ISSN 0001-4346, Mathematical Notes, 2012, Vol. 92, No. 2, pp. 151–165. © Pleiades Publishing, Ltd., 2012.
Original Russian Text © J. Brüning, V. V. Grushin, S. Yu. Dobrokhotov, 2012, published in Matematicheskie Zametki, 2012, Vol. 92, No. 2, pp. 163–180.

Averaging of Linear Operators, Adiabatic Approximation, and Pseudodifferential Operators

J. Brüning[1*], V. V. Grushin[2], and S. Yu. Dobrokhotov[3***]**

[1]*Humboldt University, Germany*

[2]*Moscow State Institute of Electronics and Mathematics, Moscow Institute of Physics and Technology*

[3]*Institute for Problems in Mechanics, Russian Academy of Sciences, Moscow Institute of Physics and Technology*

Received December 12, 2011

Abstract—An example of Schrödinger and Klein–Gordon equations with fast oscillating coefficients is used to show that they can be averaged by an adiabatic approximation based on V. P. Maslov's operator method.

DOI: 10.1134/S0001434612070188

Keywords: *Klein–Gordon equation, Schrödinger equation, adiabatic approximation, asymptotic solution, pseudodifferential operator, adiabatic principle, perturbation theory.*

1. INTRODUCTION

In linear problems for partial differential equations, averaging methods work in situations where the equation's coefficients are fast oscillating functions. Averaging methods have been investigated in numerous publications, where both very serious theoretical mathematical problems and their applications are considered; here we mention only the monographs [1]–[4]. As a rule, these methods are used to construct asymptotic solutions of the original equation in the case where the leading term is already a sufficiently smooth (but not fast oscillating) function. On the other hand, in many physical problems it is of interest to consider situations in which the leading term of the asymptotic solution is also a fast varying function. In this case, the initial problem contains several different scales, and it is reasonable to use the adiabatic approximation in the problem. We illustrate this approach with an example of two equations in the m-dimensional Euclidean space $\mathbb{R}^m_x$ with coordinates $x = (x_1, \dots, x_m)$, i.e., we consider Schrödinger- and Klein–Gordon-type equations (in particular, wave-type equations) with fast varying velocity and potential which have the form

$$C = C\biggl(\frac{\Theta(x)}{\mu}, x\biggr), \qquad V = V\biggl(\frac{\Theta(x)}{\mu}, x\biggr). \tag{1.1}$$

Here $\Theta(x) = (\Theta, \dots, \Theta_n)$, $C(y,x)$, and $V(y,x)$ are smooth real functions, $C(y,x)$ and $V(y,x)$ are 2π-periodic in each of the variables $y_1, \dots, y_n$, $n \le m$, and it is assumed that the phases Θ_j are locally noncollinear, i.e., the rank of the matrix Θ_x composed of the rows $((\Theta_1)_{x_k}, \dots, (\Theta_n)_{x_k})$, $k = 1, \dots, m$, is equal to n for all x. In many physically interesting situations, m takes the values 1, 2, and 3, and the number of "phases" n can range from 1 to m. The fact that the phases Θ_j nonlinearly depend on x means that there is a weak variation in the frequencies of spatial oscillations of the velocity and

*E-mail: bruening@mathematik.hu-berlin.de
**E-mail: vvgrushin@mail.ru
***E-mail: dobr@ipmnet.ru

the potential. In $\mathbb{R}^m_x$, we consider the operator $-\langle\nabla, C^2(\Theta)\nabla\rangle + V$ and the following two evolution equations corresponding to this operator:

$$\begin{aligned}
&\text{(a)}\quad i\frac{\partial\psi}{\partial t} = \biggl(-\biggl\langle\nabla, C^2\biggl(\frac{\Theta(x)}{\mu}, x\biggr)\nabla\biggr\rangle + V\biggr)\psi,\\
&\text{(b)}\quad \frac{\partial^2\psi}{\partial t^2} = -\biggl(-\biggl\langle\nabla, C^2\biggl(\frac{\theta(x)}{\mu}, x\biggr)\nabla\biggr\rangle + V\biggr)\psi.
\end{aligned}\tag{1.2}$$

Of course, these equations must be supplemented with initial conditions. Our goal is to construct some asymptotic solutions of these equations, more precisely, to derive some new “reduced averaged” equations with smooth coefficients in terms of which the solutions of the original equations can be expressed. To this end, following the same reasoning as in [5], we first (in Sec. 2) reduce Eqs. (1.2) to the form of equations with operator-valued symbol (see [6]) and then (in Sec. 3) apply the version (developed in [7]–[10]) of adiabatic approximation in operator form, which is based on pseudodifferential operators (functions of noncommuting operators) and on operator methods [11]. Thus, the method proposed for constructing asymptotic solutions of Eqs. (1.2) is conditionally divided into the following two parts: the first part is the reduction to “averaged equations” determined by their symbols, the so-called effective Hamiltonians; the second part is the construction of asymptotic, both slowly and fast varying, solutions of these “averaged equations”. The assumption about the behavior of solutions of the “averaged equations” is essential; they allow one to use only the expansions of effective Hamiltonians in (quasi)momenta (see Sec. 4) and hence to obtain a rather efficient description of the reduced equations (see Sec. 5) (including equations with the so-called dispersive additional terms). In the present paper, we mainly deal with the first part of our approach, which, in a sense, is of “operator character” or even of “algebraic character.” In Sec. 6, we additionally present some estimates justifying our methods including the derivation of the averaged equations and the possibility of their application to different solutions.

2. EQUATIONS WITH FAST OSCILLATING COEFFICIENTS AND AN OPERATOR-VALUED SYMBOL

We seek [5] some solutions of Eqs. (1.2) in the form of functions periodically depending on Θ_j and also on x and t:

$$\psi = \Psi\biggl(\frac{\Theta}{\mu}, x, t\biggr),\tag{2.1}$$

where $\Psi(y, x, t)$ is a smooth function that is 2π-periodic in each of the variable $y_1, \dots, y_n$. We introduce the following notation: by $\langle\,\cdot\,,\,\cdot\,\rangle$ we denote the inner product in $\mathbb{R}^m$, by $\nabla_x = \partial/\partial x$, the (column) vector expressing the gradient operator in $\mathbb{R}^m_x$, by $\nabla_y = \partial/\partial y$, the (column) vector expressing the gradient operator in $\mathbb{R}^n_y$, and by ∇^θ_y, the skew gradient vector-operator, $\nabla^\theta_y = \Theta_x(x)\nabla_y$. Substituting the functions (2.1) into Eqs. (1.2), we see that the functions ψ are their solutions if $\Psi(y, x, t)$ satisfies the corresponding equations

$$\begin{aligned}
&\text{(a)}\quad i\mu^2\Psi_t = \widehat{\mathcal{H}}\Psi,\\
&\text{(b)}\quad \mu^2\Psi_{tt} = -\widehat{\mathcal{H}}\Psi,
\end{aligned}\tag{2.2}$$

$$\widehat{\mathcal{H}} = \langle(-i\mu\nabla_x - i\nabla^\Theta_y), C^2(y, x)(-i\mu\nabla_x - i\nabla^\Theta_y)\rangle + \mu^2 V(y, x).\tag{2.3}$$

We denote $\Delta^\theta_y = \langle\nabla^\theta_y, C^2(y, x)\nabla^\theta_y\rangle$ and $D = \langle p, \nabla^\theta_y\rangle$. Then the operator-valued symbol of the operators in the right-hand side in the variables x can be rewritten as $\mathcal{H} = \mathcal{H}_0 + \mu\mathcal{H}_1 + \mu^2\mathcal{H}_2$, where

$$\mathcal{H}_0 = \langle(p - i\nabla^\Theta_y), C^2(y, x)(p - i\nabla^\Theta_y)\rangle = -\Delta^\Theta_y - i(DC^2(y, x) + C^2(y, x)D) + C^2(y, x)p^2,\tag{2.4}$$

$$\mathcal{H}_1 = -\langle\nabla_x, C^2(y, x)\nabla^\Theta_y\rangle - i\langle\nabla_x, C^2(y, x)p\rangle, \qquad \mathcal{H}_2 = V(y, x).\tag{2.5}$$

Equations (2.2) contain n more variables than the original equations. The new variables are introduced to ensure the regularization of the coefficients, namely, in (2.2), they depend on μ as $\mu \to 0$ already in a smooth way. The goal of the averaging (adiabatic approximation) methods is to eliminate these variables, i.e., to decrease the dimension of Eqs. (2.2). Although in such problems there is no actual decrease in the number of variables compared with the original equations, we nevertheless can assume that the fast phases are to a certain extent independent of the variables x, and hence their elimination in the further calculations can also be treated as a reduction of dimension.

3. GENERALIZED ADIABATIC PRINCIPLE

The approach proposed in [7]–[10], [12] (generalized adiabatic principle) can be applied to Eqs. (2.2). In the general case, the equations under study can be written in the form (2.2), where the Hamiltonian $\widehat{\mathcal{H}}$ is a pseudodifferential operator with an operator-valued symbol, which, following the notation introduced in [11], can be written as

$$\widehat{\mathcal{H}} = \mathcal{H}\biggl(-i\mu\overset{1}{\frac{\partial}{\partial x}}, \overset{2}{x}, y, -i\frac{\partial}{\partial y}, \mu\biggr).$$

The digits over the operators denote the order of action of these operators (see [11]). We assume that the symbol of the operator $\widehat{\mathcal{H}}$ satisfies the asymptotic expansion

$$\mathcal{H}\biggl(x,p,y,-i\frac{\partial}{\partial y},\mu\biggr) = \mathcal{H}_0\biggl(x,p,y,-i\frac{\partial}{\partial y}\biggr) + \mu\mathcal{H}_1\biggl(x,p,y,-i\frac{\partial}{\partial y}\biggr) + \cdots . \tag{3.1}$$

In our case, the operator $\mathcal{H}(x,p,y,-i(\partial/\partial y),\mu)$ is differential and consists only of three terms.

We seek some solutions Ψ of Eq. (2.2) as the action of a certain (so far unknown) pseudodifferential operator on a new (also still unknown) function

$$\Psi(x,y,t,\mu) = \widehat{\chi}w \equiv \chi\biggl(\overset{2}{x}, -i\mu\overset{1}{\frac{\partial}{\partial x}}, y, \mu\biggr)w(x,t,\mu). \tag{3.2}$$

Here $\widehat{\chi}$ is an "intertwining" pseudodifferential operator admitting the expansion

$$\chi(x,p,y,\mu) = \chi_0(x,p,y) + \mu\chi_1(x,p,y) + \cdots . \tag{3.3}$$

As to the function w, we assume that it satisfies the "effective" (reduced) equation

$$i\mu w_t = L\biggl(\overset{2}{x}, -i\mu\overset{1}{\frac{\partial}{\partial x}}, \mu\biggr)w, \tag{3.4}$$

or

$$\mu^2 w_{tt} = -L\biggl(\overset{2}{x}, -i\mu\overset{1}{\frac{\partial}{\partial x}}, \mu\biggr)w \tag{3.5}$$

given by the operator[1] $\widehat{L}$ whose symbol L admits a regular expansion in μ as follows:

$$L(x,p,\mu) = L_0(x,p) + \mu L_1(x,p) + \cdots . \tag{3.6}$$

The function $H_{\text{eff}}(p,x) = L_0(x,p)$ is called the (classical) *effective Hamiltonian*. We also note that an obvious change allows us to reduce the operator $\mathcal{H}_0$ defined on periodic function to the form of the operator Δ_y^{Θ} defined on Bloch functions, and then the variables (parameters) p can be linearly expressed in terms of the quasimomenta of these functions.

Lemma 1. *For the function Ψ of the form* (3.2) *to satisfy Eq.* (2.2 a) (*Eq.* (2.2 b))*, it suffices that the function w satisfy Eq.* (3.4) (*Eq.* (3.5)) *and the symbols $\chi(p,x,y,\mu)$ and $L(p,x,\mu)$ of the operators $\widehat{\chi}$ and $\widehat{L}$ satisfy the equation*

$$\chi\biggl(\overset{2}{x}, p - i\mu\overset{1}{\frac{\partial}{\partial x}}, y, \mu\biggr)L(x,p,\mu) = \mathcal{H}\biggl(\overset{2}{x}, p - i\mu\overset{1}{\frac{\partial}{\partial x}}, y, -i\frac{\partial}{\partial y}, \mu\biggr)\chi(x,p,y,\mu). \tag{3.7}$$

[1] In the physical literature (in solid-state physics), the construction of the operator $L(\overset{2}{x}, -i\mu(\overset{1}{\partial}/\partial x), \mu)$ with respect to the function $L(p,x,\mu)$ is called the *Peierls subsitution*, see [13] and more precise explanations in [8]–[10].

Proof. We substitute the function Ψ from (3.2) into Eq. (2.2) and obtain the relation $i\mu^2\widehat{\chi}w_t = \widehat{\mathcal{H}}\widehat{\chi}w$, which we rewrite as

$$(\widehat{\chi}\widehat{L} - \widehat{\mathcal{H}}\widehat{\chi})w = 0$$

using (3.4). A sufficient condition for the last relation to be satisfied is the following operator relation:

$$\widehat{\chi}\widehat{L} - \widehat{\mathcal{H}}\widehat{\chi} = 0.$$

In this relation, we pass from operators to their symbols and obtain Eq. (3.7). The derivation of (3.7) from (3.2) (2.2 b), and (3.5) is quite similar. □

Remark 1. We see that Eq. (3.7) is the same in both cases (2.2 a,b). Of course, this reduction can also be used for solutions harmonic in time, i.e., for solutions of the form $\Psi = e^{-iEt/\mu^2}\Psi'(x, y, \mu)$. It is clear that the assertion of the lemma remains valid if Eqs. (2.2) and (3.4) are replaced by the corresponding stationary equations (the prime is omitted here)

$$\widehat{\mathcal{H}}\Psi = E\Psi, \qquad L\biggl(\overset{2}{x}, -i\mu\overset{1}{\frac{\partial}{\partial x}}, \mu\biggr)\psi = E\psi.$$

Thus, the problem of reduction is reduced to the construction of symbols of the operators $\widehat{\chi}$ and $\widehat{L}$. The fact that (3.7) contains two unknown objects $\widehat{\chi}$ and $\widehat{L}$ is similar to the situation arising in the problem of determining the eigenvectors and eigenvalues. Just as in the case of determining the eigenvectors, Eq. (3.7) generally has infinitely many $\widehat{\chi}^k$ and $\widehat{L}^k$, which are often called "modes" or "terms". In this case, $\widehat{\chi}^k$ and $\widehat{L}^k$ are also determined nonuniquely for each k. Let $\mathbb{T}$ be the torus formed by $y_j \bmod 2\pi$, $j = 1, \dots, n$. In what follows, it is convenient to introduce the space $L_2(\mathbb{T})$ with respect to the variables y with a "normed" inner product by assuming that

$$(g, f)_{L_2(\mathbb{T})} = \frac{1}{(2\pi)^n}\int_{\mathbb{T}} \overline{g}(y)f(y)\,dy^n$$

for any functions $g(y)$ and $f(y)$, where the bar denotes complex conjugation. A certain ambiguity in the determination of $\widehat{\chi}^k$ and $\widehat{L}^k$ can be removed if we assume that the norms of the functions $\Psi(x, y, t)$ and w coincide in the corresponding spaces $L_2(\mathbb{T}\times\mathbb{R}^m)$ and $L_2(\mathbb{R}^m)$:

$$(\Psi, \Psi)_{L_2(\mathbb{T}\times\mathbb{R}^m)} \equiv (\widehat{\chi}w, \widehat{\chi}w)_{L_2(\mathbb{T}\times\mathbb{R}^m)} \equiv (\widehat{\chi}^*\widehat{\chi}w, w)_{L^2(\mathbb{T}\times\mathbb{R}^m)} = (w, w)_{L_2(\mathbb{R}^m)}.$$

Then, taking the relation

$$\widehat{\chi}^* = \chi\biggl(-i\mu\overset{2}{\frac{\partial}{\partial x}}, \overset{1}{x}, y, \mu\biggr)$$

into account, we see that the symbol of the operator $\widehat{\chi}^*\widehat{\chi}$ is equal to

$$\chi^*\biggl(p - i\mu\overset{2}{\frac{\partial}{\partial x}}, \overset{1}{x}, y, \mu\biggr)\chi(p, x, y, \mu),$$

where $\chi^*(p, x, y, \mu)$ is the complex conjugate of $\chi(p, x, y, \mu)$. If the integral of this function with respect to the variable y is equal to 1, then the norms of the functions Ψ and w coincide. This "normalization" condition can be written as

$$\frac{1}{(2\pi)^n}\int_{\mathbb{T}} \chi^*\biggl(p - i\mu\overset{2}{\frac{\partial}{\partial x}}, \overset{1}{x}, y, \mu\biggr)\chi(p, x, y, \mu)\,dy^n = 1. \tag{3.8}$$

Condition (3.8) does not completely remove the ambiguity in the determination of χ and L, because the operator $\widehat{\chi}$ can always be replaced by the product

$$\widehat{\chi}U\biggl(-i\mu\overset{1}{\frac{\partial}{\partial x}}, \overset{2}{x}, \mu\biggr),$$

where $U(-i\mu(\overset{1}{\partial/\partial x}), \overset{2}{x}, \mu)$ is taken as a unitary (pseudodifferential) operator. Such a variation obviously leads to a unitary transformation of the operator $\widehat{L}$, i.e., to the replacement of the corresponding operator $\widehat{L}$ by the pseudodifferential operator

$$\widehat{L}' = U\biggl(-i\mu\frac{\overset{1}{\partial}}{\partial x}, \overset{2}{x}, \mu\biggr)\widehat{L}\biggl[U\biggl(-i\mu\frac{\overset{1}{\partial}}{\partial x}, \overset{2}{x}, \mu\biggr)\biggr]^{-1}.$$

As we shall see later, the leading (with respect to the parameter μ) parts of the symbols $\widehat{L}'$ and $\widehat{L}$ coincide in this case. This ambiguity in the choice of χ does not affect the final result, i.e., the function Ψ; nevertheless, a good choice of χ can significantly simplify the calculations in specific problems.

If $\widehat{\chi}$ and $\widehat{L}$ are found, then Eq. (2.2 a) (or (2.2 b) is reduced to a simpler (reduced) equation (3.4) (or (3.5)) for the function w. After a solution of the reduced equation (3.4) (or (3.5)) is determined, the corresponding solutions Ψ of the original equation can be reconstructed by using the action of the intertwining operator by formula (3.2).

Different situations, questions, and difficulties typical of the equations with operator-valued symbol arise in the study of Eqs. (2.4) and (2.5). We make several important remarks related to this fact.

Remark 2. Here, as a rule, the "resonance" situation, which is known as the "change of multiplicity of characteristics" in the mathematical literature and as the "intersection of terms" in the physics literature, may arise. In this case, the functions L^k and $L^{k'}$ with different k and k' coincide for some values of p and x; then infinitely many different situations are possible, and only some of them have been studied (see, e.g., [14]). On the other hand, the occurrence of such situations is to some extent "local with respect to (x,p)", and they can be avoided under certain conditions imposed on the solutions of the reduced equations (3.4) and (3.5); the problem is to investigate the range of the class of solutions that have no effects of the "change of multiplicity of characteristics" (on reasonable time intervals). In the present paper, we in fact assume that the solutions of the original equation have a structure (belong to such a class) such that these effects practically do not affect their asymptotics on appropriate time intervals.

Remark 3. The problem of construction of solutions to Eq. (3.7) also arises. This problem can be solved exactly only in very rare cases, and regular perturbation theory in the parameter μ can therefore be used to determine the coefficients of the expansion of the functions χ and L. In this case, (3.3) and (3.6) can naturally be treated as *asymptotic* series in the parameter μ. Moreover, the practical calculation of the first coefficients is already a nontrivial problem, and it is necessary to consider only a minimal reasonable number of terms in expansions (3.3) and (3.6). This number can be determined from the following "asymptotic" considerations. In the case of nonstationary problems, for example, of specific Cauchy problems, the above procedure must allow us to calculate the leading term of the asymptotics of the solution Ψ. In the case of stationary (spectral) problems, this procedure must allow us to construct the asymptotics of a part of the spectrum (a "spectral series"), and the error of the eigenvalue approximation by such an asymptotics must be much less than the distance between the nearest eigenvalues.

Remark 4. The digits over the operators were introduced because of the fact that the operators $\widehat{p} = -i\mu(\partial/\partial x)$ and x do not commute, and the digits determine a method for ordering them in the construction of functions of the operators. The construction of functions of these operators is not unique and depends on their ordering. The following two ordering methods (or "quantizations") are used much more frequently than the others. In the first case, we deal with the Feynman–Maslov ordering

$$\widehat{\chi} = \chi(\overset{1}{\widehat{p}}, \overset{2}{x}, \mu), \qquad \widehat{L} = L(\overset{1}{\widehat{p}}, \overset{2}{x}, \mu),$$

which we already used and, in the second case, with the Weyl ordering

$$\widehat{\chi} = \chi_W\biggl(\frac{\overset{1}{\widehat{p}} + \overset{3}{\widehat{p}}}{2}, \overset{2}{x}, y, \mu\biggr), \qquad \widehat{L} = L_W\biggl(\frac{\overset{1}{\widehat{p}} + \overset{3}{\widehat{p}}}{2}, \overset{2}{x}, \mu\biggr)$$

with the Weyl symbols $\chi_W(p,x,t,\mu)$ and $L_W((p,x,\mu)$. We stress that the operators are the same, but their symbols are generally different because of different orderings. From the "theoretical viewpoint", Weyl quantization is more convenient, because it automatically leads to at least symmetric operators $\widehat{L}$, provided that $L^* = L$. But the practical experience of solution of many problems shows that the Feynman–Maslov ordering is significantly more convenient for deriving explicit formulas, and the pragmatic method consists, first, in finding the symbols corresponding to the Feynman–Maslov ordering and then passing to the Weyl symbols (i.e., recalculating them). Such formulas are well known (see [11], and [8], [10]), in particular,

$$L_W(p,x,\mu) = L_0(p,x) + \mu\biggl(L_1(p,x) + \frac{i}{2}\Bigl\langle\frac{\partial}{\partial p},\frac{\partial}{\partial x}\Bigr\rangle L_0\biggr) + O(\mu^2).$$

Since the initial problems are given by self-adjoint operators, the operator $\widehat{L}$ (in the case of a correct choice of $\widehat{\chi}$) is also self-adjoint. But then $L_W(p,x,\mu)$ is real-valued, and hence L_0 and the coefficient at μ in the right-hand side of the last relation are also real-valued. The calculations must be performed with this consideration in mind.

Remark 5. Finally, we note that the class of symbols of pseudodifferential equations with a parameter, which are considered here and generally in adiabatic problems, differs from the class of symbols considered, for example, in [15], [16]. In particular, we do not require that the terms of expansions (3.3) and (3.6) in the variables p decrease with the number k. Moreover, the functions of operators considered above can of course be written by using the Fourier transform, and precisely this gives their rigorous definition. In our calculations, we do not need such definitions; if necessary, they can be found in [11], [17], and also in [9], [10].

4. CALCULATION OF THE SYMBOL OF THE REDUCED EQUATION BY USING PERTURBATION THEORY IN THE PARAMETER μ

Let us use the well-known formula of composition of pseudodifferential operators. Applying the Taylor formula, at least formally, we can rewrite the left-hand side of (3.7) as

$$\chi(x,p,y,\mu)L(x,p,\mu) + \sum_{|\nu|=1}^{\infty}\frac{1}{\nu!}(-i)^{|\nu|}\mu^{|\nu|}\frac{\partial^{|\nu|}\chi}{\partial p^{\nu}}(x,p,y,\mu)\frac{\partial^{|\nu|}}{\partial x^{\nu}}L(x,p,\mu). \tag{4.1}$$

Here $\nu = (\nu_1,\dots,\nu_m)$ is a multiindex, $|\nu| = \nu_1 + \dots + \nu_m$, and $\nu! = \nu_1!\cdots\nu_m!$. A similar relation also holds for the right-hand side of (3.7), but it has finitely many terms. Substituting these expressions into (3.7), taking expansions (3.1), (3.3), and (3.6) into account, and equating the coefficients at different powers of μ to zero, we obtain the following system of recursive equations for determining $\chi_j(x,p,y)$ and $L_j(x,p,\mu)$:

$$\mathcal{H}_0\chi_0 - \chi_0L_0 = 0, \tag{4.2}$$

$$\mathcal{H}_0\chi_j - \chi_jL_0 = \chi_0L_j + \mathcal{F}_j, \qquad j = 1,2,\dots, \tag{4.3}$$

where each function $\mathcal{F}_j$ can be expressed only in terms of χ_s and L_s with $0 \le s \le j-1$. In particular,

$$\begin{aligned}
\mathcal{F}_1 &= -\mathcal{H}_1\chi_0 + i\langle\nabla_p\mathcal{H}_0,\nabla_x\chi_0\rangle - i\langle\nabla_p\chi_0,\nabla_xL_0\rangle,\\
\mathcal{F}_2 &= \chi_1L_1 - \mathcal{H}_1\chi_1 - \mathcal{H}_2\chi_0 + i\langle\nabla_p\mathcal{H}_1,\nabla_x\chi_0\rangle - i\langle\nabla_p\chi_1,\nabla_xL_0\rangle + i\langle\nabla_p\mathcal{H}_0,\nabla_x\chi_1\rangle\\
&\qquad - i\langle\nabla_p\chi_0,\nabla_xL_1\rangle + \frac{1}{2}\sum_{j,k}\biggl(\frac{\partial^2\mathcal{H}_0}{\partial p_j\partial p_k}\frac{\partial^2\chi_0}{\partial x_j\partial x_k} - \frac{\partial^2\chi_0}{\partial p_j\partial p_k}\frac{\partial^2L_0}{\partial x_j\partial x_k}\biggr).
\end{aligned} \tag{4.4}$$

Now let us analyze Eqs. (4.2) and (4.3) in the case where the symbol $\mathcal{H}_0$ is given by formula (2.4). Equation (4.2) means that, for fixed values of the parameters x and p, χ_0 is an eigenfunction of the elliptic operator $\mathcal{H}_0$ on a compact manifold $\mathbb{T}$. It follows from the general theory of elliptic operators on compact manifolds that $\mathcal{H}_0$ with the domain $C^\infty(\mathbb{T})$ is essentially self-adjoint in $L_2(\mathbb{T})$ and has a complete system of eigenfunctions χ_0^k, $k = 0,1,\dots$, corresponding to the real eigenvalues $L_0^k(x,p)$ (see, e.g., [16]). We note that for $p = 0$, the operator $\mathcal{H}_0$ becomes $-\Delta_y^{\Theta}$, i.e., an elliptic operator on the torus $\mathbb{T}$, and its minimal eigenvalue and the corresponding normed eigenfunction can be found easily and are well known:

$$L_0^0(x,0)=0, \qquad \chi_0^0(x,0,y)=1; \tag{4.5}$$

in this case, the eigenvalue $L_0^0(x,0)=0$ is nondegenerate and hence is separated from the others. These assertions can standardly be derived from the "energy relation"

$$(-\Delta_y^\Theta u,u)_{L^2(\mathbb{T})}=(\nabla_y^\Theta u,C^2\nabla_y^\Theta u)_{L^2(\mathbb{T})}.$$

Since, in what follows, we consider only the minimal eigenvalue and the corresponding eigenfunction, we omit the superscript 0 to avoid the cumbersome notation and write

$$L_0^0=L_0, \qquad \chi_0^0=\chi_0.$$

As was shown in [18], for x from a compact set K in $\mathbb{R}^3$ and sufficiently small p, the eigenvalue L_0 of the operator $\mathcal{H}_0$ is nondegenerate and analytic in p and the function $\chi_0^0(x,p,y)$ can be chosen smooth in (x,p) and analytic in p (a simple proof of these assertions is also given in [19]).

In what follows, we assume that x and p are chosen precisely in this way. For χ_0 and L_0 in (4.2) we take χ_0 and L_0, respectively. We shall solve Eqs. (4.3) in succession. If all equations with numbers less than j have already been solved, then to determine L_j^0, it suffices, in the sense of the inner product in $L_2(\mathbb{T})$, to take the scalar product of the left- and right-hand sides by χ_0, to use the fact that the operator $\mathcal{H}_0-L_0$ is self-adjoint, and to apply this operator to the second multiplier of χ_0. Since (4.2) is satisfied and the function χ_0 is normed, we obtain

$$L_j(x,p)=-(\mathcal{F}_j(x,p,y),\chi_0(x,p,y))_{L^2(\mathbb{T})}. \tag{4.6}$$

After this, the existence of the solution χ_j of Eq. (4.3) is a consequence of the following assertion.

Lemma 2. *Suppose that $F(y)$ is a smooth function on the torus $\mathbb{T}$ and is orthogonal in $L_2(\mathbb{T})$ to the function $\chi^0(x,p,y)$ (for fixed (x,p)). For $x\in K$ and sufficiently small p, there exists a solution $f\in C^\infty(\mathbb{T})$ of the equation*

$$(\mathcal{H}_0-L_0)f=F; \tag{4.7}$$

this solution is unique and satisfies the condition that f is orthogonal to the function $\chi_0(x,p,y)$.

If F is a smooth function of the variables $x\in K$, $y\in\mathbb{T}$, and p (for small p) and of some additional parameters z, then the solution $f(x,p,y,z)$ is also a smooth function. Any other smooth solution $f_1(x,p,y,z)$ on the torus $\mathbb{T}$ of Eq. (4.7) is expressed in terms of f by the formula $f_1=f+g\chi_0$, where $g(x,p,z)$ is a smooth function of the parameters (x,p,z).

Proof. It follows from the general theory of elliptic equations on compact manifolds that $\mathcal{H}_0-L_0$, as an operator of the Sobolev space $H^2(\mathbb{T})$, is Fredholm in $L^2(\mathbb{T})$. Therefore, the solvability condition for Eq. (4.7) means that the right-hand side is orthogonal to the solution χ_0 of the homogeneous equation, and the solution f orthogonal to χ_0 can be determined uniquely.

The infinite differentiability of the functions contained in them can easily be proved by using the general theory of elliptic operators on compact manifolds (see, e.g., [16]). To this end, we consider the following problem: for fixed (x,p), according to a given function $F(y)$ and a number d, it is required to find a function $u(y)$ and a number g satisfying the equations

$$(\mathcal{H}_0-L_0)u(y)-g\chi_0(x,p,y)=F(y), \qquad (u(y),\chi_0)_{L^2(\mathbb{T})}=d. \tag{4.8}$$

Since the corresponding operator

$$A(x,p)\colon \mathcal{H}^{s+2}(\mathbb{T})\times C^1\to\mathcal{H}^s(\mathbb{T})\times C^1,$$

where C^1 is a one-dimensional complex space, is invertible and $A(x,p)$ smoothly depends on the parameters (x,p), the inverse operator $A^{-1}(x,p)$ also smoothly depends on the parameters (x,p). If for F we take the function $F(x,p,y,z)$, then we see that the solution $u(x,p,y,d,z)$ of problem (4.8) smoothly depends on the parameters (x,p,z). But $f(x,p,y,z)$ is precisely the solution of problem (4.8) for $d=0$ and $g=0$; therefore, $f(x,p,y,z)$ is a smooth function of the parameters (x,p,z) ranging in the space $\mathcal{H}^{s+2}(\mathbb{T})$ for any s. Now it follows from the standard embedding theorems for Sobolev spaces that $f(x,p,y,z)$ is an infinitely differentiable function with respect to the set of the variables. $\square$

5. PERTURBATION THEORY FOR THE EFFECTIVE HAMILTONIANS IN THE MOMENTUM VARIABLES

5.1. Heuristic Considerations about the Minimal Number of Terms L_j and Their Expansion in p in the Case of Simplified Reduced Equations for Different Solutions

As was already noted, it is not easy to write L_j in a sufficiently simple form in the general situation even for $j = 0, 1, 2$. But it seems that L_j with numbers $j \geq 3$ are, as a rule, not very interesting for physical applications. Moreover, depending on the behavior of C^2 and V and on some additional conditions determining the solutions of the reduced equations (and of the original equations), it is possible to simplify the reduced equations (3.4) themselves. Considerations about such a simplification, which are based on the WKB method, were illustrated in [8], [10] by the example of the Cauchy problem

$$i\mu\frac{\partial w}{\partial t} = \widehat{L}w, \qquad w|_{t=0} = a^0(x)e^{iS^0(x)/h}$$

with fast oscillating initial data characterized by the parameter $h = \mu^\kappa$, where $\kappa \leq 1$ is a nonnegative number. Here we present only the conclusions.

For small p, we write $L_j(x,p)$ and $\chi_j(x,p,y)$ as

$$\begin{aligned} L_j(x,p) &= \sum_{|k|\leq K} L_j^{(k)}(x,p) + O(|p|^{K+1}), \\ \chi_j(x,p,y) &= \sum_{|k|\leq K} \chi_j^{(k)}(x,p,y) + O(|p|^{K+1}), \end{aligned} \tag{5.1}$$

where $L_j^{(k)}(x,p)$ and $\chi_j^{(k)}(x,p,y)$ are homogeneous polynomials in p of degree k. In fact, expansions (5.1) are the Taylor series of the right-hand sides, and hence they admit the termwise differentiation with respect to the parameters x, p, and y.

We point out at once that

$$L_0 = L_0^{(2)} + O(|p|^4)$$

for the problems under study. Therefore, for $\kappa = 0$ (the "long-wave" mode), the solution oscillates weakly, and to describe the leading term of the asymptotics, it is necessary to calculate $L_0^{(2)}$, $L_1^{(1)}$, and $L_2^{(0)}|_{p=0}$, which gives the limit equations that are well known in the "classical" averaging theory [1]–[4]. To construct the leading term of the asymptotics for $0 < \kappa < 2/3$ (the "medium-wave" mode), it suffices to calculate $L_0^{(2)}$ and $L_1^{(1)}$. If κ is close to $2/3$, i.e., h is close to $\mu^{3/2}$, then these functions must be supplemented with $L_0^{(4)}$ for which it is still possible to write some constructive formulas; in the case of the wave equation, these formulas give the so-called "weak dispersion" and a reduced Boussinesq-type equation. If κ continues to increase, then the coefficients of L_0 and L_1 must be determined exactly, and we have the "short-wave" mode. A further increase in the parameter κ destroys the adiabatic approximation, although, for some problems, the parameter κ can be increased even to $3/2$, and then the WKB method becomes the Born approximation. In the present paper, we consider only the long- and medium-wave modes. The derivation of Boussinesq-type equations requires a great amount of calculations and place and will be considered in a separate paper. Here we present only the final formulas for $L_0^{(4)}$. We note that the above considerations can also be used in the case of fast decreasing functions and in the case where the asymptotics of the solution contains focal points and caustics, i.e., the above version of the adiabatic approximation covers all these cases.

5.2. Perturbation Formulas for Effective Hamiltonians in the Momentum Variables

Now let us derive formulas for the expansions of L_k in small p, which permits constructing the "medium-" and "long-wave" asymptotic solutions of the reduced equations. To find them, it suffices to solve Eqs. (4.2) and (4.3) for small p.

Let $G(y)$ be a function 2π-periodic in each of the variables y_j, which can possibly depend on some other variables. By $\langle G\rangle_{\mathbb{T}}$ we denote the mean value of this function on the torus $\mathbb{T}$:

$$\langle G\rangle_{\mathbb{T}} = \frac{1}{(2\pi)^n}\int_{\mathbb{T}} G(y)\,dy. \tag{5.2}$$

We need the following simple useful assertion. Let $F(y,z)$ be a smooth function on the torus $\mathbb{T}$ with zero mean $\langle F\rangle_{\mathbb{T}} = 0$ which also smoothly depends on the parameters $z = (z_1,\dots,z_l)$ that belong to a compact set. On the torus $\mathbb{T}$, we consider the following equation (the so-called "problem on the cell") for the function $f(y,z)$ (see [1]–[4]):

$$\Delta_y^{\Theta} f = F, \qquad \langle f\rangle_{\mathbb{T}} = 0. \tag{5.3}$$

Lemma 3. *Problem* (5.3) *has a smooth solution on the torus $\mathbb{T}$, and this solution is unique. Any other smooth solution $f_1(y,z)$ on the torus $\mathbb{T}$ of the equation $\Delta_y^{\Theta} f_1 = F$ can be expressed in terms of f by the formula $f_1 = f + q$, where $q(z)$ is a smooth function of the parameters z.*

Proof. The proof of this lemma follows from the assertion of Lemma 2 for $p = 0$. □

In what follows, this solution f with zero mean of the problem on the cell (5.3) will be denoted by

$$f(y,z) = \frac{1}{\Delta_y^{\Theta}} F(y,z), \qquad \langle f\rangle_{\mathbb{T}} = 0. \tag{5.4}$$

This lemma allows us not only to prove the statement about the behavior of L_0 and χ_0 for small p, but also to obtain sufficiently explicit formulas for the first terms of the expansion of these functions in p. To formulate that statement, we introduce the following notation. We set

$$\widetilde{C}^2(y,x) = C^2(y,x) - \langle C^2\rangle_{\mathbb{T}}, \tag{5.5}$$

and, in addition to the operator $D = \langle p, \Theta_x(x)\nabla_y\rangle$ linearly depending on p, we introduce the operator

$$Q = DC^2 + C^2 D. \tag{5.6}$$

that also linearly depends on p. By $g_0(y,x)$ and $g_1(y,x,p)$ we denote the solutions with zero mean of the problem on the cell:

$$g_0 = \frac{1}{\Delta_y^{\Theta}} \widetilde{C}^2(y,x), \qquad g_1 = \frac{1}{\Delta_y^{\Theta}} (D\widetilde{C}^2(y,x)), \qquad \langle g_{1,2}\rangle_{\mathbb{T}} = 0. \tag{5.7}$$

We note that $g_1(y,x,p)$ is a linear function of p.

Lemma 4. *For x belonging to the compact set K and sufficiently small p, the minimal eigenvalue $L_0(x,p)$ of the operator $\mathcal{H}_0$ is nondegenerate. The functions $L_0(x,p)$, L_1, and L_2 are analytic in p, the functions $\chi_0(x,p,y)$, $\chi_1(x,p,y)$, and $\chi_1(x,p,y)$ can also be chosen analytic in p, and hence the following relations hold*:

$$\begin{aligned} &L_0(x,p) = p^2\langle C^2\rangle_{\mathbb{T}} - \langle Qg_1\rangle_{\mathbb{T}} + O(|p|^4) \equiv p^2\langle C^2\rangle_{\mathbb{T}} - \langle C^2 Dg_1\rangle_{\mathbb{T}} + O(|p|^4), \\ &\chi_0(x,p) = 1 - ig_1(y,x,p) + O(|p|^2), \end{aligned} \tag{5.8}$$

where $\|1 - ig_1(y,x,p)\|_{L_2(\mathbb{T})} = 1 + O(|p|^2)$,

$$L_1(x,p) = -\frac{i}{2}\langle\nabla_x,\nabla_p\rangle L_0 + O(p^2) \equiv i\langle\langle\nabla_x, C^2\nabla_y^{\Theta} g_1\rangle\rangle_{\mathbb{T}} - i\langle\nabla_x, p\langle C^2\rangle_{\mathbb{T}}\rangle + O(p^2), \tag{5.9}$$

$$L_2(x,p) = \langle V(y,x)\rangle_{\mathbb{T}} + O(p). \tag{5.10}$$

Corollary. *The Weyl symbol of the operator $\widehat{L}$ satisfies the relation*

$$L_W(p,x,\mu) = L_0 + \mu^2\langle V(y,x)\rangle_{\mathbb{T}} + O(|p|^3) + \mu^2 O(|p|) + O(\mu^3). \tag{5.11}$$

Remark 6. We present the formulas for the following correction to the expansions of L_0 and χ_0 without proofs:

$$L_0(p,x) = p^2\langle C^2\rangle_{\mathbb{T}} - \langle Qg_1\rangle_{\mathbb{T}} + p^4\langle g_0\widetilde{C}^2\rangle_{\mathbb{T}} + 2p^2\langle g_1 Q g_0\rangle_{\mathbb{T}} + \langle g_1^2\rangle_{\mathbb{T}}\langle Qg_1\rangle_{\mathbb{T}} + p^2\langle g_1^2\widetilde{C}^2\rangle_{\mathbb{T}}$$
$$+ \left\langle (Qg_1 - \langle Qg_1\rangle_{\mathbb{T}})\frac{1}{\Delta_y^\Theta}(Qg_1 - \langle Qg_1)\right\rangle_{\mathbb{T}} + O(|p|^6),$$
$$\chi_0 = 1 - ig_1(y,x,p) + p^2 g_0(y,x) - \frac{1}{\Delta_y^\Theta}(Qg_1 - \langle Qg_1\rangle_{\mathbb{T}}) - \frac{1}{2}\langle g_1^2\rangle_{\mathbb{T}} + O(|p|^3),$$
$$\left\| 1 - ig_1(y,x,p) + p^2 g_0(y,x) - \frac{1}{\Delta_y^\Theta}(Qg_1 - \langle Qg_1\rangle_{\mathbb{T}}) - \frac{1}{2}\langle g_1^2\rangle_{\mathbb{T}} \right\|_{L_2(\mathbb{T})} = 1 + O(|p|^3).$$

Proof of Lemma 4. As was already noted, the first part of this assertion follows from the general theorems given in [18], and hence it remains to show that the first coefficients of the expansion of L_0, χ_0, L_1, etc. in the power series (5.1) have the same form as in (5.8)–(5.10).

1. *Calculation of the coefficients of the expansions of* L_0 *and* χ_0. Substituting (5.1) into (4.2) and setting the homogeneous polynomials of the same order in p equal to zero, we obtain the following system of recursive equations for $L_0^{(k)}$ and $\chi_0^{(k)}$:

$$-\Delta_y^\Theta \chi_0^{(0)} - L_0^{(0)}\chi_0^{(0)} = 0,$$
$$(-\Delta_y^\Theta - L_0^{(0)})\chi_0^{(1)} - L_0^{(1)}\chi_0^{(0)} = i(DC^2(y,x) + C^2(y,x)D)\chi_0^{(0)},$$
$$(-\Delta_y^\Theta - L_0^{(0)})\chi_0^{(2)} - L_0^{(2)}\chi_0^{(0)} = i(DC^2(y,x) + C^2(y,x)D)\chi_0^{(1)}$$
$$- p^2 C^2(y,x)\chi_0^{(0)} + L_0^{(1)}\chi_0^{(1)},$$
$$(-\Delta_y^\Theta - L_0^{(0)})\chi_0^{(k)} - L_0^{(k)}\chi_0^{(0)} = F_k, \qquad k = 3,4,\dots,$$

where F_k is expressed in terms of $\chi_0^{(j)}$ and $L_0^{(j)}$ with $j = 0,\dots,k-1$.

The first relation together with the normalization condition implies $L_0^{(0)} = 0$ and $\chi_0^{(0)} = 1$. All subsequent relations have the following general form: the unknowns in the left-hand side are $L_0^{(k)}$ and $\chi_0^{(k)}$ (if we take into account that $L_0^{(0)}$ and $\chi_0^{(0)}$ have already been determined), and the right-hand side contains a function determined by the preceding relations. Then, to determine $L_0^{(k)}$, it suffices, to take the scalar product in $L_2(\mathbb{T})$ of the left- and right-hand sides by the function $\chi_0^{(0)}$, to use the condition that the operator $-\Delta_y^\Theta - L_0^{(0)}$ is self-adjoint, to apply this operator to the second multiplier of $\chi_0^{(0)}$ in the scalar product, which is equal to zero according to the first relation, and to use the fact that χ_0 is normed, i.e., the fact that the obtained coefficient of $L_0^{(k)}$ is equal to 1. Thus, $L_0^{(k)}$ is uniquely determined by the preceding $\chi_0^{(j)}$ and $L_0^{(j)}$, $j = 0,\dots,k-1$. Therefore, the second relation implies

$$L_0^{(1)} = -i\langle DC^2\rangle = 0,$$

because the mean of the differentiated function is equal to zero. Taking the relations

$$L_0^{(0)} = L_0^{(1)} = 0 \qquad \text{and} \qquad \chi_0^{(0)} = 1$$

into account, we can rewrite the equation for $\chi_0^{(1)}$ in the form

$$\Delta_y^\Theta \chi_0^{(1)} = -i(D\widetilde{C}^2(y,x)),$$

which implies (5.8).

To determine $\chi_0^{(k)}$, it is necessary to transfer $L_0^{(k)}\chi_0^{(k)}$ to the right-hand side, to note that the mean of the right-hand side is equal to zero by the definition of $L_0^{(k)}$, and to apply Lemma 3. It follows from this lemma that $\chi_0^{(k)}$ is determined not uniquely but up to a constant (with respect to y) $q_k = O(p^k)$. We can study the problem only on the cell, i.e., choose solutions with zero mean (as was performed for $\chi_0^{(1)}$), but then the normalization condition $(\chi_0, \chi_0) = 1$ is not guaranteed. We immediately note that the normalization condition $\chi_0^{(0)} + \chi_0^{(1)}$ is satisfied with the required accuracy of $O(p^2)$, because

$$(\chi_0^{(0)} + \chi_0^{(1)}, \chi_0^{(0)} + \chi_0^{(1)}) = 1 + ((\chi_0^{(0)}, \chi_0^{(1)}) + (\chi_0^{(1)}, \chi_0^{(0)})) + (\chi_0^{(1)}, \chi_0^{(1)}),$$

the middle term is equal to zero because of the cellular structure of $\chi_0^{(1)}$, and the third term is a homogeneous polynomial of degree two $(\chi_0^{(1)}, \chi_0^{(1)}) = O(p^2)$.

It is easy to see that $\chi_0^{(k)}$ is also not uniquely determined by the normalization condition (just as the entire function, $\chi_0^{(0)}$ is determined up to multiplication by the phase multiplier $e^{i\theta(x,p)}$ with a real $\theta(x, p)$). The term $L_0^{(2)}$ is determined independently of the choice of the constant q_1, which can easily be verified from the second relation, just as the fact that

$$L_0^{(2)} = p^2\langle C^2\rangle - \langle C^2 D\chi_0^{(1)}\rangle.$$

The analyticity of the further expansion in the series $\sum \chi_0^{(k)}$ could, in principle, depend on this choice. But this is not the case, because the analytic function χ_0 can be modified by using the phase multiplier $e^{i\theta} = 1 + i\theta_1 + \cdots$ without changing its analyticity. And choosing the first term, we can obtain the desired term $\chi_0^{(1)}$, because $\operatorname{Re} q_1 = 0$ by the normalization condition.

We note that the system of recursive relations for $L_0^{(k)}$ and $\chi_0^{(k)}$ can be solved even if the explicit form of the eigenfunction χ_0 is unknown. For example, let us solve the equation for $\chi_0^{(1)}$ by the method of undetermined coefficients presenting $\chi_0^{(1)}$ as

$$\chi_0^{(1)} = \langle b, p\rangle = b_1(x, y)p_1 + b_2(x, y)p_2 + b_3(x, y)p_3,$$

where the coefficients $b_j(x, y)$, $j = 1, 2, 3$, are unknown. Then, to determine each $b_j(x, y)$, it is necessary to solve a cellular problem of the form (5.3). Similarly, the problem of determining each of $\chi_0^{(k)}$ can be reduced to solving several cellular problems of the form (5.3) for the coefficients of the corresponding homogeneous polynomial in the variables p.

2. *Calculation of the coefficients of the expansions of L_1 and L_2.* Similarly, the series (5.1) can be used to solve the other equations in (4.3). But it is possible to use formulas (4.4), (4.6) at once. Taking these formulas and (2.4), (2.5) into account, we obtain

$$\begin{aligned}\mathcal{F}_1 &= i\langle\nabla_x, C^2(y, x)p\rangle + \langle\nabla_x, C^2(y, x)\nabla_y^\theta\rangle(-ig_1) + i\langle\nabla_p(-iQ), \nabla_x(-ig_1)\rangle + O(p^2)\\ &\equiv i\langle\nabla_x, C^2(y, x)p\rangle - i\langle\nabla_x, C^2(y, x)\nabla_y^\theta\rangle g_1 - i\langle\nabla_p Q, \nabla_x g_1\rangle + O(p^2).\end{aligned} \tag{5.12}$$

Hence, using the definition of Q, we can write

$$L_1 = -i\langle\langle\nabla_x, C^2(y, x)p\rangle\rangle_{\mathbb{T}} + i\langle\nabla_p Q, \nabla_x g_1\rangle_{\mathbb{T}} + O(p^2), \qquad \chi_1 = O(p). \tag{5.13}$$

Taking the last relations and the relations for L_0 and χ_0 into account, we obtain

$$\nabla_x\chi_0 = O(p), \quad \nabla_x L_0 = O(p), \quad \nabla_x\chi_1 = O(p), \quad \nabla_x L_1 = O(p),$$
$$\frac{\partial^2\chi_0}{\partial x_j\partial x_k} = O(p), \quad \frac{\partial^2 L_0}{\partial x_j\partial x_k} = O(p^2).$$

This immediately implies

$$\mathcal{F}_2 = -\mathcal{H}_2\chi_0 + O(p) = -V(y, x) + O(p), \qquad L_2 = \langle V(y, x)\rangle_{\mathbb{T}} + O(p). \tag{5.14}$$

To prove the lemma, it remains to write the term $\nabla_p\langle Qg_1\rangle_{\mathbb{T}}$ in a different way. We have

$$\nabla_p\langle Qg_1\rangle_{\mathbb{T}} = \langle\nabla_p Qg_1\rangle_{\mathbb{T}} + \langle Q\nabla_p g_1\rangle_{\mathbb{T}} = \langle C^2\nabla_y^\theta g_1\rangle_{\mathbb{T}} + \left\langle C^2 D\frac{1}{\Delta_y^\theta}\nabla_y^\theta C^2\right\rangle_{\mathbb{T}}$$

$$= \langle C^2\nabla_y^\theta g_1\rangle_{\mathbb{T}} + \left\langle \nabla_y^\theta \frac{1}{\Delta_y^\theta} DC^2C^2 \right\rangle_{\mathbb{T}} = 2\langle C^2\nabla_y^\theta g_1\rangle_{\mathbb{T}}.$$

Now we apply the operator ∇_y, move this operator under the averaging operation, and obtain the first relations in (5.9). □

6. AVERAGED EQUATIONS

Recall that we are solving the operator equation $\widehat{\chi}\widehat{L} = \widehat{\mathcal{H}}\widehat{\chi}$ equivalent to (3.7) by using perturbation theory. First, we expand the corresponding symbols L and χ in asymptotic power series in the parameter μ and obtain Eqs. (4.2) and (4.3); then for L_j and χ_j we pass to the expansions in homogeneous polynomials in the variables p and obtain the corresponding system of recursive relations for $L_j^{(k)}$ and $\chi_j^{(k)}$ for each j, similarly to the proof of Lemma 4.

We introduce an integer $N \geq 2$ and treat $(L_j)_N$ and $(\chi_j)_N$ as the corresponding sums of terms $L_j^{(k)}$ and $\chi_j^{(k)}$ with $k \leq N - j$; then we treat $(L)_N$ and $(\chi)_N$ as the sums of terms $\mu^j L_j$ and $\mu^j\chi_j$ for $j \leq N$, i.e., $(L)_N$ and $(\chi)_N$ are polynomials of degree N in the variables (p,μ). Taking into account that, for such symbols, the composition formulas of the form (4.1) contain only finitely many terms, we obtain the operator relation

$$(\widehat{\chi})_N(\widehat{L})_N = \widehat{\mathcal{H}}(\widehat{\chi})_N + \widehat{r}_N, \tag{6.1}$$

where $r_N(x,p,y,\mu)$ is a polynomial in the variables (p,μ) and

$$r_N = O(|p| + |\mu|)^{N+1}.$$

Replacing p by ξ, we write $r_N(x,p,y,\mu)$ as

$$r_N(x,p,y,\mu) = \mu^{N+1}R_N(x,\xi,y,\mu),$$

where $R_N(x,\xi,y,\mu)$ is a polynomial in the variables (ξ,μ); hence $\widehat{r}_N$ in (6.1) can be replaced by

$$\widehat{r}_N = \mu^{N+1}R_N\left(x, -i\frac{\partial}{\partial x}, y, \mu\right). \tag{6.2}$$

Instead of (3.4) and (3.5), for the reduced equations we take

$$i\mu^2 w_t = (\widehat{L}^{(0)})_N w, \tag{6.3}$$

$$\mu^2 w_{tt} = -(\widehat{L}^{(0)})_N w. \tag{6.4}$$

First, we consider the case $N = 2$. Then

$$(L)_2 = L_0^{(0)} + L_0^{(1)} + L_0^{(2)} + \mu(L_1^{(0)} + L_1^{(1)}) + \mu^2 L_2^{(0)},$$
$$(\chi)_2 = \chi_0^{(0)} + \chi_0^{(1)} + \chi_0^{(2)} + \mu(\chi_1^{(0)} + \chi_1^{(1)}) + \mu^2\chi_2^{(0)},$$

and the separate terms are determined by the corresponding recursive relations obtained in the proof of Lemma 4, so that the assertions of this lemma hold for them, in particular, $L_0^{(0)} = L_0^{(1)} = L_1^{(0)} = 0$. Now the reduced equation (6.3) becomes

$$i\mu^2 w_t = (L_0^{(2)}(x,\widehat{p}) + \mu L_1^{(0)}(x,\widehat{p}) + \mu^2 L_2^{(0)}(x,\widehat{p}))w, \qquad \widehat{p} = -i\mu\nabla_x. \tag{6.5}$$

We represent $\chi_0^{(1)} = -ig_1$ as

$$\chi_0^{(1)} = -i(b_1(y,x)p_1 + b_2(y,x)p_2 + b_3(y,x)p_3)$$

and apply formulas (5.8). Then we can write the symbol $L_0^{(2)}$ as

$$L_0^{(2)} = \langle C^2\rangle_{\mathbb{T}} + \sum_{k,j}\langle C^2\nabla_y^\Theta b_j\rangle_{\mathbb{T}}^k p_j p_k,$$

where $\langle C^2\nabla_y^\Theta b_j\rangle_{\mathbb{T}}^k$ is the kth component of the vector $\langle C^2\nabla_y^\Theta b_j\rangle_{\mathbb{T}}$. Similarly, from (5.9) and (5.10) we obtain

$$L_1^{(1)} = i\sum_{k,j}\frac{\partial\langle C^2\nabla_y^\Theta b_j\rangle_{\mathbb{T}}^k}{\partial x_k}p_j - i\sum_k\frac{\partial\langle C^2\rangle_{\mathbb{T}}}{\partial x_k}p_k, \qquad L_{2\,0}^0 = \langle V\rangle_{\mathbb{T}}.$$

Taking into account the relation

$$\langle C^2\rangle_{\mathbb{T}}\frac{\partial^2}{\partial x_k^2}w + (\langle C^2\rangle_{\mathbb{T}})_{x_k}\frac{\partial}{\partial x_k} = \frac{\partial}{\partial x_k}\langle C^2\rangle_{\mathbb{T}}\frac{\partial}{\partial x_k}$$

and the fact that the same relation holds for $\langle C^2\nabla_y^\Theta b_j\rangle_{\mathbb{T}}^k$, we can write the reduced equation (6.5) after eliminating the multiplier μ^2 in the form

$$i\frac{\partial}{\partial t}w = -\sum_{k,j}\frac{\partial}{\partial x_k}\biggl((\langle C^2\nabla_y^\Theta b_j\rangle_{\mathbb{T}}^k + \langle C^2\rangle_{\mathbb{T}})\frac{\partial w}{\partial x_j}\biggr) + \langle V\rangle_{\mathbb{T}}w, \tag{6.6}$$

and Eq. (6.4) for $N = 2$ after eliminating the multiplier μ^2, in the form

$$\frac{\partial^2}{\partial t^2}w = \sum_{k,j}\frac{\partial}{\partial x_k}\biggl((\langle C^2\nabla_y^\Theta b_j\rangle_{\mathbb{T}}^k + \langle C^2\rangle_{\mathbb{T}})\frac{\partial w}{\partial x_j}\biggr) - \langle V\rangle_{\mathbb{T}}w. \tag{6.7}$$

These are precisely the averaged equations according to the terminology introduced in [1]. Appropriately changing the argument in the proof of Lemma 1 and taking (6.1), (6.3), and (6.4) into account, we obtain the following statement.

Theorem 1. *If a function $w(x,t)$ is a solution of* (6.6) (*or of* (6.7))*, then $\Psi = (\widehat{\chi}^0)_2 w$ satisfies the equation*

$$i\mu^2\Psi_t = \widehat{H}\Psi + \mu^3 R_2\biggl(x, -i\frac{\partial}{\partial x}, y, \mu\biggr)w \tag{6.8}$$

or

$$\mu^2\Psi_{tt} = -\widehat{H}\Psi - \mu^3 R_2\biggl(x, -i\frac{\partial}{\partial x}, y, \mu\biggr)w. \tag{6.9}$$

Thus, the function Ψ satisfies Eqs. (2.2 a,b) respectively up to $O(\mu^3)$.

Proof of Theorem 1. We repeat the argument of the proof of Lemma 1, use (6.1), (6.2), (6.3), and (6.6), and obtain

$$i\mu^2\Psi_t = i\mu^2(\widehat{\chi})_2 w_t = (\widehat{\chi})_2(\widehat{L})_2 w = (\widehat{\mathcal{H}}(\widehat{\chi})_2 + \widehat{r}_2)w = \widehat{\mathcal{H}}\Psi + \mu^3 R_2\biggl(x, -i\frac{\partial}{\partial x}, y, \mu\biggr)w.$$

Relation (6.9) can be proved similarly. $\square$

Now let us consider the case of an arbitrary $N \ge 2$. The operator $(\widehat{L})_N$, which is obtained by replacing p by $\widehat{p} = -i\mu\nabla_x$ (under the assumption that the operator $\widehat{p}$ acts first, and x acts second), has the form

$$(\widehat{L})_N = \mu^2 A_2\biggl(x, -i\frac{\partial}{\partial x}\biggr) + \mu^3 A_3\biggl(x, -i\frac{\partial}{\partial x}\biggr) + \dots + \mu^N A_N\biggl(x, -i\frac{\partial}{\partial x}\biggr),$$

$$A_s(x,\xi) = \sum_{j=0}^{s} L_j^{(s-j)}(x,\xi),$$

where $A_s(x, -i\partial/\partial x)$ are differential operators of order s. Now the reduced equation (6.3) after the elimination of μ^2 becomes

$$iw_t = \biggl(A_2\biggl(x, -i\frac{\partial}{\partial x}\biggr) + \mu A_3\biggl(x, -i\frac{\partial}{\partial x}\biggr) + \dots + \mu^{N-2}A_N\biggl(x, -i\frac{\partial}{\partial x}\biggr)\biggr)w, \tag{6.10}$$

and Eq. (6.4) after the elimination of μ^2 becomes

$$w_{tt} = -\left(A_2\left(x, -i\frac{\partial}{\partial x}\right) + \mu A_3\left(x, -i\frac{\partial}{\partial x}\right) + \cdots + \mu^{N-2}A_N\left(x, -i\frac{\partial}{\partial x}\right)\right)w. \tag{6.11}$$

Assume that we have a function $w(y, x, t)$ that, for a fixed μ, satisfies Eq. (6.10) (or (6.11)) with a "discrepancy" in the right-hand side

$$iw_t = \left(A_2\left(x, -i\frac{\partial}{\partial x}\right) + \mu A_3\left(x, -i\frac{\partial}{\partial x}\right) + \cdots + \mu^{N-2}A_N\left(x, -i\frac{\partial}{\partial x}\right)\right)w + F(y, x, t) \tag{6.12}$$

or

$$w_{tt} = -\left(A_2\left(x, -i\frac{\partial}{\partial x}\right) + \mu A_3\left(x, -i\frac{\partial}{\partial x}\right) + \cdots + \mu^{N-2}A_N\left(x, -i\frac{\partial}{\partial x}\right)\right)w - F(y, x, t); \tag{6.13}$$

then $i\mu^2 w_t = (L^0)_N w + \mu^2 F$ or $\mu^2 w_{tt} = -(L^0)_N w - \mu^2 F$, respectively. In this case, we can use the intertwining operator $(\widehat{\chi}^0)_N$ to construct the function $\Psi = (\widehat{\chi}^0)_N w$ for which Eq. (2.2 a) (or (2.2 b)) is also satisfies with a certain "discrepancy" in the right-hand side.

Theorem 2. *If a function $w(x, t)$ is a solution of* (6.12) (*or* (6.13))*, then $\Psi = (\widehat{\chi}^0)_N w$ satisfies the equation*

$$i\mu^2\Psi_t = \widehat{H}\Psi + \mu^{N+1}R_N\left(x, -i\frac{\partial}{\partial x}, y, \mu\right)w + \mu^2(\widehat{\chi})_N F \tag{6.14}$$

or

$$\mu^2\Psi_{tt} = -\widehat{H}\Psi - \mu^{N+1}R_N\left(x, -i\frac{\partial}{\partial x}, y, \mu\right)w - \mu^2(\widehat{\chi})_N F. \tag{6.15}$$

Proof. Repeating the argument of the proof of Theorem 1, we use (6.1), (6.2), and (6.12) and obtain

$$\begin{aligned} i\mu^2\Psi_t &= i\mu^2(\widehat{\chi})_N w_t = (\widehat{\chi})_N(\widehat{L})_N w + \mu^2(\widehat{\chi})_N F = (\widehat{\mathcal{H}}(\widehat{\chi})_N + \widehat{r}_N)w + \mu^2(\widehat{\chi})_N F \\ &= \widehat{\mathcal{H}}\Psi + \mu^{N+1}R_N\left(x, -i\frac{\partial}{\partial x}, y, \mu\right)w + \mu^2(\widehat{\chi})_N F. \end{aligned}$$

Relation (6.15) can be proved similarly. □

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 11-01-00973) and by DFG—Russian Academy of Sciences grant no. 436 RUS 113/990/0-1.

REFERENCES

1. N. S. Bakhvalov and G. P. Panasenko, *Averaging Processes in Periodic Media: Mathematical Problems of Mechanics of Composite Materials* (Nauka, Moscow, 1984) [in Russian].
2. V. V. Zhikov, S. M. Kozlov, and O. A. Oleinik, *Homogenization of Differential Operators* (Fizmatlit, Moscow, 1993) [in Russian].
3. V. A. Marchenko and E. Ya. Khruslov, *Averaged Models of Micro-Inhomogeneous Media* (B. I. Verkin Institute for Low Temperature Physics and Engineering of the National Academy of Sciences of Ukraine, Kharkov, 2003) [in Russian].
4. A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic Analysis of Periodic Structures*, in *Stud. Math. Appl.* (North-Holland Publ., Amsterdam, 1978), Vol. 5.
5. S. Yu. Dobrokhotov, "Application of Maslov's theory to two problems with operator-valued symbol," Uspekhi Mat. Nauk **39** (4), 125 (1984).
6. V. P. Maslov, *Perturbation Theory and Asymptotic Methods* (Izdat. Moscow State University, Moscow, 1965) [in Russian].
7. L. V. Berlyand and S. Yu. Dobrokhotov, ""Operator separation of variables" in problems of short-wave asymptotic behavior for differential equations with fast oscillating coefficients," Dokl. Akad. Nauk SSSR **296** (1), 80–84 (1987) [Soviet Math. Dokl. **32** (9), 714–716 (1987)].

8. V. V. Belov, S. Yu. Dobrokhotov, and T. Ya. Tudorovskiy, "Operator separation of variables for adiabatic problems in quantum and wave mechanics," J. Engrg. Math. **55** (1-4), 183–237 (2006).
9. V. V. Grushin and S. Yu. Dobrokhotov, "Peierls substitution and the Maslov operator method," Mat. Zametki **87** (4), 554–571 (2010) [Math. Notes **87** (3–4), 521–536 (2010)].
10. J. Brüning, V. V. Grushin, S. Yu. Dobrokhotov, and T. Ya. Tudorovskii, "Generalized Foldy–Wouthuysen transformation and pseudodifferential operators," Teoret. Mat. Fiz. **167** (2), 171–192 (2011) [Theoret. and Math. Phys. **167** (2), 547–566 (2011)].
11. V. P. Maslov, *Operator Methods* (Nauka, Moscow, 1973) [in Russian].
12. V. V. Belov, S. Yu. Dobrokhotov, V. P. Maslov, and T. Ya. Tudorovskii, "A generalized adiabatic principle for electron dynamics in curved nanostructures," Uspekhi Fiz. Nauk **175** (9), 1004–1010 (2005) [Physics–Uspekhi **48** (9), 962–968 (2005)].
13. E. M. Lifshits and L. P. Pitaevskii, *Theoretical Physics*, Vol. 9: *Statistical Physics*, Pt. 2: *Theory of Condensed State* (Nauka, Moscow, 1978) [in Russian].
14. V. V. Kucherenko, "Asymptotics of the solution of the system $A(x, -ih\frac{\partial}{\partial x})$ as $h \to 0$ in the case of characteristics of variable multiplicity," [in Russian]. Izv. Akad. Nauk SSSR Ser. Mat. **38** (3), 625–662 (1974) [Math. USSR-Izv. **8** (3), 631–666 (1974)].
15. L. Hörmander, *The Analysis of Linear Partial Differential Operators*, Vol. 3: *Pseudodifferential Operators* (Springer-Verlag, Heidelberg, 1983; Mir, Moscow, 1987).
16. M. A. Shubin, *Pseudodifferential Operators and Spectral Theory* (Nauka, Moscow, 1978) [in Russian].
17. V. P. Maslov and M. V. Fedoryuk, *Semiclassical Approximation for Equations of Quantum Mechanics* (Nauka, Moscow, 1976) [in Russian].
18. T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Heidelberg, 1966; Mir, Moscow, 1972).
19. V. V. Grushin, "Multiparameter perturbation theory of Fredholm operators applied to Bloch functions," Mat. Zametki **86** (6), 819–828 (2009) [Math. Notes **86** (5–6), 767–774 (2009)].