^\infty \frac{d}{dt}U_0'(-t)U'(t)\Psi \, dt.$$

We have to prove the convergence of the integral in norm in space H. First, observe that

$$\frac{d}{dt}U_0'(t)\Psi = \mathcal{A}_0'U_0'(t)\Psi, \qquad \frac{d}{dt}U'(t)\Psi = \mathcal{A}'U'(t)\Psi,$$

where $\mathcal{A}_0'$ and $\mathcal{A}'$ are the generators to groups $U_0'(t)$, $U'(t)$, respectively. Similarly to (7.5), we have

$$\mathcal{A}' = \begin{pmatrix} 0 & A \\ 1 & 0 \end{pmatrix}, \tag{10.8}$$

where $A = \sum_{j=1}^{n} (\partial_j - iA_j)^2 - m^2$. Therefore,

$$\frac{d}{dt} U'_0(-t)U'_1(t)\Psi = U'_0(-t)(\mathcal{A}' - \mathcal{A}'_0)U'(t)\Psi. \tag{10.9}$$

Now (10.8) and (7.5) imply

$$\mathcal{A}' - \mathcal{A}'_0 = \begin{pmatrix} 0 & L \\ 0 & 0 \end{pmatrix}.$$

Furthermore, **E2** implies that $L = \sum_{j=1}^{n} (\partial_j - iA_j)^2 - \Delta$ is a first order partial differential operator with the coefficients vanishing for $|x| \geq R_0$. Thus, (10.1) and (10.3) imply that

$$\begin{aligned} \|U'_0(-t)(\mathcal{A}' - \mathcal{A}'_0)U'(t)\Psi\|_H &\leq C\ \|(\mathcal{A}' - \mathcal{A}'_0)U'(t)\Psi\|_H \\ &= C\ \|\Big((\mathcal{A}' - \mathcal{A}'_0)U'(t)\Psi\Big)^0\|_{L^2(B_{R_0})} \\ &\leq C_1\ \|\Big(U'(t)\Psi\Big)^1\|_{H^1(B_{R_0})} \\ &\leq C(R)\varepsilon(t)\|\Psi\|_{(R)},\ t \geq 0. \end{aligned} \tag{10.10}$$

Hence (10.9) implies

$$\int_s^\infty \|\frac{d}{dt} U'_0(-t)U'(t)\Psi\|_H\, dt \leq C(R)\varepsilon_1(s)\|\Psi\|_{(R)}, \quad s \geq 0. \tag{10.11}$$

Therefore, (10.5) and (10.6) follow by (10.1). It remains to prove (10.7). First, similarly to (7.16),

$$E\langle Y_0, r(t)\Psi\rangle^2 = \langle q_0(x-y), r(t)\Psi(x) \otimes r(t)\Psi(y)\rangle. \tag{10.12}$$

Therefore, the Shur Lemma implies (similarly to (7.21))

$$E\langle Y_0, r(t)\Psi\rangle^2 \leq \|q_0\|_{L^1}\ \|r(t)\Psi\|_{L^2}\ \|r(t)\Psi\|_{L^2}, \tag{10.13}$$

where the norms $\|\cdot\|_{L^p}$ have an obvious meaning. Finally, (10.6) implies for $\Psi \in H_{(R)}$,

$$\|r(t)\Psi\|_{L^2} \leq C\|r(t)\Psi\|_H \leq C(R)\varepsilon_1(t)\|\Psi\|_{(R)}. \tag{10.14}$$

Therefore, (10.7) follows from (10.13) since $\|q_0\|_{L^1} < \infty$ by (6.1). □

11. Convergence to Equilibrium for Variable Coefficients

The assertion of Theorem A follows from two propositions below:

Proposition 11.1. *The family of the measures* $\{\mu_t, t \in \mathbb{R}\}$, *is weakly compact in* $\mathcal{H}^{-\varepsilon}$, $\forall \varepsilon > 0$.

Proposition 11.2. *For any* $\Psi \in \mathcal{D}$,

$$\hat{\mu}_t(\Psi) \equiv \int \exp(i\langle Y, \Psi\rangle)\ \mu_t(dY) \to \exp\{-\frac{1}{2}\mathcal{Q}_\infty(W\Psi, W\Psi\ )\}, \quad t \to \infty. \tag{11.1}$$

We deduce these propositions from Propositions 3.2 and 3.3, respectively, with the help of Theorem 10.4.

Proof of Proposition 11.1. Similarly to Proposition 3.2, Proposition 11.1 follows from the bounds

$$\sup_{t\geq 0} E\|U(t)Y_0\|_R < \infty, \quad R > 0. \tag{11.2}$$

For the proof, write the solution to (1.3) in the form

$$u(x,t) = v(x,t) + w(x,t). \tag{11.3}$$

Here $v(x,t)$ is the solution to (3.1), and $w(x,t)$ is the solution to the following Cauchy problem:

$$\left\{\begin{array}{l} \ddot{w}(x,t) = \sum_{k=1}^n (\partial_k - iA_k(x))^2 w(x,t) - m^2\, w(x,t) \\ \qquad - \sum_{k=1}^n 2iA_k(x)\partial_k v(x,t) - \sum_{k=1}^n (i\partial_k A_k(x) + A_k^2(x))v(x,t), \\ w|_{t=0} = 0, \quad \dot{w}|_{t=0} = 0, \quad x \in \mathbb{R}^n. \end{array}\right. \tag{11.4}$$

Then (11.3) implies

$$E\|U(t)Y_0\|_R \leq E\|U_0(t)Y_0\|_R + E\|(w(\cdot,t), \dot{w}(\cdot,t))\|_R. \tag{11.5}$$

By Proposition 3.1 we have

$$\sup_{t\geq 0} E\|U_0(t)Y_0\|_R < \infty. \tag{11.6}$$

It remains to estimate the second term in the right-hand side of (11.5). The Duhamel representation for the solution to (11.4) gives

$$(w, \dot{w}) = \int_0^t U(t-s)(0, \psi(\cdot,s))\, ds, \tag{11.7}$$

where $\psi(x,s) = -2i \sum\limits_{k=1}^n A_k(x)\partial_k v(x,s) - \sum\limits_{k=1}^n (i\partial_k A_k(x) + A_k^2(x))v(x,s)$. Assumption **E2** implies that supp $\psi(\cdot,s) \subset B_{R_0}$. Moreover,

$$\|(0, \psi(\cdot,s))\|_{R_0} \leq C\|v(\cdot,s)\|_{H^1(B_{R_0})} \leq C\|U_0(s)Y_0\|_{R_0}. \tag{11.8}$$

The decay estimates of type (10.3) hold for the group $U(t)$, as well as for $U'(t)$, as both groups correspond to the same equation by Lemma 7.1. Hence, we have from (11.8),

$$\begin{aligned} \|U(t-s)(0, \psi(\cdot,s))\|_R &\leq C(R)\varepsilon(t-s)\|(0, \psi(\cdot,s))\|_{R_0} \\ &\leq C_1(R)\varepsilon(t-s)\|U_0(s)Y_0\|_{R_0}, \end{aligned} \tag{11.9}$$

where $\varepsilon(\cdot)$ is defined in (10.2). Therefore, (11.7) and (11.6) imply

$$E\|(w(\cdot,t),\dot w(\cdot,t))\|_R \le C(R)\int_0^t \varepsilon(t-s)E\|U_0(s)Y_0\|_{R_0}\,ds$$
$$\le C_2(R)<\infty, \quad t\ge 0. \tag{11.10}$$

Then (11.6) and (11.5) imply (11.2). □

Proof of Proposition 11.2. Equations (10.5) and (10.7) imply by Cauchy–Schwartz,

$$|E\exp i\langle U(t)Y_0,\Psi\rangle - E\exp i\langle Y_0, U_0'(t)W\Psi\rangle| \le E|\langle Y_0, r(t)\Psi\rangle|$$
$$\le (E|\langle Y_0, r(t)\Psi\rangle|^2)^{1/2}\to 0,\ \ t\to\infty.$$

It remains to prove that

$$E\exp i\langle Y_0, U_0'(t)W\Psi\rangle \to \exp\Big\{-\frac{1}{2}\mathcal{Q}_\infty(W\Psi,W\Psi\,)\Big\},\ \ t\to\infty. \tag{11.11}$$

This does not follow directly from Proposition 3.3 since generally, $W\Psi\ \notin\ \mathcal{D}$. We approximate $W\Psi$ by functions from $\mathcal{D}$. $W\Psi\in H$, and $\mathcal{D}$ is dense in H. Hence, for any $\epsilon>0$ there exists $\Phi\in\mathcal{D}$ such that

$$\|W\Psi-\Phi\|_H\le\epsilon. \tag{11.12}$$

Therefore, we can derive (11.11) by the triangle inequality

$$\Big|E\exp i\langle Y_0, U_0'(t)W\Psi\rangle - \exp\Big\{-\frac{1}{2}\mathcal{Q}_\infty(W\Psi,W\Psi\,)\Big\}\Big|$$
$$\le \Big|E\exp i\langle Y_0, U_0'(t)W\Psi\rangle - E\exp i\langle Y_0, U_0'(t)\Phi\rangle\Big|$$
$$+E\Big|\exp i\langle U_0(t)Y_0,\Phi\rangle - \exp\Big\{-\frac{1}{2}\mathcal{Q}_\infty(\Phi,\Phi\,)\Big\}\Big|$$
$$+\Big|\exp\Big\{-\frac{1}{2}\mathcal{Q}_\infty(\Phi,\Phi\,)\Big\} - \exp\Big\{-\frac{1}{2}\mathcal{Q}_\infty(W\Psi,W\Psi\,)\Big\}\Big|. \tag{11.13}$$

Applying Cauchy–Schwartz, we get, similarly to (10.12)-(10.14), that

$$E|\langle Y_0, U_0'(t)(W\Psi-\Phi)\rangle| \le (E|\langle Y_0, U_0'(t)(W\Psi-\Phi)\rangle|^2)^{1/2}\le C\|U_0'(t)(W\Psi-\Phi)\|_H.$$

Hence, (10.1) and (11.12) imply

$$E|\langle Y_0, U_0'(t)(W\Psi-\Phi)\rangle|\le C\epsilon,\ \ t\ge 0. \tag{11.14}$$

Now we can estimate each term on the right-hand side of (11.13). The first term is $\mathcal{O}(\epsilon)$ uniformly in $t>0$ by (11.14). The second term converges to zero as $t\to\infty$ by Proposition 3.3 since $\Phi\in\mathcal{D}$. Finally, the third term is $\mathcal{O}(\epsilon)$ owing to (11.12) and the continuity of the quadratic form $\mathcal{Q}_\infty(\Psi,\Psi)$ in $L^2(\mathbb{R}^n)\otimes\mathbb{C}^2$. The continuity follows from the Shur Lemma since the integral kernels $q_\infty^{ij}(z)\in L^1(\mathbb{R}^n)\otimes M^2$ by Corollary 6.3. Now the convergence in (11.11) follows since $\epsilon>0$ is arbitrary. □

12. Appendix A. Fourier Transform Calculations

Consider the covariance functions of the solutions to the system (3.2). Let $F: w \mapsto \hat{w}$ denote the FT of a tempered distribution $w \in S'(\mathbb{R}^n)$ (see, e.g. [13]). We also use this notation for vector- and matrix-valued functions.

12.1. Dynamics in the FT space. In the FT representation, the system (3.2) becomes $\dot{\hat{Y}}(k,t) = \hat{\mathcal{A}}_0(k)\hat{Y}(k,t)$, hence

$$\hat{Y}(k,t) = \hat{\mathcal{G}}_t(k)\hat{Y}_0(k), \quad \hat{\mathcal{G}}_t(k) = \exp(\hat{\mathcal{A}}_0(k)t). \tag{12.1}$$

Here we denote

$$\hat{\mathcal{A}}_0(k) = \begin{pmatrix} 0 & 1 \\ -|k|^2 - m^2 & 0 \end{pmatrix}, \qquad \hat{\mathcal{G}}_t(k) = \begin{pmatrix} \cos \omega t & \dfrac{\sin \omega t}{\omega} \\ -\omega \sin \omega t & \cos \omega t \end{pmatrix}, \tag{12.2}$$

where $\omega = \omega(k) = \sqrt{|k|^2 + m^2}$.

12.2. Covariance matrices in the FT space.

Lemma 12.1. *In the sense of matrix-valued distributions,*

$$q_t(x-y) := E\Big(Y(x,t) \otimes Y(y,t)\Big) = F^{-1}_{k \to x-y} \hat{\mathcal{G}}_t(k)\hat{q}_0(k)\hat{\mathcal{G}}'_t(k), \quad t \in \mathbb{R}. \tag{12.3}$$

Proof. Translation invariance (1.8) implies

$$E\Big(Y_0(x) \otimes_C Y_0(y)\Big) = C_0^+(x-y), \quad E\Big(Y_0(x) \otimes_C \overline{Y_0(y)}\Big) = C_0^-(x-y), \tag{12.4}$$

where $\otimes_C$ stands for the tensor product of complex vectors. Therefore,

$$\begin{aligned} E\Big(\hat{Y}_0(k) \otimes_C \hat{Y}_0(k')\Big) &= F_{x \to k} F_{y \to k'}\, C_0^+(x-y) = (2\pi)^n \delta(k+k')\hat{C}_0^+(k), \\ E\Big(\hat{Y}_0(k) \otimes_C \overline{\hat{Y}_0(k')}\Big) &= F_{x \to k} F_{y \to -k'} C_0^-(x-y) = (2\pi)^n \delta(k-k')\hat{C}_0^-(k). \end{aligned} \tag{12.5}$$

Now (12.1) and (12.2) give in matrix notation that

$$\begin{aligned} E\Big(\hat{Y}(k,t) \otimes_C \hat{Y}(k',t)\Big) &= (2\pi)^n \delta(k+k')\hat{\mathcal{G}}_t(k)\hat{C}_0^+(k)\hat{\mathcal{G}}'_t(k), \\ E\Big(\hat{Y}(k,t) \otimes_C \overline{\hat{Y}(k',t)}\Big) &= (2\pi)^n \delta(k-k')\hat{\mathcal{G}}_t(k)\hat{C}_0^-(k)\hat{\mathcal{G}}'_t(k). \end{aligned} \tag{12.6}$$

Therefore, by the inverse FT formula we get

$$\begin{aligned} E\Big(Y(x,t) \otimes_C Y(y,t)\Big) &= F^{-1}_{k \to x-y} \hat{\mathcal{G}}_t(k)\hat{C}_0^+(k)\hat{\mathcal{G}}'_t(k), \\ E\Big(Y(x,t) \otimes_C \overline{Y(y,t)}\Big) &= F^{-1}_{k \to x-y} \hat{\mathcal{G}}_t(k)\hat{C}_0^-(k)\hat{\mathcal{G}}'_t(k). \end{aligned} \tag{12.7}$$

Then (12.3) follows by linearity. □

13. Appendix B. Measures in Sobolev's Spaces

Here we formally verify the bound (4.5) for $s, \alpha < -n/2$. Definition (4.2) implies for $u \in H^{s,\alpha}$,

$$\|u\|_{s,\alpha}^2 = \frac{1}{(2\pi)^{2n}} \int \langle x\rangle^{2\alpha} \Big[\int e^{-ix(k-k')} \langle k\rangle^s \langle k'\rangle^s \hat{u}(k)\overline{\hat{u}(k')}\, dk dk' \Big] dx. \tag{13.1}$$

Let $\mu(du)$ be a translation-invariant measure in $H^{s,\alpha}$ with a CF $Q(x, y) = q(x - y)$. Similarly to (12.5), (12.4), we get

$$\int \hat{u}(k)\overline{\hat{u}(k')}\mu(du) = (2\pi)^n \delta(k - k')\, \mathrm{tr}\,\hat{q}(k). \tag{13.2}$$

Then, integrating (13.1) with respect to the measure $\mu(du)$, we get the formula

$$\int \|u\|_{s,\alpha}^2 \mu(du) = \frac{1}{(2\pi)^n} \int \langle x\rangle^{2\alpha} dx \int \langle k\rangle^{2s}\, \mathrm{tr}\,\hat{q}(k)\, dk. \tag{13.3}$$

Applying it to $\hat{q}(k) = T$ with $\alpha, s < -n/2$ and to $\hat{q}(k) = T(k^2 + m^2)^{-1}$ with $1 + s$ instead of s, we get (4.5).

Acknowledgements. The authors thank V. I. Arnold, A. Bensoussan, I. A. Ibragimov, H. P. McKean, J. Lebowitz, A. I. Shnirelman, H. Spohn, B. R. Vainberg and M. I. Vishik for fruitful discussions and remarks.

References

1. Billingsley, P.: *Convergence of Probability Measures*. New York, London, Sydney, Toronto: John Wiley, 1968
2. Boldrighini, C., Dobrushin, R.L.Sukhov, Yu.M.: Time asymptotics for some degenerate models of evolution of systems with an infinite number of particles. Technical Report, University of Camerino, 1980
3. Boldrighini, C., Pellegrinotti, A., Triolo, L.: Convergence to stationary states for infinite harmonic systems, J. Stat. Phys. **30**, 123–155 (1983)
4. Botvich, D.D., Malyshev, V.A.: Unitary equivalence of temperature dynamics for ideal and locally perturbed fermi-gas. Commun. Math. Phys. **91**, no. 4, 301–312 (1983)
5. Bournaveas, N.: Local existence for the Maxwell–Dirac equations in three space dimensions. Comm. Partial Diff. Equs. **21**, no. 5–6, 693–720 (1996)
6. Bulinskii, A.V., Molchanov, S.A.: Asymptotic Gaussian property of the solution of the Burgers equation with random initial data. Theory Probab. Appl. **36**, no. 2, 217–236 (1991)
7. Cornfeld, I.P., Fomin, S.V., Sinai, Ya.G.: *Ergodic Theory*. New York–Berlin: Springer, 1982
8. Dobrushin, R.L., Pellegrinotti, A., Suhov, Yu.M.: One-dimensional harmonic lattice caricature of hydrodynamics: A higher correction. J. Stat. Phys. **61**, no. 1/2, 387–402 (1990)
9. Dobrushin, R.L., Sinai, Ya.G., Sukhov, Yu.M.: Dynamical systems of statistical mechanics. In: *Dynamical Systems, Ergodic Theory and Applications, Encyclopaedia of Mathematical Sciences, V.* **100**. Berlin: Springer, 2000, pp. 384–431
10. Dobrushin, R.L., Suhov, Yu.M.: On the problem of the mathematical foundation of the Gibbs postulate in classical statistical mechanics. In: *Mathematical Problems in Theoretical Physics, Lecture Notes in Physics, V.* **80**. Berlin: Springer-Verlag 1978, pp. 325–340
11. Dudnikova, T.V., Komech, A.I., Ratanov, N.E., Suhov, Yu.M.: On convergence to equilibrium distribution. II. Wave equations with mixing. Submitted to J. Stat. Phys.
12. Dudnikova, T.V. Komech, A.I., Spohn, H.: On convergence to statistic equilibrium in two-temperature problem for wave equation with mixing. Preprint Max-Planck Institute for Mathematics in the Sciences, N. **26**. Leipzig, 2000 (http://www.mis.mpg.de)
13. Egorov, Yu.V., Komech, A.I., Shubin, M.A.: *Elements of the Modern Theory of Partial Differential Equations*. Berlin: Springer, 1999
14. Fedoryuk, M.V.: The stationary phase method and pseudodifferential operators. Russ. Math. Surveys **26**, no. 1, 65–115 (1971)

15. Gikhman, I.I., Skorokhod, A.V.: *The Theory of Stochastic Processes, Vol.* **I**. Berlin: Springer, 1974
16. Hörmander, L.: *The Analysis of Linear Partial Differential Operators III: Pseudo-Differential Operators*. Berlin–Heidelberg–New York: Springer-Verlag, 1985
17. Ibragimov, I.A., Linnik, Yu.V.: *Independent and Stationary Sequences of Random Variables*. Groningen: Wolters-Noordhoff, 1971
18. Jaksic, V., Pillet, C.-A.: Ergodic properties of classical dissipative systems. I. Acta Math. **181**, no. 2, 245–282 (1998)
19. Komech, A.I.: Stabilisation of statistics in wave and Klein–Gordon equations with mixing. Scattering theory for solutions of infinite energy. Rend. Sem. Mat. Fis. Milano **65**, 9–22 (1995)
20. Kopylova, E.A.: Stabilization of statistical solutions of the Klein–Gordon equation. Mosc. Univ. Math. Bull. **41**, no. 2, 72–75 (1986)
21. Kopylova, E.A.: Stabilisation of Statistical Solutions of Klein–Gordon Equations. PhD Thesis, Moscow State University, 1986
22. Mikhailov, V.P.: Partial Differential Equations. Moscow: Mir, 1978
23. Morawetz, C.S., Strauss, W.A.: Decay and scattering of solutions of a nonlinear relativistic wave equation. Comm. Pure Appl. Math. **25**, 1–31 (1972)
24. Petrov, V.V.: *Limit Theorems of Probability Theory*. Oxford: Clarendon Press, 1995
25. Planck, M.: *The Theory of Heat Radiation*. New York: Dover Publications, 1959
26. Ratanov, N.E.: Stabilisation of statistic solutions of second order hyperbolic equations. Russian Mathematical Surveys **39**, no. 1, 179–180 (1984)
27. Ratanov, N.E., Shuhov, A.G., Suhov, Yu.M.: Stabilisation of the statistical solution of the parabolic equation, Acta Appl. Math. **22**, no. 1, 103–115 (1991)
28. Reed, M., Simon, B.: *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness*. New York: Academic Press, 1975
29. Reed, M., Simon, B.: *Methods of Modern Mathematical Physics III: Scattering Theory*. New York: Academic Press, 1979
30. Rosenblatt, M.A.: A central limit theorem and a strong mixing condition. Proc. Nat. Acad. Sci. U.S.A. **42**, no. 1, 43–47 (1956)
31. Seitz, F.: *The Modern Theory of Solids.* New York: McGraw-Hill, 1940
32. Shuhov, A.G., Suhov, Yu.M.: Ergodic properties of groups of the Bogoliubov transformations of CAR C^*-algebras. Ann. Phys. **175**, 231–266 (1987)
33. Spohn, H., Lebowitz, J.L.: Stationary non equilibrium states of infinite harmonic systems. Commun. Math. Phys. **54**, 97–120 (1977)
34. Sommerfeld, A.: *Thermodynamics and Statistical Mechanics.* New York: Academic Press, 1956
35. Vainberg, B.R.: Behaviour for large time of solutions of the Klein–Gordon equation. Trans. Moscow Math. Soc. **30**, 139–158 (1976)
36. Vainberg, B.R.. *Asymptotic Methods in Equations of Mathematical Physics*. New York–London–Paris: Gordon and Breach, 1989
37. Vishik, M.I., Fursikov, A.V.: *Mathematical Problems of Statistical Hydromechanics*. Dordrecht; Kluwer Academic Publishers, 1988

Communicated by H. Spohn