Commun. Math. Phys. 117, 387–403 (1988)

Communications in Mathematical Physics

High Order Corrections to the Time-Dependent Born–Oppenheimer Approximation. II: Coulomb Systems

George A Hagedorn*

Department of Mathematics and Center for Transport Theory and Mathematical Physics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-4097, USA

Abstract. We study the dynamics of molecular systems with Coulomb forces. We prove that if the nuclear masses are proportional to ε^{-4}, then certain solutions to the time dependent Schrödinger equation have asymptotic expansions to arbitrarily high order in powers of ε, as $\varepsilon \searrow 0$.

1. Introduction

In this paper we consider the quantum dynamics of molecular systems with Coulomb potentials. We study these systems by exploiting the smallness of the ratio ε^4 of the electron masses to nuclear masses. We prove that certain solutions to the time dependent Schrödinger equation have arbitrarily high order asymptotic expansions in powers of ε as $\varepsilon \searrow 0$.

The smallness of the mass ratio ε^4 was used by Born and Oppenheimer [1] in 1927 to study the energy levels of molecules. They formally showed that these levels had asymptotic expansions through fourth order in ε, and that the non-zero terms had direct physical interpretations. Approximately fifty years later, these results were rigorously proved by Aventini, Combes, Duclos, Grossman, and Seiler [2–4] (see also [15], where a simple, exactly soluble Harmonic oscillator example is presented). In [10] the energy level expansions were extended to arbitrarily high order for the case of smooth potentials. The high order Coulomb case has only been analyzed [11] for diatomic molecules, and is technically very complicated.

Born and Oppenheimer did not consider time dependent problems in [1], but in 1928, London [14] had the proper intuition for qualitatively understanding the dynamics for small ε. The only rigorous time-dependent results of which we are aware deal exclusively with smooth potentials [8, 9]. In [8], a zeroth order expansion was obtained, and in [9], the expansion was developed to arbitrarily high order. Thus, the present paper is an extension of [9] to accommodate Coulomb potentials.

* Supported in part by the National Science Foundation under grant number DMS-8601536

From our point of view, the validity of Born–Oppenheimer type expansions depends primarily on the presence of two spatial scales in the nuclear coordinates. The principal technical tool of [9–11] and the present paper is the method of multiple scales. Coulomb potentials cause difficulties because of cusps in the electron wave functions at the positions of the nuclei. In [11] this difficulty was handled by a global distortion of the electron coordinate system. This global distortion is not appropriate if there are more than two nuclei. To handle general polyatomic systems, we will use a local distortion technique. This technique is a time dependent modification of the local distortion technique of [12] (see also [16, 17] for related techniques). Although we do not yet know how to do it, we believe that such local distortions can be used to study the time independent Born–Oppenheimer expansion for polyatomic Coulomb systems.

The paper is organized as follows: In Sect. 2, notation is developed, and the main theorem is stated. The asymptotic expansion is formally derived in Sect. 3. In Sect. 4 we sketch a proof of the main theorem that uses the calculations of Sect. 3. We only sketch the proof because it is virtually identical to the proof in [9]. Our main new ideas and techniques are in Sect. 3.

2. Notation and Results

In this section we develop notation and state our theorem. The notation is similar to that of [9].

The Hamiltonian for a molecular system with Coulomb potentials is

$$H(\varepsilon) = H_0(\varepsilon) + h$$

on $L^2(\mathbb{R}^{3N})$, where

$$H_0(\varepsilon) = -\sum_{j=1}^{K} \frac{\varepsilon^4}{2M_j} \Delta_j,$$

and

$$h = -\sum_{j=K+1}^{N} \frac{1}{2m_j} \Delta_j + \sum_{i<j} \frac{Z_i Z_j}{|x_i - x_j|}.$$

Here, $x_j \in \mathbb{R}^3$ denotes the position of the j^{th} particle; Δ_j denotes the Laplacian in x_j; the mass of the j^{th} nucleus is $\varepsilon^{-4} M_j$ (for $1 \leqq j \leqq K$); the mass of the j^{th} electron is m_j (for $K+1 \leqq j \leqq N$); and Z_j is the charge on the j^{th} particle. The operator $H_0(\varepsilon)$ represents the nuclear kinetic energy, and h represents the electron Hamiltonian plus the internuclear potential energy. *To simplify some notation, we will assume that* $M_j = 1$ *for* $1 \leqq j \leqq K$.

We let $R = (x_1, \ldots, x_K) = (R_1, \ldots, R_{3K})$, and $r = (x_{K+1}, \ldots, x_N) = (R_{3K+1}, \ldots, R_{3N})$. The operator h decomposes under the direct integral decomposition

$$L^2(\mathbb{R}^{3N}) = \int_{\mathbb{R}^{3K}} \oplus L^2(\mathbb{R}^{3(N-K)}) dR,$$

and we let $h(R)$ denote the fiber of h; i.e., given any $f(R, r) \in L^2(\mathbb{R}^{3N})$, we have $[hf](R, r) = [h(R) f(R, \cdot)](r)$.

We make two basic assumptions:

Hypothesis I. The nuclear charges are all positive, i.e., $Z_j > 0$ for $1 \leqq j \leqq K$.

Hypothesis II. There exists an open set $\mathscr{U} \subset \mathbb{R}^{3K}$, such that $R \in \mathscr{U}$ implies that $h(R)$ has a non-degenerate discrete eigenvalue $E(R)$ that depends continuously on R.

Remarks. 1. Hypothesis *I* is a stronger condition than we actually need, but without it, we must restrict the classical trajectories $a(t)$ that are defined below. The restriction is that $a(t)$ cannot describe a nuclear configuration in which two nuclei have coincident positions in $\mathbb{R}^3$.

2. In the physically interesting case, Hypothesis *I* is satisfied, and Hypothesis *II* is frequently satisfied. The electron masses are all equal to 1, and the electron charges are all equal to -1.

3. If $h(R)$ has any discrete eigenvalues, then the lowest one is non-degenerate. Our assumptions guarantee that all discrete eigenvalues of $h(R)$ are analytic in R as long as R is real and does not describe a nuclear configuration in which two nuclei coincide [12].

It follows from Hypothesis *II* that the eigenvalue $E(R)$ has an eigenfunction, $\varphi(R, \cdot)$, that is real, normalized, and continuous in R, for $R \in \mathscr{U}$. For $R \notin \mathscr{U}$, we set $\varphi(R, r) = 0$. We fix this choice permanently, and note that $\varphi(R, \cdot)$ and the orthogonal projection $P_{\varphi(R)}$ onto the subspace generated by $\varphi(R, \cdot)$ are actually C^2 functions of R for $R \in \mathscr{U}$ (see [4]). We set $P_{\varphi(R)}$ equal to zero for $R \notin \mathscr{U}$, and let $Q_{\varphi(R)} = (1 - P_{\varphi(R)})$.

Our goal is to study certain solutions to the Schrödinger equation $i(\partial \Psi / \partial s) = H(\varepsilon) \Psi$ for times s in an interval $[0, \varepsilon^{-2} T]$, and with initial velocities of the nuclei on the order of ε^2. For convenience we replace the variable s by $t = \varepsilon^2 s$, and study the equation $i\varepsilon^2(\partial \Psi / \partial t) = H(\varepsilon) \Psi$ for times t in the fixed interval $[0, T]$, and with initial velocities (with respect to t) of the nuclei fixed. In this limit, classical trajectories of the nuclei are fixed, and the initial kinetic energy of the nuclei is fixed.

In the t variables, the $\varepsilon \to 0$ limit is a semiclassical limit for the heavy particles, so we will use the semiclassical techniques that have been developed for this purpose [5–7], with $\hbar$ replaced by ε^2. In particular, we will make use of the special wave functions $\phi_k(A, B, \varepsilon^2, a, \eta, x)$ which are defined in [6, 7]. The definition is reproduced below. It requires a substantial amount of notation, which can also be found in [9].

To simplify some expressions, we set $n = 3K$. A multi-index $k = (k_1, k_2, \ldots, k_n)$ is an ordered n-tuple of non-negative integers. The order of k is defined to be $|k| = \sum_{i=1}^{n} k_i$, and the factorial of k is defined to be $k! = (k_1!)(k_2!) \cdots (k_n!)$. The symbol D^k denotes the differential operator

$$D^k = \frac{\partial^{|k|}}{(\partial x_1)^{k_1} (\partial x_2)^{k_2} \cdots (\partial x_n)^{k_n}},$$

and the symbol x^k denotes the monomial $x^k = x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$. We denote the gradient of a function f by $f^{(1)}$, and we denote the matrix of second partial derivatives of f by $f^{(2)}$. With a slight abuse of notation, we view $\mathbb{R}^n$ as a subset of $\mathbb{C}^n$, and let e_i denote the i^{th} standard basis vector in $\mathbb{R}^n$ or $\mathbb{C}^n$. The inner product on $\mathbb{R}^n$ or $\mathbb{C}^n$ is $\langle v, w \rangle = \sum_{i=1}^{n} \bar{v}_i w_i$.

We generalize the zeroth and first order Hermite polynomials by

$$\tilde{\mathscr{H}}_0(x) = 1$$

and

$$\tilde{\mathscr{H}}_1(v; x) = 2\langle v, x\rangle,$$

where v is an arbitrary non-zero vector in $\mathbb{C}^n$. Our generalizations of the higher order Hermite polynomials are defined recursively as follows: Let $v_1, v_2, \ldots, v_m$ be m arbitrary non-zero vectors in $\mathbb{C}^n$. Then,

$$\tilde{\mathscr{H}}_m(v_1, v_2, \ldots, v_m; x) = 2\langle v_m, x\rangle \tilde{\mathscr{H}}_{m-1}(v_1, v_2, \ldots, v_{m-1}; x) - 2\sum_{i=1}^{m-1} \langle v_m, \bar{v}_i\rangle \tilde{\mathscr{H}}_{n-2}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{m-1}; x).$$

One can prove inductively that these functions do not depend on the ordering of the vectors $v_1, v_2, \ldots, v_m$. Furthermore, if the space dimension is $n = 1$ and the vectors $v_1, v_2, \ldots, v_m$ are all equal to $1 \in \mathbb{C}^1$, then $\tilde{\mathscr{H}}_m(v_1, v_2, \ldots, v_m; x)$ is equal to the usual Hermite polynomial of order m, $H_m(x)$.

Now suppose A is a complex invertible $n \times n$ matrix. We define $|A| = [AA^*]^{1/2}$, where A^* denotes the adjoint of A. By the polar decomposition theorem, there exists a unique unitary matrix U_A so that $A = |A| U_A$. Given a multi-index k, we define the polynomial

$$\mathscr{H}_k(A; x) = \tilde{\mathscr{H}}_{|k|}(\underbrace{U_A e_1, \ldots, U_A e_1}_{k_1 \text{ entries}}, \underbrace{U_A e_2, \ldots, U_A e_2}_{k_2 \text{ entries}}, \ldots, \underbrace{U_A e_n, \ldots, U_A e_n}_{k_n \text{ entries}}; x).$$

We are now in a position to define the functions $\phi_k(A, B, \varepsilon^2, a, \eta, x)$.

Definition. Let A and B be complex $n \times n$ matrices with the following properties:

A and B are invertible; (2.1)

BA^{-1} is symmetric ([real symmetric] + i[real symmetric]); (2.2)

$\operatorname{Re} BA^{-1} = \frac{1}{2}[(BA^{-1}) + (BA^{-1})^*]$ is strictly positive definite; (2.3)

$$(\operatorname{Re} BA^{-1})^{-1} = AA^*. \tag{2.4}$$

Let $a \in \mathbb{R}^n, \eta \in \mathbb{R}^n, \varepsilon > 0$. Then for each multi-index k, we define

$$\phi_k(A, B, \varepsilon^2, a, \eta, x) = 2^{-|k|/2}(k!)^{-1/2}\pi^{-n/4}\varepsilon^{-n/2}[\det A]^{-1/2} \cdot \mathscr{H}_k(A; \varepsilon^{-1}|A|^{-1}(x-a)) \cdot \exp\{-\langle (x-a), BA^{-1}(x-a)\rangle/2\varepsilon^2 + i\langle \eta, (x-a)\rangle/\varepsilon^2\}.$$

The choice of the branch of the square root of $[\det A]^{-1}$ in this definition will depend on the context, and will always be specified.

Let us now impose Hypotheses *I* and *II*. The semiclassical wavefunction of the nuclei will be localized near a classical trajectory. The potential which determines this trajectory is the eigenvalue $E(R)$ of $h(R)$. Let a_0 be an element of the open set $\mathscr{U}$, and let η_0 be an arbitrary real $3K$ dimensional vector. By standard theorems on ordinary differential equations, there is an interval $[0, T]$ of non-zero length, such that there exists a unique solution $(a(t), \eta(t))$ to the system of ordinary

differential equations

$$\frac{\partial a}{\partial t}(t) = \eta(t), \tag{2.5}$$

$$\frac{\partial \eta}{\partial t}(t) = -E^{(1)}(a(t)), \tag{2.6}$$

subject to the conditions $a(0) = a_0, \eta(0) = \eta_0$, and $a(t) \in \mathscr{U}$ for $0 \leqq t \leqq T$. We define the action $S(t)$ along the classical path $(a(t), \eta(t))$ by

$$S(t) = \int_0^t \left(\frac{(\eta(s))^2}{2} - E(a(s)) \right) ds. \tag{2.7}$$

Next, we choose $3K \times 3K$ matrices A_0 and B_0 which satisfy conditions (2.1)–(2.4). Again, standard theorems on ordinary differential equations guarantee the existence of a solution $(A(t), B(t))$ to the system

$$\frac{\partial A}{\partial t}(t) = iB(t), \tag{2.8}$$

$$\frac{\partial B}{\partial t}(t) = iE^{(2)}(a(t))A(t), \tag{2.9}$$

subject to the initial conditions $A(0) = A_0$ and $B(0) = B_0$. As in [5–9], $A(t)$ and $B(t)$ are determined by the relations

$$A(t) = \frac{\partial a(t)}{\partial a(0)} A(0) + i \frac{\partial a(t)}{\partial \eta(0)} B(0),$$

$$B(t) = \frac{\partial \eta(t)}{\partial \eta(0)} B(0) - i \frac{\partial \eta(t)}{\partial a(0)} A(0).$$

For technical reasons and because $\mathscr{U}$ may be a proper subset of $\mathbb{R}^{3K}$, we must insert some cut-offs. In Sect. 3 we will choose open sets $\mathscr{U}_1$ and $\mathscr{U}_2$ in $\mathbb{R}^{3K}$ with compact closures $\bar{\mathscr{U}}_1$ and $\bar{\mathscr{U}}_2$, so that we have $a(t) \in \mathscr{U}_1 \subset \bar{\mathscr{U}}_1 \subset \mathscr{U}_2 \subset \bar{\mathscr{U}}_2 \subset \mathscr{U}$, for all $t \in [0, T]$. Since $\mathscr{U}_2$ is open and $\bar{\mathscr{U}}_1$ is compact, there exists a C^∞ function $F: \mathbb{R}^{3K} \to [0, 1]$, such that $F(R) = 0$ for $R \notin \mathscr{U}_2$ and $F(R) = 1$ for $R \in \bar{\mathscr{U}}_1$.

The main result of this paper is the following:

Theorem 2.1. *Let $H(\varepsilon)$ satisfy Hypotheses I and II. Assume $\varphi(R, r), a(t), \eta(t)$, etc. are defined as above. Choose $T > 0$, a positive integer α, and complex numbers d_k for $|k| \leqq \alpha$. Then, for $0 < \varepsilon \leqq 1, 0 \leqq t \leqq T$, and arbitrary J, there exist functions $\Psi_{j,\varepsilon}(R, r, t)$, $0 \leqq j \leqq J$, with the following properties:*

(1) $\Psi_{0,\varepsilon}(R, r, t) = e^{iS(t)/\varepsilon^2} F(R) \varphi(R, r) \cdot \sum_{|k| \leqq \alpha} d_k \phi_k(A(t), B(t), \varepsilon^2, a(t), \eta(t), R)$.

(2) *For each j, $\| \Psi_{j,\varepsilon}(R, r, t) \|$ is bounded as $\varepsilon \searrow 0$ for $0 \leqq t \leqq T$.*

(3) *In the variable R, the function $\Psi_{j,\varepsilon}(R, r, t)$ is localized near the classical path $a(t)$*

in the sense that there exist constants M, N_j, b_j, *such that*

$$\int |\Psi_{j,\varepsilon}(R,r,t)|^2 dr \leqq b_j\left(1+\left[\frac{|R-a(t)|}{\varepsilon}\right]^{N_j}\right)\exp\{-M|R-a(t)|^2/\varepsilon^2\},$$

for $0<\varepsilon\leqq 1$ *and* $0\leqq t\leqq T$.

(4) *The function* $\sum_{j=0}^{J}\Psi_{j,\varepsilon}(R,r,t)\varepsilon^j$ *is a* J^{th} *order asymptotic approximation to the solution of the Schrödinger equation in the sense that*

$$\left\| e^{-itH(\varepsilon)/\varepsilon^2}\sum_{j=0}^{J}\Psi_{j,\varepsilon}(R,r,0)\varepsilon^j - \sum_{j=0}^{J}\Psi_{j,\varepsilon}(R,r,t)\varepsilon^j\right\| \leqq C_J\varepsilon^{J+1},$$

for $0\leqq\varepsilon\leqq 1$, *and* $0\leqq t\leqq T$.

Remarks. 1. The first order term, $\Psi_{1,\varepsilon}(R,r,t)$, is only a semiclassical correction term. It contains no adiabatic corrections in the sense that its r dependence has no components orthogonal to $\varphi(R,r)$ in $L^2(dr)$ for fixed R and t. In particular,

$$\Psi_{1,\varepsilon}(R,r,t)=e^{iS(t)/\varepsilon^2}F(R)\varphi(R,r)\sum_{|k|\leqq\alpha+3}d_k(t)\phi(A(t),B(t),\varepsilon^2,a(t),\eta(t),R),$$

where the functions $d_k(t)$ solve a system of differential equations. In Sect. 3 we will describe how one can derive this system.

2. A procedure for computing the functions $\Psi_{j,\varepsilon}(R,r,t)$ is given in Sect. 3. Since these computations parallel those of [9], we have not given many details. The reader who is interested in explicit computation should consult [7, 9] and note the erratum to [9]. For details of the semiclassical techniques in particular, he should see Theorem 1.1 of [7] and the remarks which follow that theorem.

3. Formal Derivation of the Expansion

In this section we give a formal derivation of the results that were stated in Sect. 2. The main ideas involved are the same as in the smooth potential case [9]. However, because $\varphi(R,r)$ is only a C^2 function of R in the Coulomb case, we must make significant alterations. The alterations involve complicated nonlinear changes of variables before the application of the multiple scales and semiclassical ideas of [9].

As described in the previous sections, we wish to study the small ε asymptotics of the solutions to the equation

$$i\varepsilon^2\frac{\partial\Psi}{\partial t}=-\frac{\varepsilon^4}{2}\Delta\Psi+h\Psi, \tag{3.1}$$

where Δ is the $3K$ dimensional Laplacian in the variable R, and h decomposes under the direct integral decomposition with fibers $h(R)$.

Although the eigenfunction $\varphi(R,r)$ does not have third order partial derivatives with respect to the R variables in $L^2(dr)$, it is analytic in certain directions in the configuration space of all the particles [12]. We change variables in $(R,r,t)\in\mathbb{R}^{3N+1}$, so that the electronic wave function is C^∞ in $3K+1$ of the $3N+1$ new coordinate directions. We do so in two steps. In the first step we change coordinates

in a way that solves the difficulty in the R variables, but transfers part of the non-differentiability into the t variable. The motivation for this step comes from [12]. In the second step we change variables to fix up the resulting non-differentiability in t. After the change of coordinates, we mimic the techniques of [9] to derive the expansion.

3A. The First Change of Variables. Our first change of variables is done by mimicking [12] for each fixed time t. In the new coordinates, the Coulomb singularity is at the nuclear configuration $a(t)$ rather than R.

In the case of smooth potentials, the wave functions of interest are highly concentrated where $|R - a(t)|$ is small. A posteriori, the same is true in the Coulomb case. So, we will not change variables on all of $\mathbb{R}^{3N} \times [0, T]$, but instead will work on a set $W \subset \mathbb{R}^{3N} \times [0, T]$. Outside this set, our wave functions will vanish.

We take W to be the set

$$W = \bigcup_{t \in [0,T]} (\mathscr{W}(t) \times \mathbb{R}^{3(N-K)} \times \{t\}),$$

where $\mathscr{W}(t) = \{(R_1, \ldots, R_{3K}) \in \mathbb{R}^{3K}: |R_j - a_j(t)| \leqq \gamma \text{ for } 1 \leqq j \leqq 3K\}$. Here γ is a number that will be chosen later. We also choose the sets $\mathscr{U}_1$ and $\mathscr{U}_2$ of Sect. 2 to be

$$\mathscr{U}_1 = \{(R_1, \ldots, R_{3K}): |R - a(t)| < \gamma/2\},$$

and

$$\mathscr{U}_2 = \{(R_1, \ldots, R_{3K}): |R - a(t)| < \gamma\}.$$

We define $b_j(t) \in \mathbb{R}^3$ for $1 \leqq j \leqq K$ by

$$b_j(t) = (a_{3j-2}(t), a_{3j-1}(t), a_{3j}(t)).$$

We permanently fix a choice of $\delta > 0$, so that $\delta < \frac{1}{4} \min\{|b_i(t) - b_j(t)|: 1 \leqq i < j \leqq K, 0 \leqq t \leqq T\}$, and we restrict $\gamma < \delta/2$. We choose a function $f: [0, \infty) \to [0, 1]$, so that $f \in C^\infty$, $f(x) = 0$ if $x > 2\delta$, and $f(x) = 1$ if $x < \delta$. We define a vector field $v_{R,t}: \mathbb{R}^3 \to \mathbb{R}^3$ by

$$v_{R,t}(\zeta) = \sum_{j=1}^{K} [b_j(t) - x_j] f(|\zeta - b_j(t)|).$$

One can easily prove [12] that by shrinking γ, the maximum of the norm of the derivative matrix $\partial v_{R,t}/\partial \zeta$ can be made arbitrarily small for all $R \in \mathscr{W}(t)$, $t \in [0, T]$, and $\zeta \in \mathbb{R}^3$. As a consequence [12], for small enough γ, the mapping $u_{R,t}: \mathbb{R}^3 \to \mathbb{R}^3$ given by $u_{R,t}(\zeta) = \zeta + v_{R,t}(\zeta)$ is a C^∞ diffeomorphism of $\mathbb{R}^3$. It is also a C^∞ function of all the variables R, ζ, and t.

We let $J_1(R, r, t) = \prod_{i=K+1}^{N} \det(I + [(\partial v_{R,t}/\partial \zeta)(x_i)])$, and define a fibered unitary operator $U_1(t)$ on $L^2(\mathbb{R}^{3N})$ by the following two conditions: We define $U_1(t)\Omega = \Omega$ for functions Ω whose support does not intersect the set $\mathscr{W}(t) \times \mathbb{R}^{3(N-K)}$. For functions Ω whose supports are contained in the set $\mathscr{W}(t) \times \mathbb{R}^{3(N-K)}$, we define

$$[U_1(R, t)\Omega](R, r) = [J_1(R, r, t)]^{1/2} \Omega(R, u_{R,t}(x_{K+1}), \ldots, u_{R,t}(x_N)).$$

By linearity, this uniquely determines $U_1(t)$, which unitarily implements our first change of variables: If (1) γ is sufficiently small; (2) Ψ is an approximate solution

to Eq. (3.1) in the sense that

$$i\varepsilon^2 \frac{\partial \Psi}{\partial t} = -\frac{\varepsilon^4}{2}\Delta \Psi + h\Psi + O(\varepsilon^n);$$

and (3) Ψ has support in W, then $\Omega = U_1(t)^{-1}\Psi$ is supported in W, and satisfies an equation of the form

$$i\varepsilon^2\left(\frac{\partial}{\partial t} + \mathscr{A}_1\right)\Omega = -\frac{\varepsilon^2}{2}(\Delta_R + \mathscr{B}_1)\Omega + h_1\Omega + O(\varepsilon^n). \tag{3.2}$$

Here, $\mathscr{A}_1$ is a first order differential operator in r, whose coefficients are C^∞ functions of compact support in the variables $(R, r, t) \in W$. For fixed (r, t), the coefficients are analytic in R, for (R, r, t) in W. $\mathscr{B}_1$ is a second order differential operator in R and r, with coefficients that are C^∞ functions of compact support in the variables $(R, r, t) \in W$. For fixed (r, t), its coefficients are analytic in R, for (R, r, t) in W. The fibered operator h_1 has the form

$$h_1 = -\sum_{j=K+1}^{N} \frac{1}{2m_j}\Delta_j + \mathscr{C}_1 + V_1.$$

Here $\mathscr{C}_1$ is a second order differential operator in the r variables with C^∞ coefficients of compact support in the variables $(R, r, t) \in W$. For fixed (r, t), its coefficients are analytic in R, for (R, r, t) in W. It is relatively bounded with respect to the electronic kinetic energy, with relative bound less than one. The potential term V_1 has three parts. The nuclear–nuclear repulsion is the same as in the original potential V. The electron–electron potential is relatively bounded with respect to the electronic kinetic energy. The electron–nuclear interaction has the form

$$\sum_{j=1}^{K}\sum_{i=K+1}^{N} [Z_i Z_j + p_{i,j}(R, x_i, t)]\frac{1}{|x_i - b_j(t)|},$$

where each $p_{i,j}$ is C^∞ for $(R, r, t) \in W$. The function $p_{i,j}$ is analytic in R for $R \in \mathscr{W}(t)$, and $p_{i,j}$ and all of its derivatives vanish in the limit as x_i goes to infinity in $\mathbb{R}^3$, uniformly in R for $R \in \mathscr{W}(t)$ and $t \in [0, T]$.

All of these facts concerning equation (3.2) follow from direct computation and simple extensions of the ideas of [12]. One should note that the coefficients of equation (3.2) have no singularities in the R variables for $(R, r, t) \in W$. In fact, the fibers $h_1(R, t)$ on $L^2(\mathbb{R}^{3(N-K)})$ form an analytic family in the R variables for $R \in \mathscr{W}(t)$ if γ is sufficiently small. However, the electron–nuclear interaction has time dependence with moving Coulomb centers in the new variables. So, in the new variables, the electron wave functions for nondegenerate eigenvalues are only C^2 functions of t.

We now permanently fix a choice of the number γ, small enough for all the above conditions to hold.

3B. The Second Change of Variables. To determine the next change of variables, we define a vector field $X(\zeta, s)$ whose associated flow near a nucleus is parallel to the motion of that nucleus. The flow will then determine the change of coordinates.

The goal is to obtain a coordinate system in which the centers of the electron–nuclear Coulomb potentials do not move in time.

We define the vector field to be

$$X(\zeta, s) = \sum_{j=1}^{K} \xi_j(s) f(|\zeta - b_j(s)|),$$

where $\xi_j(t) = (\eta_{3j-2}(t), \eta_{3j-1}(t), \eta_{3j}(t))$. Given any fixed $t \in [0, T]$, we solve the differential equation

$$\frac{\partial \zeta}{\partial s} = X(\zeta, s)$$

backwards in time, from time t to time 0. If the initial condition at time t is $\zeta(t) = x \in \mathbb{R}^3$, then we define $w(x, t) = \zeta(0)$. By standard theorems of ordinary differential equations (see e.g., [13, Satz 4, p. 125]), the mapping that carries x to $w(x, t)$ exists, is unique, and is C^∞. It is also a diffeomorphism because one can construct the smooth inverse by flowing forward in time rather than backwards.

We let

$$J_2(r, t) = \prod_{i=K+1}^{N} \det\left[\frac{\partial w}{\partial x}(x_i, t)\right],$$

and define a unitary operator $U_2(t)$ by

$$[U_2(t)\Phi](R, r) = [J_2(r, t)]^{1/2} \Phi(R, w(x_{K+1}, t), \ldots, w(x_N, t)).$$

This operator unitarily implements our second change of variables.

If $\Omega(R, r, t)$ is a solution to Eq. (3.2) with support in W, then $\Xi = U_2(t)^{-1}\Omega$ is supported in W, and satisfies an equation of the form

$$i\varepsilon^2\left(\frac{\partial}{\partial t} + \mathscr{A}_2\right)\Xi = -\frac{\varepsilon^4}{2}(\Delta_R + \mathscr{B}_2)\Xi + h_2\Xi + O(\varepsilon^n). \tag{3.3}$$

The differential operators $\mathscr{A}_2$ and $\mathscr{B}_2$ have the properties enumerated above for $\mathscr{A}_1$ and $\mathscr{B}_1$, respectively. The fibered operator h_2 has the form

$$h_2 = -\sum_{j=K+1}^{N} \frac{1}{2m_j}\Delta_j + \mathscr{C}_2 + V_2.$$

Since $U_2(t)$ has no R dependence, and $h_1(R, t)$ is an analytic family in R for $R \in \mathscr{W}(t)$, $h_2(R, t) = U_2(t) h_1(R, t) U_2(t)^{-1}$ is also an analytic family in the R variables for $R \in \mathscr{W}(t)$. In the t variable, $h_2(R, t)$ is a C^∞ family in the sense that the resolvent $(z - h_2(R, t))^{-1}$ acting on $L^2(\mathbb{R}^{3(N-K)})$ is an infinitely differentiable function of t. To see this, one computes the derivatives explicitly. For example, the first derivative is

$$(z - h_2(R, t))^{-1} \frac{\partial h_2}{\partial t}(R, t)(z - h_2(R, t))^{-1}.$$

The operator $(\partial h_2/\partial t)(R, t)$ is a second order differential operator in r. With the exception of the electron–nuclear interaction terms, the coefficients of $(\partial h_2/\partial t)(R, t)$ are C^∞ functions of (R, r, t) with compact support in r. Thus, the terms of

$(\partial h_2/\partial t)(R,t)$, other than the electron–nuclear terms, are bounded operators from the range of $(z-h_2(R,t))^{-1}$ (with the usual Sobolev norm) to $L^2(\mathbb{R}^{3(N-K)})$, with smooth dependence on t. The electron–nuclear interaction terms give rise to bounded operators on these spaces with smooth time dependence, but this is less obvious. These terms in the potential V_2 have the form

$$\sum_{j=1}^{K}\sum_{i=K+1}^{N}[Z_iZ_j+q_{i,j}(R,x_i,t)]\frac{1}{|x_i-b_j(0)|},$$

where each $q_{i,j}$ is C^∞ for $(R,r,t)\in W$. The function $q_{i,j}$ is analytic in R for $R\in\mathscr{W}(t)$, and all of its derivatives vanish in the limit as x_i goes to infinity in $\mathbb{R}^3$, uniformly in R for $R\in\mathscr{W}(t)$ and $t\in[0,T]$. Since the only t dependence of the electron–nuclear terms in V_2 is in the $q_{i,j}$'s, it follows that the electron–nuclear interaction terms in $(\partial h_2/\partial t)(R,t)$ are bounded operators from the range of $(z-h_2(R,t))^{-1}$ to $L^2(\mathbb{R}^{3(N-K)})$, with smooth dependence on t. The higher order derivatives of the resolvent of $h_2(R,t)$ are dealt with in a similar manner.

Thus, the operators $h_2(R,t)$ are C^∞ functions of (R,t) in the resolvent sense described above for $R\in\mathscr{W}(t)$ and $t\in[0,T]$. From this, it follows that the eigenfunction $\varphi_2(R,r,t)=U_2(t)^{-1}U_1(t)^{-1}\varphi(R,r)$ of $h_2(R,t)$ is an analytic function of $R\in\mathscr{W}(t)$ and a C^∞ function of $t\in[0,T]$, with values in $L^2(\mathbb{R}^{3(N-K)},dr)$.

3C. The Multiple Scales Expansion. In the remainder of this section of the paper, we apply the ideas of Sect. 3 of [9] to Eq. (3.3).

Instead of directly seeking a solution $\Xi(R,r,t)$ to (3.3), we will seek a solution $\Phi(x,y,r,t)$ to a higher dimensional problem. We will then recover $\Xi(R,r,t)=\Phi(R,(R-a(t))/\varepsilon,r,t)$. This is the technique of the "multiple scales." It is useful for our problem because as $\varepsilon\to 0$, the variables $x=R$ and $y=(R-a(t))/\varepsilon$ behave in a more or less independent fashion. Moreover, semiclassical effects occur in the variable y, and adiabatic effects occur in the variable x. Without the independent treatments of these variables, the two types of effects become intertwined and much more difficult to understand.

To obtain the equation that is satisfied by $\Phi(x,y,r,t)$, we make a formal change of variables in Eq. (3.3) from (R,r,t) to (x,y,r,t), with $x=R$ and $y=(R-a(t))/\varepsilon$. We also make a careful choice of when to replace the variable R by x or $[a(t)+\varepsilon y]$. In particular, we always replace R by x in the coefficients of the equation, but we add on $[E(a(t)+\varepsilon y)-E(x)]\Phi(x,y,r,t)$. The equation satisfied by $\Phi(x,y,r,t)$ has the form:

$$i\varepsilon^2\left(\frac{\partial}{\partial t}+\mathscr{A}\right)\Phi=-\frac{\varepsilon^4}{2}\mathscr{F}\Phi-\varepsilon^3\mathscr{G}\cdot\nabla_y\Phi-\frac{\varepsilon^2}{2}\Delta_y\Phi+i\varepsilon\eta(t)\cdot\nabla_y\Phi+E(a(t)+\varepsilon y)\Phi$$
$$+h_2(x,t)\Phi-E(x)\Phi+O(\varepsilon^n). \tag{3.4}$$

Here, $\mathscr{A}$ is a first order differential operator in the r variables; $\mathscr{F}$ is a second order differential operator in x and r; and $\mathscr{G}$ is a first order operator in x and r. The coefficients of $\mathscr{A}$, $\mathscr{F}$, and $\mathscr{G}$ are C^∞ functions of compact support in $(x,r,t)\in W$. The operator $h_2(x,t)$ is analytic in x and is C^∞ in t in the resolvent sense described in Sect. 3B.

For later reference, we note that $\mathscr{G} = \nabla +$ (first order operator in r only), and that $\mathscr{F} = \mathscr{G} \cdot \mathscr{G}$.

If the $O(\varepsilon^n)$ in Eq. (3.4) is understood as an $L^\infty(dx\,dt)$-norm estimate of a function with values in $L^2(dy\,dr)$, and the $O(\varepsilon^n)$ in Eq. (3.3) is understood as an $L^\infty(dt)$-norm estimate of a function with values in $L^2(dR\,dr)$, then any solution $\Phi(x, y, r, t)$ to Eq. (3.4) gives rise to a solution $\Xi(R, r, t) = \Phi(R, (R - a(t))/\varepsilon, r, t)$ to Eq. (3.3). Then $\Psi(R, r, t) = U_1(t) U_2(t) \Xi(R, r, t)$ is an approximate solution to Eq. (3.1). In Sect. 4 we will rigorously prove this for the solutions Φ to Eq. (3.4) that we construct below.

To formally derive our solutions to Eq. (3.4), we begin by making the ansatz that Eq. (3.4) possesses an approximate solution of the form

$$\Phi(x, y, r, t) = e^{iS(t)/\varepsilon^2} e^{i\eta(t) \cdot y/\varepsilon} F(x) \cdot (\lambda_0(x, y, r, t) + \varepsilon \lambda_1(x, y, r, t) + \cdots), \tag{3.5}$$

where F is the cutoff function of Sect. 2 which is determined uniquely by the choice of γ at the end of Sect. 3A.

We define $\psi_n(x, y, r, t) = U_1(x, t) U_2(t) \lambda_n(x, y, r, t)$. In certain circumstances, it is pedagogically advantageous to consider the functions ψ_n rather than deal directly with the functions λ_n.

Remark. As in [9], the factor $F(x)$ plays an uninteresting role, but it causes an extra complication in the formal computations. In particular, certain terms which occur below contain derivatives of F. As in [9], these terms are basically irrelevant. Whenever we encounter one of these terms in this section, we will make a comment, drop the term, and refer the reader to Sect. 4 for the explanation of why the term can be dropped. Heuristically, these terms can be dropped because derivatives of $F(R)$ are supported in a region of configuration space where the wave function is exponentially small in ε as $\varepsilon \to 0$.

We now need to determine the functions λ_n. Since $F(x) = 0$ for $x \notin \mathscr{U}_2$, we can arbitrarily set $\lambda_n(x, y, r, t) = 0$ for all $x \notin \mathscr{U}_2$. To determine these functions for $x \in \mathscr{U}_2$ we substitute the expression (3.5) into Eq. (3.4) and expand the factor $E(a(t) + \varepsilon y)$ in its Taylor series in powers of ε. Then we multiply everything out and equate coefficients of like powers of ε on the two sides of the equation. Since we are assuming that Eqs. (2.5)–(2.8) are satisfied, a large amount of cancellation takes place. After making these cancellations, we find that various conditions are forced upon us by the terms of different orders in ε. These conditions and the initial conditions more or less completely determine the functions λ_n.

After using Eq. (2.7) to make a cancellation, we see that the zeroth order terms force us to take

$$F(x)[h_2(x, t) - E(x)] \lambda_0 = 0.$$

Thus, for $x \in \mathscr{U}_2$, we are forced to take λ_0 to be some function $g_0(x, y, t)$ times $\varphi_2(x, r, t)$. Without loss of generality, we can choose g_0 to be independent of x because any x dependence can be absorbed into y dependence in the higher order terms. Thus, we have

$$\lambda_0(x, y, r, t) = g_0(y, t) \varphi_2(x, r, t).$$

Equivalently, we can state this in terms of ψ_0 as

$$\psi_0(x, y, r, t) = g_0(y, t)\varphi(x, r, t).$$

Similarly, after using both Eqs. (2.6) and (2.7) to make cancellations, we see that the first order terms force us to take

$$F(x)[h_2(x, t) - E(x)]\lambda_1 = 0.$$

Thus, for $x \in \mathscr{U}_2$, λ_1 has the form

$$\lambda_1(x, y, r, t) = g_1(x, y, t)\varphi_2(x, r, t),$$

for some function g_1. Equivalently, ψ_1 has the form

$$\psi_1(x, y, r, t) = g_1(x, y, t)\varphi(x, r, t).$$

The second order terms impose the following condition:

$$iF(x)(\dot{\lambda}_0 + \mathscr{A}\lambda_0) = F(x)\Big[-\tfrac{1}{2}\Delta_y\lambda_0 + E^{(2)}(a)\frac{y^2}{2}\lambda_0 - i\eta\cdot\mathscr{G}\lambda_0 + [h_2(x, t) - E(x)]\lambda_2 \Big] - i\eta\cdot(\nabla_x F)\lambda_0, \quad (3.6a)$$

where the dot $\dot{}$ stands for time derivative and $E^{(2)}(a)y^2/2$ is a shorthand abuse of notation which stands for

$$\frac{1}{2}\sum_{i,j} y_i y_j \frac{\partial^2 E}{\partial y_i \partial y_j}(a(t)).$$

As we remarked earlier, arguments from Sect. 4 show that the final term in Eq. (3.6a) is actually of finite order in ε in an appropriate norm. So, we may drop it to obtain the following condition for $x \in \mathscr{U}_2$:

$$i(\dot{\lambda}_0 + \mathscr{A}\lambda_0) = -\tfrac{1}{2}\Delta_y\lambda_0 + E^{(2)}(a)\frac{y^2}{2}\lambda_0 - i\eta\cdot\mathscr{G}\lambda_0 + [h_2(x, t) - E(x)]\lambda_2. \quad (3.6b)$$

We separate this equation into two parts. The first part involves those components on the two sides of the equation which are (x, y, and t dependent) multiples of $\varphi_2(x, r, t)$. The second part involves those terms which are orthogonal to $\varphi_2(x, r, t)$ in $L^2(dr)$ for each fixed x, y, and t. Thus, Eq. (3.6b) imposes two conditions:

$$iP_{\varphi_2}(\dot{\lambda}_0 + \mathscr{A}\lambda_0) = -\tfrac{1}{2}\Delta_y\lambda_0 + E^{(2)}(a)\frac{y^2}{2}\lambda_0 - iP_{\varphi_2}\eta\cdot\mathscr{G}\lambda_0 \quad (3.7a)$$

and

$$[h_2(x, t) - E(x)]\lambda_2 = iQ_{\varphi_2}\eta\cdot\mathscr{G}\lambda_0 + iQ_{\varphi_2}\dot{\lambda}_0 + iQ_{\varphi_2}\mathscr{A}\lambda_0. \quad (3.7b)$$

Here, $P_{\varphi_2} = P_{\varphi_2}(x, t)$ denotes the orthogonal projection in $L^2(dr)$ onto multiples of $\varphi_2(x, r, t)$ and $Q_{\varphi_2} = Q_{\varphi_2}(x, t) = (1 - P_{\varphi_2}(x, t))$ is the complementary orthogonal projection.

Because there are cancellations that are not obvious in Eqs. (3.7), it is instructive to consider the equations in the ψ representation that are equivalent to Eqs. (3.7). They are:

$$i\dot{\psi}_0 = -\tfrac{1}{2}\Delta_y\psi_0 + E^{(2)}(a)\frac{y^2}{2}\psi_0 \tag{3.8a}$$

and

$$[h(x) - E(x)]\psi_2 = i\eta\cdot\nabla_x\psi_0. \tag{3.8b}$$

Equation (3.8a) involves the x, y, and t dependent multiples of $\varphi(x,r)$, and Eq. (3.8b) involves functions that are orthogonal to $\varphi(x,r)$ in $L^2(dr)$. We note that Eqs. (3.8) are the same as Eqs. (3.5) of [9], and they can be solved as in [9]. To obtain the solutions in which we are interested, we arbitrarily impose certain initial conditions. Then ψ_0 is determined for $x\in\mathcal{U}_2$ to be

$$\psi_0 = \varphi(x,r)\sum_{|k|\leqq\alpha} d_k\phi_k(A(t), B(t), \varepsilon^2, 0, 0, y).$$

Equivalently, Eq. (3.7a) is solved by

$$\lambda_0 = \varphi_2(x,r,t)\sum_{|k|\leqq\alpha} d_k\phi_k(A(t), B(t), \varepsilon^2, 0, 0, y).$$

The numbers d_k are arbitrary, and the number α is arbitrary, but indicates that the sum is finite. The functions $\phi_k(A, B, \varepsilon^2, a, \eta, y)$ were defined precisely so that these expressions would satisfy Eq. (3.8a) and (3.7a). In view of the form of our ansatz (3.5), we see that conclusion (1) of Theorem 2.1 is fulfilled.

Equation (3.8b) does not uniquely determine ψ_2 because $[h(x) - E(x)]$ has a non-trivial kernel. However, the restriction of $[h(x) - E(x)]$ to the orthogonal complement of the multiples of $\varphi(x,r)$ in $L^2(dr)$ has an inverse $r(x)$. As a consequence, Eq. (3.5b) shows that $\psi_2 = \psi_2^{\perp} + \psi_2^{\parallel}$, where $\psi_2^{\parallel}$ is some function of x, y, and t times $\varphi(x,r)$, and

$$\psi_2^{\perp} = ir(x)\eta\cdot\nabla_x\psi_0, \tag{3.9a}$$

for $x\in\mathcal{U}_2$. Equivalently, we let $r_2(x)$ denote the inverse of the restriction of $[h_2(x,r,t) - E(x)]$ to the orthogonal complement of $\varphi_2(x,r,t)$ in $L^2(dr)$, and decompose λ_2 into a component $\lambda_2^{\perp}$ orthogonal to $\varphi_2(x,r,t)$ and a component $\lambda_2^{\parallel}$ in the $\varphi_2(R,r,t)$ direction. Then

$$\lambda_2^{\perp} = ir_2(x)Q_{\varphi_2}(\eta\cdot\mathcal{G}\lambda_0 + \dot{\lambda}_0 + \mathcal{A}\lambda_0). \tag{3.9b}$$

Thus, the second order terms in ε determine λ_0 and $\lambda_2^{\perp}$.

We now consider the third order terms. They impose the following condition:

$$\begin{aligned} iF(x)(\dot{\lambda}_1 + \mathcal{A}\lambda_1) = F(x)\Big[&-\tfrac{1}{2}\Delta_y\lambda_1 + E^{(2)}(a)\frac{y^2}{2}\lambda_1 + E^{(3)}(a)\frac{y^3}{6}\lambda_0 - i\eta\cdot\mathcal{G}\lambda_1 \\ &- \mathcal{G}\cdot\nabla_y\lambda_0 + [h_2(x,t) - E(x)]\lambda_3\Big] \\ &- i\eta\cdot(\nabla_x F)\lambda_1 - (\nabla_x F)\cdot(\nabla_y\lambda_0), \end{aligned} \tag{3.10a}$$

where we have again made the obvious abuse of notation. In Sect. 4 we will see that the last two terms in this equation can be dropped because they are of infinite order in ε. We drop them to obtain the following condition for $x \in \mathscr{U}_2$:

$$i(\dot{\lambda}_1 + \mathscr{A}\lambda_1) = -\tfrac{1}{2}\Delta_y \lambda_1 + E^{(2)}(a)\frac{y^2}{2}\lambda_1 + E^{(3)}(a)\frac{y^3}{6}\lambda_0$$
$$- i\eta\cdot\mathscr{G}\lambda_1 - \mathscr{G}\cdot\nabla_y\lambda_0 + [h_2(x,t) - E(x)]\lambda_3. \tag{3.10b}$$

As with the second order terms, we separate these terms into those which are multiples of $\varphi_2(x,r,t)$ and those which are orthogonal to $\varphi_2(x,r,t)$ in $L^2(dr)$. We obtain

$$iP_{\varphi_2}(\dot{\lambda}_1 + \mathscr{A}\lambda_1) = -\tfrac{1}{2}\Delta_y\lambda_1 + E^{(2)}(a)\frac{y^2}{2}\lambda_1 + E^{(3)}(a)\frac{y^3}{6}\lambda_0$$
$$- P_{\varphi_2}(i\eta\cdot\mathscr{G}\lambda_1 + \mathscr{G}\cdot\nabla_y\lambda_0), \tag{3.11a}$$

and

$$[h_2(x,t) - E(x)]\lambda_3 = Q_{\varphi_2}(i\eta\cdot\mathscr{G}\lambda_1 + \mathscr{G}\cdot\nabla_y\lambda_0 + \dot{\lambda}_1 + \mathscr{A}\lambda_1). \tag{3.11b}$$

The equivalent, but simpler equations in the ψ representation are

$$i\dot{\psi}_1 = -\tfrac{1}{2}\Delta_y\psi_1 + E^{(2)}(a)\frac{y^2}{2}\psi_1 + E^{(3)}(a)\frac{y^3}{6}\psi_0 - iP_\varphi\eta\cdot\nabla_x\psi_1, \tag{3.12a}$$

and

$$[h(x) - E(x)]\psi_3 = iQ_\varphi\eta\cdot\nabla_x\psi_1 + \nabla_y\cdot\nabla_x\psi_0. \tag{3.12b}$$

Equations (3.12) are the same as Eqs. (3.8) of [9], as corrected in the erratum [9].

With the initial condition $\psi_1(x,y,r,0) = 0$, Eq. (3.12a) can be fairly explicitly solved as described in [9] and its erratum [9]. Alternatively, one can use those techniques to directly solve Eq. (3.11a) with $\lambda_1(x,y,r,0) = 0$.

Remarks. 1. The solution to (3.12a) has the form $\psi_1(x,y,r,t) = g_1(x,y,t)\varphi(x,r)$, where g_1 turns out to have no x dependence. Also, $\lambda_1(x,y,r,t) = g_1(x,y,t)\varphi_2(x,r,t)$. The function g_1 is a time dependent linear combination of the functions $\phi_k(A(t), B(t), \varepsilon^2, 0, 0, y)$ with $|k| \leqq \alpha + 3$ (see [9]).

2. Since λ_1 is a y and t dependent multiple of $\varphi_2(x,r,t)$, $\lambda_0 + \varepsilon\lambda_1$ is only a semiclassical correction to $\lambda_{0,\varepsilon}$. The corrections to the adiabatic approximation first show up in λ_2.

To solve Eq. (3.12b), we let $\psi_3 = \psi_3^\perp + \psi_3^\parallel$, where $\psi_3^\parallel = g_3(x,y,t)\ \varphi(x,r)$ and

$$\psi_3^\perp = r(x)[iQ_\varphi\eta\cdot\nabla_x\psi_1 + \nabla_y\cdot\nabla_x\psi_0]. \tag{3.13a}$$

Equivalently, we have

$$\lambda_3^\perp = r_2(x,t)Q_{\varphi_2}[i\eta\cdot\mathscr{G}\lambda_1 + \mathscr{G}\cdot\nabla_y\lambda_0 + i\dot{\lambda}_1 + i\mathscr{A}\lambda_1]. \tag{3.13b}$$

We now turn to the fourth order terms. They impose

$$iF(x)(\dot{\lambda}_2 + \mathscr{A}\lambda_2) = F(x)\Big[-\tfrac{1}{2}\Delta_y\lambda_2 + E^{(2)}(a)\frac{y^2}{2}\lambda_2 + E^{(3)}(a)\frac{y^3}{6}\lambda_1$$

$$+ E^{(4)}(a)\frac{y^4}{24}\lambda_0 - i\eta\cdot\mathscr{G}\lambda_2 - \mathscr{G}\cdot\nabla_y\lambda_1$$

$$-\tfrac{1}{2}\mathscr{F}\lambda_0 + [h_2(x,t) - E(x)]\lambda_4\Big]$$

$$- i\eta\cdot(\nabla_x F)\lambda_2 - (\nabla_x F)\cdot(\nabla_y\lambda_1) - \tfrac{1}{2}(\Delta_x F)\lambda_0 - (\nabla_x F)\cdot(\mathscr{G}\lambda_0). \tag{3.14a}$$

We drop the last four terms on the right-hand side of this equation since they are of infinite order in ε. Then for $x\in\mathscr{U}_2$, we impose

$$i(\dot{\lambda}_2 + \mathscr{A}\lambda_2) = -\tfrac{1}{2}\Delta_y\lambda_2 + E^{(2)}(a)\frac{y^2}{2}\lambda_2 + E^{(3)}(a)\frac{y^3}{6}\lambda_1$$

$$+ E^{(4)}(a)\frac{y^4}{24}\lambda_0 - i\eta\cdot\mathscr{G}\lambda_2 - \mathscr{G}\cdot\nabla_y\lambda_1 - \tfrac{1}{2}\mathscr{F}\lambda_0 + [h_2(x,t) - E(x)]\lambda_4. \tag{3.14b}$$

We solve this equation by splitting it into components that are multiples of $\varphi_2(x,r,t)$ and components that are orthogonal to $\varphi_2(x,r,t)$. We can solve the resulting equations by the techniques of [9] and its erratum [9]. Once again, it is technically less complicated to work in the ψ representation rather than the λ representation. The initial condition $\lambda_2^{\parallel}(x,y,r,0) = \psi_2^{\parallel}(x,y,r,0) = 0$ and the fourth order terms thus determine $\psi_2^{\parallel}$ and $\psi_2^{\perp}$.

By the obvious induction, one determines the higher order terms. The n^{th} order terms of the equation and the condition $\lambda_2^{\parallel}(x,y,r,0) = 0$ determine $\lambda_{n-2}^{\parallel}$ and $\lambda_n^{\perp}$. Unfortunately, for $n \geqq 5$, one does not have the luxury of working in the ψ representation because of the lack of existence of third and higher order derivatives of $\varphi(x,r)$ with respect to x. Of course, this was the reason for introducing the λ representation in the first place.

Remark. If one is interested only in the 0^{th}, 1^{st}, or 2^{nd} order expansion for the solution Ψ to Eq. (3.1), one can get the correct result by calculating in the ψ representation with the formal equations of [9]. However, the proof in [9] does not apply to the Coulomb case at any order because $\varphi(x,r)$ is only C^2 in x. Without going to the λ representation, we do not know how to construct arbitrarily high order terms or prove that the expansion is asymptotic to the solution of the Schrödinger equation in a norm sense.

Conclusions (1) and (2) of Theorem 2.1 are now clear upon substitution of R for x and $(R - a(t))/\varepsilon$ for y in the formulas for $\lambda_{n,\varepsilon}$, and then computing $\psi_n = U_1(t)U_2(t)\lambda_n$. A proof of Conclusion (4) will be presented in the next section. Conclusion (3) follows from the fact that the y dependence of λ_n exhibits itself only through the functions $\phi_k(A,B,\varepsilon^2,0,0,y)$, which are polynomials in y times

Gaussians in y. The x dependence involves smooth functions which are bounded for $x \in \mathscr{U}_2$.

4. Proof of the Main Theorem

In this section we outline a proof of Theorem 2.2. The proof is virtually identical to that of [9].

As we observed at the end of the last section, conclusions (1), (2), and (3) of the theorem are satisfied. Thus, we need only prove conclusion (4).

We begin by computing the functions λ_n of Sect. 3. By Eq. (3.5), this gives us a formal asymptotic expansion for $\Phi(x, y, r, t)$. This in turn gives us a formal asymptotic expansion for $\Psi(R, r, t) = U_1(t) U_2(t) \Xi(R, r, t)$, where $\Xi(R, r, t) = \Phi(R, (R - a(t))/\varepsilon, r, t)$.

We let Ψ_J denote the truncated asymptotic expansion for Ψ that only contains the terms that arose from $\lambda_0, \lambda_1, \ldots, \lambda_J$, $\lambda_{J+1}^{\perp}$, and $\lambda_{J+2}^{\perp}$. We substitute Ψ_J into Eq. (3.1), and add $[G_{a(t)}^{[J+2]} - E(R)] \Psi_J$, where $G_{a(t)}^{[J+2]}$ is the $(J+2)^{\text{nd}}$ order Taylor series expansion of $E(R)$ about the point $a(t)$. By virtue of the computations of Sect. 3, we observe the cancellation of all terms of order ε^{ℓ}, that do not involve derivatives of $F(R)$, for $\ell \leqq J + 2$. Lemma 4.3 of [9] shows that the norm of the term $[G_{a(t)}^{[J+2]} - E(R)] \Psi_J$ is bounded by a constant times ε^{J+3}. Lemma 4.2 of [9] shows that the terms involving derivatives of $F(R)$ are bounded by constants times ε^{ℓ} for arbitrary ℓ. Theorem 2.1 then follows from Lemma 4.1 of [9].

References

1. Born, M., Oppenheimer, R.: Zur Quantentheorie der Molekeln. Ann. Phys. (Leipzig) **84**, 457–484 (1927)
2. Combes, J.-M.: On the Born–Oppenheimer Approximation. In: International Symposium on Mathematical Problems in Theoretical Physics Araki, H. (ed.), pp. 467–471. Berlin, Heidelberg, New York: Springer 1975
3. Combes, J.-M.: The Born–Oppenheimer Approximation. In: The Schrödinger Equation Thirring, W., Urban, P. (eds.), pp. 139–159. Wien, New York: Springer 1977
4. Combes, J.-M., Duclos, P., Seiler, R.: The Born–Oppenheimer approximation. In: Rigorous Atomic and Molecular Physics Velo, G., Wightman, A. (eds.), pp. 185–212. New York: Plenum Press 1981
5. Hagedorn, G. A.: Semiclassical quantum mechanics. I: The $\hbar \to 0$ limit for coherent states. Commun. Math. Phys. **71**, 77–93 (1980)
6. Hagedorn, G. A.: Semiclassical quantum mechanics. III: The large order asymptotics and more general states. Ann. Phys. **135**, 58–70 (1981)
7. Hagedorn, G. A.: Semiclassical quantum mechanics. IV: Large order asymptotics and more general states in more than one dimension. Ann. Inst. H. Poincaré Sect. A **42**, 363–374 (1985)
8. Hagedorn, G. A.: A time dependent Born–Oppenheimer approximation. Commun. Math. Phys. **77**, 1–19 (1980)
9. Hagedorn, G. A.: High order corrections to the time-dependent Born–Oppenheimer approximation. I: Smooth potentials. Ann. Math. **124**, 571–590 (1986). Erratum. Ann. Math. **126**, 219 (1987)
10. Hagedorn, G. A.: Higher order corrections to the time-independent Born–Oppenheimer approximation. I: Smooth potentials. Ann. Inst. H. Poincaré Sect. A. **47**, 1–19 (1987)
11. Hagedorn, G. A.: High order corrections to the time-independent Born–Oppenheimer approximation. II: Diatomic coulomb systems. Commun. Math. Phys. **116**, 23–44 (1988)
12. Hunziker, W.: Distortion analyticity and molecular resonance curves. E.T.H. Preprint

13. Kamke, E.: Differentialgleichungen I. Gewöhnliche Differentialgleichungen (Auflage 6) Leipzig: Akademische Verlagsgesellschaft Geest und Portig K.-G. 1969
14. London, F.: Über den Mechanismus der Homoöpolaren Bindung. In: Probleme der Modernen Physik Debye, P. (ed.). Leipzig: S. Hirzel 1928
15. Seiler, R.: Does the Born–Oppenheimer approximation work? Helv. Phys. Acta **46**, 230–234 (1973)
16. Wüller, U.: Existence of time evolution for Schrödinger operators with time dependent singular potentials. Ann. Inst. H. Poincaré Sect. A. **44**, 155–171 (1986)
17. Yajima, K.: Existence of solutions for Schrödinger evolution equations. Commun. Math. Phys. **110**, 415–426 (1987)

Communicated by B. Simon

Received July, 14, 1987; in revised form December 30, 1987