JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 71, 36–47 (1979)

On the Nonrelativistic Limits of the Klein–Gordon and Dirac Equations*

ANDREW Y. SCHOENE[1]

General Motors Research Laboratories, Warren, Michigan 48090

Submitted by Peter D. Lax

INTRODUCTION

In this paper we give a mathematical proof that the Klein-Gordon and Dirac equations of relativistic quantum mechanics have the correct nonrelativistic limits. Our proof applies to the physically important case of an external electromagnetic field, albeit with some restrictions on the size of the field (in a previous paper [7] only the free-field case was treated). After applying transformations to facilitate study of the nonrelativistic limiting behavior, both equations can be represented in operator form as $\epsilon u^\epsilon_{tt} - i\hbar u_t^\epsilon + Su^\epsilon = 0$, where S is a self-adjoint operator on an appropriate Hilbert space and ϵ is a small constant. The correct nonrelativistic equations are formally obtained by setting $\epsilon = 0$. On physical grounds it is argued that u^ϵ_{tt} cannot be appreciable, and thus ϵu^ϵ_{tt} can be neglected. Since this is a singular perturbation problem, it has been established only comparatively recently (Zlamal [9], Smoller [8], Friedman [1]; an excellent survey is given in O'Malley [4]) that solutions of the above operator equation converge as $\epsilon \to 0$ to solutions of $-i\hbar u_t + Su = 0$.

Our results go further by estimating the difference between solutions of the relativistic and nonrelativistic equations in terms of the parameter ϵ, but are restricted to operators S which are time-independent, continuous, and "not too large." In particular, electostatic potentials with Coulomb type $(1/r)$ singularities are formally excluded. However, by definition of the nonrelativistic limit all other energies in the system must be small compared to the rest energy mc^2. Physically this means that the nonrelativistic limit cannot be employed in a region of space where the electostatic potential is very large, for example, in the neighborhood of the nucleus.

The proof is an extension of the methods employed in [7], which treated the

* This research was partially supported by NSF Grant GP-8856 and based in part on work submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of New Mexico.

[1] Current address: Department of Mathematics, University of Michigan, Dearborn, Michigan 48128.

0022-247X/79/090036-12$02.00/0

case of a positive semi-definite operator S, i.e., no external field. A somewhat involved perturbation argument is used here to reduce the lower semi-bounded case to the positive one. Conditions on the electromagnetic potential sufficient to insure that the Hamilotnian (S) is lower semi-bounded are carefully analyzed. The treatment of the nonrelativistic limit of the Dirac equation by semi-group methods is also believed to be new.

Our results are based on the following theorem which will be proved in Section 2. Theorem 2 of Section 2 gives operator-theoretic conditions under which the lower order perturbations due to the electrostatic potential can be neglected.

THEOREM 1. *Let w^ϵ and w satisfy, respectively,*

$$\begin{gathered}\epsilon w^\epsilon_{tt} - 2aiw_t{}^\epsilon + Sw^\epsilon = 0, \qquad w^\epsilon(0) = g \in D(S^2), \\ w_t{}^\epsilon(0) = g_1 \in D(S)\end{gathered} \tag{1}$$

$$-2aiw_t + Sw = 0, \qquad w(0) = g \tag{2}$$

where S (the Hamiltonian) is a self-adjoint and lower semi-bounded operator with lower bound $-d$ on a Hilbert space H, D denotes domain, and ϵ, a, d are positive. Then

$$\| w^\epsilon(t) - w(t)\| \leqslant \epsilon C(t+1)\left(\sum_{k=0}^{2} \| S^k g \| + \sum_{k=0}^{1} \| S^k g_1 \|\right)$$

for a constant C.

1. The Klein-Gordon Equation

In what follows $\mathbf{A} = (a_1, a_2, a_3)$ denotes the magnetic vector potential, ϕ the electrostatic potential, and $\mathbf{x} = (x_1, x_2, x_3)$. Following Schiff [6, p. 469] we make the substitution $u^\epsilon(\mathbf{x}, t) = u(\mathbf{x}, t)\exp(imc^2t/\hbar)$ in the Klein-Gordon equation for a particle of charge e and mass m:

$$\hbar^2 u_{tt} + 2ie\phi u_t - e^2\phi^2 u + c^2(i\hbar\nabla + (e/c)\mathbf{A})^2 u + m^2c^4 u = 0 \tag{3}$$

to factor out the relativistic rest energy mc^2. There results

$$\begin{aligned}(\hbar^2/2mc^2)\, u^\epsilon_{tt} &- i\hbar(1 - e\phi/mc^2)\, u_t{}^\epsilon \\ &+ [e\phi(1 - e\phi/2mc^2) + (1/2m)\,(i\hbar\nabla + (e/c)\,\mathbf{A})^2]\, u^\epsilon = 0.\end{aligned} \tag{4}$$

The Schrödinger equation

$$-i\hbar u_t + [(1/2m)(i\hbar\nabla + (e/c)\mathbf{A})^2 + e\phi]u = 0 \tag{5}$$

can now be obtained from (4) by dropping the first term and the two terms in $e\phi/mc^2$, arguing that in the nonrelativistic limit (all other energies in the system are small compared with mc^2) $(\hbar^2/2mc^2)u_{tt}^\epsilon$ is small compared to $i\hbar u_t^\epsilon$ and $e\phi/mc^2$ is small compared to 1. It is not difficult to justify neglecting the latter two terms if their smallness is made precise. This is done in Theorem 2 of the Appendix. More challenging is the first term because it represents a singular perturbation (its neglect reduces the order of the equation). Consequently, we shall concentrate on the effects of this term and drop the other two, obtaining

$$\epsilon w_{tt}^\epsilon - i\hbar w_t^\epsilon + [(1/2m)(i\hbar\nabla + (e/c)\mathbf{A})^2 + e\phi]\, w^\epsilon = 0. \tag{6}$$

In order to utilize the results of operator theory we must choose an appropriate function space for the solutions of the Klein-Gordon equation. Our interest in the nonrelativistic limit dictates the choice $w^\epsilon(\mathbf{x}, t) \in L^2(R^3) \times C^2[0, T]$ for some large time T. That is, we assume that each solution of the modified Klein-Gordon (6) has a finite L^2 norm at each point in time. This is reasonable because the corresponding solution of the Schrödinger equation (5) belongs to $L^2(R^3) \times C^1[0, T]$, and allows us to estimate the divergence of the corresponding solutions over time. To make the transition to the Hilbert space $L^2(R^3)$ we let the operator in brackets in (6) operate on functions in $C_0^\infty(R^3)$. It has been shown by Kato (see [2] for an excellent survey article) that the symmetric operator so defined has a unique self-adjoint extension on $L^2(R^3)$ which is lower semi-bounded provided: (i) $\mathbf{A}(\mathbf{x})$ is continuously differentiable (it may behave arbitrarily as $|\mathbf{x}| \to \infty$), (ii) ϕ can be written as the sum of a function in L^2 and a bounded function. It is worth noting that (i) allows a uniform magnetic field (the Zeeman effect) and (ii), electrostatic potentials of Coulomb type. However, we cannot justify the discarding of the $e^2\phi^2/2mc^2$ term from (4) to obtain (6) for such potentials.

We are now in a position to apply Theorem 1. Inspection of its proof shows that the leading term in the estimate is actually $\epsilon/a = 2\hbar/(mc^2)$, and the complete estimate has the form

$$\begin{aligned}\| u^\epsilon(t) - u(t)\| \leqslant C(\hbar/mc^2)[d\,\| g \| + \hbar^{-1} \| Sg \| + td^2 \| g \| \\ + t\, d\hbar^{-1} \| Sg \| + t\hbar^{-2} \| S^2 g \| + \text{terms involving } g_1].\end{aligned} \tag{7}$$

Here C is a dimensionless constant. Each of the terms in brackets can be related to the energy of the system. Since the transformation $u \to u^\epsilon$ is equivalent to subtracting out the rest energy, the remaining energies and hence the terms in brackets are small compared to mc^2 in the nonrelativistic limit. In fact if $\phi = 0$ (no external field), then S is nonnegative and we may take $d = 0$ [7].

The Dirac Equation. We let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ and β denote the four 4×4 Dirac matrices. If $\mathbf{F}$ and $\mathbf{G}$ are two vector operators, then the operator $\mathbf{F} \cdot \mathbf{G}$ is defined by $(\mathbf{F} \cdot \mathbf{G})f = (F_1G_1 + F_2G_2 + F_3G_3)f$, and the superscript ^{T}

applied to a row vector will denote the corresponding column vector. With this notation we may write the Dirac equation as

$$[i\hbar\, \partial/\partial t - e\phi + (i\hbar c\nabla + e\mathbf{A}) \cdot \boldsymbol{\alpha} - mc^2\beta]\mathbf{u} = 0 \tag{1}$$

where $\mathbf{u} = (u_1, u_2, u_3, u_4)^T$ is the 4-component wave function. For the nonrelativistic reduction it is more convenient to rewrite (1) as a coupled system of two equations

$$(i\hbar\, \partial/\partial t - e\phi)\mathbf{u}_a + (i\hbar c\nabla + e\mathbf{A}) \cdot \boldsymbol{\sigma}\mathbf{u}_b - mc^2\mathbf{u}_a = 0 \tag{2}$$

$$(i\hbar\, \partial/\partial t - e\phi)\mathbf{u}_b + (i\hbar c\nabla + e\mathbf{A}) \cdot \boldsymbol{\sigma}\mathbf{u}_a + mc^2\mathbf{u}_b = 0 \tag{3}$$

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the 2×2 Pauli spin matrices, and $\mathbf{u}_a = (u_1, u_2)^T$, $\mathbf{u}_b = (u_3, u_4)^T$ are the "large" and "small" components, respectively, of $\mathbf{u}$. Our aim is to show that in the nonrelativistic limit $\mathbf{u}_a$ satisfies the Pauli equation

$$i\hbar\mathbf{w}_t = (1/2m)(i\hbar\nabla + (e/c)\mathbf{A})^2 + e\phi - (e\hbar/2mc)\mathbf{H} \cdot \boldsymbol{\sigma}]\mathbf{w} \tag{4}$$

and $\mathbf{u}_b$ is negligible.

In order to uncouple (2) and (3) we define new functions $\mathbf{w}_+$ and $\mathbf{w}_-$ by $\mathbf{w}_+ = \mathbf{u}_a + \mathbf{u}_b$, $\mathbf{w}_- = \mathbf{u}_a - \mathbf{u}_b$, where the addition is componentwise. Let $\mathbf{L}$ denote $(i\hbar c\nabla + e\mathbf{A}) \cdot \boldsymbol{\sigma}$. By first adding and then subtracting (2) and (3) we obtain

$$(i\hbar\, \partial/\partial t - e\phi)\mathbf{w}_+ - \mathbf{L}\mathbf{w}_+ - mc^2\mathbf{w}_- = 0 \tag{5}$$

$$(i\hbar\, \partial/\partial t - e\phi)\mathbf{w}_- + \mathbf{L}\mathbf{w}_- - mc^2\mathbf{w}_+ = 0. \tag{6}$$

Solving (5) for $\mathbf{w}_-$ and substituting into (6) yields

$$(Q^2 + \mathbf{L}Q - Q\mathbf{L} - \mathbf{L}^2 - m^2c^4)\mathbf{w}_+ = 0 \tag{7}$$

where Q denotes $i\hbar\partial/\partial t - e\phi$. It can be shown [6, p. 478] that $\mathbf{L}Q - Q\mathbf{L}$ reduces to $-ie\hbar c\mathbf{E} \cdot \boldsymbol{\sigma}$ and $\mathbf{L}^2$ reduces to $(i\hbar c\nabla + e\mathbf{A})^2 - eh c\mathbf{H} \cdot \boldsymbol{\sigma}$ where $\mathbf{E} = -c^{-1}\partial\mathbf{A}/\partial t - \nabla\phi = -\nabla\phi$ (since $\mathbf{A}$ is assumed time-independent) and $\mathbf{H} = \nabla \times \mathbf{A}$ are the electric and magnetic field strengths, respectively. Hence (7) becomes, again setting $\epsilon = \hbar^2/2mc^2$, and $\mathbf{w}_+^\epsilon(\mathbf{x}, t) = \mathbf{w}_+(\mathbf{x}, t)\exp(imc^2t/\hbar)$,

$$\begin{aligned}\epsilon\mathbf{w}^\epsilon_{+tt} - i\hbar(1 - e\phi/mc^2)\,\mathbf{w}^\epsilon_{+t} + e\phi(1 - e\phi/2mc^2)\,\mathbf{w}_+{}^\epsilon + [(1/2m)\,(i\hbar\nabla + e/c\mathbf{A})^2 \\ - (e\hbar/2mc)\,\mathbf{H} \cdot \boldsymbol{\sigma} + (ie\hbar/2mc)\,\mathbf{E} \cdot \boldsymbol{\sigma}]\,\mathbf{w}_+{}^\epsilon = 0.\end{aligned} \tag{8}$$

Following a similar procedure we obtain the same equation for $\mathbf{w}_-{}^\epsilon$ except that $+(ie\hbar/2mc)\mathbf{E} \cdot \boldsymbol{\sigma}$ is replaced by $-(ie\hbar/2mc)\mathbf{E} \cdot \boldsymbol{\sigma}$.

In situations of physical interest the $\mathbf{E}$ term in (8) is of order $(v/c)^2$ times the $e\phi$ term, according to Schiff [6, p. 479], where v is the velocity of the particle described by $\mathbf{u}$. Thus, in the nonrelativistic limit ($v \ll c$) the $\mathbf{E}$ term in (8) and in the similar equation for $\mathbf{w}_-^\epsilon$ may be dropped as a small perturbation of a lower order term. This done we have, to order $(v/c)^2$, $\mathbf{w}_+^\epsilon = \mathbf{w}_-^\epsilon$ or $\mathbf{u}_b^\epsilon = 0$ and $\mathbf{u}_a^\epsilon$ satisfying (8) without the $\mathbf{E}$ term. In the nonrelativistic limit we are also justified in dropping the two $e\phi/mc^2$ terms in (8), following the argument of Section 1. We thus obtain the two-component Pauli equation for $\mathbf{u}_a$, namely,

$$\epsilon\partial^2\mathbf{u}_a/\partial t^2 - i\hbar\partial\mathbf{u}_a/\partial t + [(1/2m)(i\hbar\nabla + (e/c)\mathbf{A})^2 + e\phi - (e\hbar/2mc)\mathbf{H}\cdot\boldsymbol{\sigma}]\mathbf{u}_a = 0. \tag{9}$$

It is not necessary to impose growth conditions on $\mathbf{A}$ directly since it is part of a positive operator. However, the $\mathbf{H}$ term is not negligible and so we must assume, e.g., that $\mathbf{H}$ as well as ϕ can be written as the sum of a square integrable function and a bounded function. Since $\mathbf{H} = \nabla \times \mathbf{A}$, this is a requirement on the growth of the first order derivatives of $\mathbf{A}$. From such assumptions it follows from the result quoted in Section 1 that the operator in brackets in (9) is self-adjoint and lower semi-bounded on $L_2(R^3) \times L_2(R^3)$. We can now apply Theorem 1 to deduce that the solutions of (9) and (4) differ by an estimate of the form (7) of Section 1.

Strictly speaking we have shown that the transformed ($\psi \to \psi^\epsilon \exp(i\hbar t/\epsilon)$) Klein-Gordon and Dirac (the two "large" components) wave functions are closely approximated by the Schrödinger and Pauli wave functions, respectively. However, the time factor $\exp(i\hbar t/\epsilon)$ represents a very high frequency oscillation which is not physically observable. From a physical standpoint, therefore, it is sufficient to obtain the nonrelativistic reduction for the transformed wave functions.

2

THEOREM 1. *Let w^ϵ and u satisfy, respectively,*

$$\epsilon w_{tt}^\epsilon - 2aiw_t^\epsilon + Sw^\epsilon = 0, \qquad w^\epsilon(0) = g \in D(S^2), \qquad w_t^\epsilon(0) = g_1 \in D(S) \tag{1}$$

$$-2aiu_t + Su = 0, \qquad u(0) = g \tag{2}$$

where S is a self-adjoint and lower semibounded operator on a Hilbert space H, D denotes domain, and ϵ, a are positive. Then

$$\| w^\epsilon(t) - u(t)\| \leqslant \epsilon C(t+1)\left(\sum_{k=0}^{2} \| S^k g\| + \sum_{k=0}^{1} \| S^k g_1\|\right)$$

for a constant C.

Proof. We first note that the "energy" estimates developed in the proof of Theorem 2 below can be used to show the solutions of (1) and (2) are unique, but we omit the verification. By hypothesis, there is a constant $d > 0$ such that

$S + 2ad$ is positive, i.e., $((S + 2ad)f, f) \geqslant \epsilon(f, f)$ for all $f \in D(S)$. Hence, $S + 2ad$ has a unique positive (self-adjoint) square root, which we shall denote by T [3, p. 281]. We consider the first-order system

$$\frac{\partial}{\partial t}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = i\left(\frac{1}{\epsilon^{1/2}}\begin{bmatrix} T & 0 \\ 0 & -T \end{bmatrix} + \frac{1}{\epsilon}\begin{bmatrix} a & -a \\ -a & a \end{bmatrix} + \begin{bmatrix} 0 & d \\ d & 0 \end{bmatrix}\right)\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \equiv iL^\epsilon \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad (3)$$

with $u_1(0) = u_2(0) = g \in D(S^2)$. If we multiply the second of these two equations by $i(d - a/\epsilon)$ and the first by $\partial/\partial t + (i/\epsilon^{1/2})T - ia/\epsilon$, and then substitute for u_2 in the first equation we find that u_1 satisfies the equation

$$\begin{aligned} \epsilon v_{tt}^\epsilon - 2aiv_t^\epsilon + (S + \epsilon d^2)\, v^\epsilon &= 0, \\ v^\epsilon(0) = g, \qquad v_t^\epsilon(0) &= i/\epsilon^{1/2}Tg + idg. \end{aligned} \quad (4)$$

A similar procedure shows that u_2 also satisfies this equation (with different initial conditions). The proof is now accomplished in two main stages. We first compare the solutions of (2) and (3) by taking Laplace transforms and working backwards by use of the identity (6). Then the difference between (1) and (4) is estimated by "energy" methods.

The operator L^ϵ of (3) is evidently self-adjoint on $H \times H$, and, consequently, iL^ϵ generates a group of unitary operators on $H \times H$ which we denote by $\exp(iL^\epsilon t)$. From the theory of semi-groups we know that for $g \in D(T)$, the unique solution of (3) is given by

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \exp(iL^\epsilon t)\begin{pmatrix} g \\ g \end{pmatrix}. \quad (5)$$

Now the Laplace transform of a semi-group is also the resolvent of its infinitesimal generator, here $(\lambda I - iL^\epsilon)^{-1}$, where I denotes the identity operator on $H \times H$. Taking determinants we find

$$(\lambda - iL^\epsilon)^{-1} = \begin{bmatrix} \epsilon\lambda + i\epsilon^{1/2}T - ia & -i(a - \epsilon d) \\ -i(a - \epsilon d) & \epsilon\lambda - i\epsilon^{1/2}T - ia \end{bmatrix} D$$

where $D = (\epsilon\lambda^2 - 2ai\lambda + \epsilon d^2 - S)^{-1}$. Because S is self-adjoint, D is everywhere defined and bounded for $\lambda > 0$. Evidently

$$\lim_{\epsilon \to 0}(\lambda - iL^\epsilon)^{-1} = -i\begin{bmatrix} a & a \\ a & a \end{bmatrix}(-2ai\lambda - S)^{-1} \equiv R.$$

We would like that R be the resolvent of $iL \equiv (1/2ai)\left[\begin{smallmatrix} S & 0 \\ 0 & S \end{smallmatrix}\right]$, a two-dimensional representation of (2). It may be verified by multiplication that $\lambda - iL$ and R are inverse to each other, provided both are restricted to operate on elements of the form $\left(\begin{smallmatrix} g \\ g \end{smallmatrix}\right)$. To get from the convergence of the Laplace transforms (resolvents) back to convergence of the solutions we need the following identity [7, p. 255], valid for any two semi-groups $\exp(S_1 t)$ and $\exp(S_2 t)$ on a Banach space B and any complex λ for which $(\lambda - S_1)^{-1}$ and $(\lambda - S_2)^{-1}$ exist:

$$\begin{aligned}[\exp(S_2 t) - \exp(S_1 t)]\, w &= [(\lambda - S_2)^{-1} - (\lambda - S_1)^{-1}] \exp(S_2 t)\, (\lambda - S_2)\, w \\ &\quad - \exp(S_1 t)\, [(\lambda - S_2)^{-1} - (\lambda - S_1)^{-1}]\, (\lambda - S_2)\, w \\ &\quad + \int_0^t \exp(S_1(t-s))\, [(\lambda - S_2)^{-1} - (\lambda - S_1)^{-1}] \exp(S_2 s)\, (\lambda - S_2)^2\, w\, ds \end{aligned} \tag{6}$$

where $w \in D(S_2^2)$. To apply this we take $S_1 = iL^\epsilon$, $S_2 = iL$, $\lambda > 0$, and $w = \binom{g}{g}$ for $g \in D(S^2)$. By the unitary property $\| \exp(iL^\epsilon t)\| = \| \exp(iLt)\| = 1$. Hence, to bound the right-hand-side of (6) in terms of $\|(\lambda - iL)g\|$ and $\|(\lambda - iL)^2 g\|$ we only need an estimate for $\|(\lambda - iL)^{-1} - (\lambda - iL^\epsilon)^{-1}\|$. Now

$$\begin{aligned}(\lambda - iL)^{-1} - (\lambda - iL^\epsilon)^{-1} &= -i \begin{bmatrix} a & a \\ a & a \end{bmatrix} ((-2ai\lambda - S)^{-1} - D) - \epsilon \begin{bmatrix} \lambda & id \\ id & \lambda \end{bmatrix} D - i\epsilon^{1/2} \begin{bmatrix} -T & 0 \\ 0 & T \end{bmatrix} D \\ &\equiv R_1 + R_2 + R_3 . \end{aligned} \tag{7}$$

Each of these terms can be estimated using the functional calculus for self-adjoint operators [5, Chap. IX]. We have

$$\begin{aligned}\| D - (-2ai\lambda - S)^{-1} \|^2 &= \int_{-\infty}^{\infty} \left| \frac{1}{\epsilon\lambda^2 - 2ai\lambda + \epsilon d^2 - 2ad + x^2} - \frac{1}{-2ai\lambda - 2ad + x^2} \right|^2 d\| E_x \|^2 \end{aligned}$$

where $\{E_x\}$ is the spectral resolution of T. Here we have used the fact that $S = T^2 - 2ad$. A calculation shows that the integrand is bounded by $[\epsilon(\lambda^2 + d^2)/(4a^2\lambda^2)]^2$; consequently, an upper bound for R_1 is $\epsilon(\lambda^2 + d^2)/2a\lambda^2$. Similarly,

$$\| D \| \leqslant \max_{x \in (-\infty, \infty)} 1/((x^2 - 2ad + \epsilon\lambda^2 + \epsilon d^2)^2 + 4a^2\lambda^2)^{1/2} \leqslant 1/2a\lambda,$$

$$\left\| \begin{bmatrix} \lambda & id \\ id & \lambda \end{bmatrix} \right\| = (\lambda^2 + d^2)^{1/2}, \qquad \text{and} \qquad \| R_2 \| \leqslant \epsilon(\lambda^2 + d^2)^{1/2}/2a\lambda.$$

R_3 may also be estimated using the functional calculus, and we would find $\| R_3 \| \leqslant C\epsilon^{1/2}$. Combining these 3 estimates yields $\| \exp(iLt) - \exp(iL^\epsilon t)\| = O(\epsilon^{1/2})$. This estimate on the difference of the semi-groups can be carried over immediately to the difference of the solutions of (2) and (4), and this would conclude the first stage.

However, a closer analysis shows that the estimate on the difference of the solutions can be improved to $O(\epsilon)$. This has already been done for the case $d = 0$ in [7] and will only be outlined here. First we rewrite (7) as

$$\begin{aligned}
&[\exp(iLt) - \exp(iL^{\epsilon}t)]\, w \\
&\quad = R_3 \exp(iLt)\,(\lambda - iL)\, w - \exp(iL^{\epsilon}t)\, R_3(\lambda - iL)\, w \\
&\qquad + \int_0^t \exp(iL^{\epsilon}(t-s))\, R_3 \exp(iLs)\,(\lambda - iL)^2\, w\, ds + O(\epsilon)
\end{aligned} \tag{7'}$$

where $O(\epsilon)$ denotes all the terms involving R_1 and R_2, and $w = (g, g)^T$. Since $(u_1 + u_2)/2$ also satisfies (4), together with the initial conditions $v^{\epsilon}(0) = g$, $v_t^{\epsilon}(0) = idg$, we add the two equations represented by (7′). Denoting the elements of $\exp(iL^{\epsilon}t)$ by $S_{ij}(t)$, $1 \leqslant i, j \leqslant 2$, we find

$$\begin{aligned}
u_1 + u_2 - 2u &= \sum_{i,j=1}^{2} [S_{ij}^{\epsilon}(t) - S_{ij}(t)]\, g = 0 + i(\epsilon^{1/2})\,(S_{22}^{\epsilon} - S_{11}^{\epsilon})\, TDT_1 g \\
&\quad - i(\epsilon^{1/2}) \int_0^t (S_{22}^{\epsilon} - S_{11}^{\epsilon})\,(t-s)\, TDT_1{}^2 \exp(St/2ai)\, g\, ds + O(\epsilon)
\end{aligned} \tag{8}$$

where $T_1 \equiv \lambda - S/2ai$. The above utilizes the fact that $S_{12}^{\epsilon} \equiv S_{21}^{\epsilon}$ and $\exp(St/2ai)$ commutes with $T_1{}^2$. The first part of the proof could be quickly concluded if it were known that $\| S_{22}^{\epsilon} - S_{11}^{\epsilon} \| = O(\epsilon^{1/2})$. This is stated in Lemma 1 below. Using the estimate from Lemma 1 and $\| \exp(St/2ai)\| = 1$ in (8) yields, for $a > \epsilon d$,

$$\| u_1 + u_2 - 2u \| < (\epsilon/(a - \epsilon d))\, \| T^2 D \|\, (\| T_1 g \| + t \, \| T_1{}^2 g \|) + O(\epsilon). \tag{9}$$

The functional calculus shows that $\| T^2 D \| \leqslant (d^2 + \lambda^2)^{1/2}/\lambda$ provided $\epsilon(\lambda^2 + d^2) \leqslant 4ad$. Since $\lambda \geqslant 0$ is arbitrary, we set $\lambda = d$ for convenience and assume that $\epsilon d \leqslant a/2$. The above inequality is thus satisfied, and $\epsilon/(a - \epsilon d) \leqslant 2\epsilon/a$. With this and the $O(\epsilon)$ terms we find from (9)

$$\| u_1(t) + u_2(t) - 2u(t)\| \leqslant (8\epsilon/a)(\| T_1 g \| + t \, \| T_1{}^2 g \|). \tag{10}$$

Let $(v_1, v_2)^T = \exp(iL^{\epsilon}t)(h, h)^T$. Then $\epsilon^{1/2}(v_1 - v_2)/2$ satisfies (4) with the initial data $v^{\epsilon}(0) = 0$, $v_t^{\epsilon}(0) = iTh$. Further, $\| v_1 - v_2 \| = \|[S_{11} - S_{22}]h \| \leqslant (\epsilon^{1/2}/(a - \epsilon d))\, \| Th \|$ by Lemma 1. By linearity $y = (u_1 + u_2)/2 + \epsilon^{1/2}(v_1 - v_2)/2$ has the initial data $y(0) = g$, $y_t(0) = i(dg + Th)$, and combining the last estimate with (10) we obtain

$$\| y(t) - u(t)\| \leqslant (4\epsilon/a)(\| T_1 g \| + t \, \| T_1{}^2 g \|) + (\epsilon/a)\, \| Th \|. \tag{11}$$

We have now estimated the difference between the solutions of (2) and (4).

Next comes a comparison of the solutions of (1) and (4) (for the same initial data). We shall generalize (4) somewhat and prove

THEOREM 2. *Let w and y satisfy, respectively,* (1) *and* $\epsilon y_{tt} - 2ai(1 - Q) y_t + (S - V) y = 0$, with the same initial data. Let $z = w - y$. Then $\| z \| = O(\epsilon)$. *Here we assume that Q and V are self-adjoint and either* Ia *and* II *or* Ib *and* II *holds*:

(Ia$_1$) $(Qf, f) \leqslant 0$ *and*

(Ia$_2$) $\| Qf \| \leqslant r_1 \| f \| + r_2 \| Sf \|$, $f \in D(S)$, *and* $r_1, r_2 < \epsilon$.

(Ib) *Q is bounded with* $\| Q \| < \epsilon$.

(II) $\| Vf \| \leqslant s_1 \| f \| + s_2 \| Sf \|$, $f \in D(S)$, *and* $s_1, s_2 < \epsilon$.

Proof. z satisfies

$$\begin{gathered} \epsilon z_{tt} - 2ai(1 - Q) z_t + (S - V) z = 2aiQw_t - Vw \equiv \epsilon f(t), \\ z(0) = z_t(0) = 0. \end{gathered} \tag{12}$$

Taking the innerproduct of (12) with z_t first on the right and then on the left and adding the resulting two equations yields

$$\epsilon(d/dt)(z_t, z_t) + (d/dt)((S - V)z, z) = 2\epsilon \operatorname{Re}(f, z_t),$$

or

$$\epsilon \| z_t \|^2 + ((S - V) z, z) = 2\epsilon \operatorname{Re} \int_0^t (f(s), z_t(s))\, ds, \tag{13}$$

where we have used the self-adjointness of Q and $S - V$. If we do the same with z, only this time subtract, we find

$$\epsilon[(z_{tt}, z) - (z, z_{tt})] - 2ai(d/dt)((1 - Q)z, z) = 2i\epsilon \operatorname{Im}(f, z). \tag{14}$$

Now $(z_{tt}, z) = (d/dt)(z_t, z) - (z_t, z_t)$, $(z, z_{tt}) = (d/dt)(z, z_t) - (z_t, z_t)$, and substituting this into (14) and integrating we obtain

$$\epsilon \operatorname{Im}(z_t, z) - a((1 - Q) z, z) = \epsilon \operatorname{Im} \int_0^t (f(s), z(s))\, ds. \tag{15}$$

For simplicity, we choose hypothesis Ia and obtain from (15), after discarding the term in Q,

$$a \| z \|_M^2 \leqslant \epsilon \| z_t \|_M \| z \|_M + \epsilon K t \| z \|_M . \tag{16}$$

where $\| \cdot \|_M = \sup_{0 \leqslant s \leqslant t} \| \cdot \|$ and K denotes an upper bound for $\| f(s) \|_M$ to be estimated. It follows from the assumption on V that $((S - V)z, z) \geqslant -d_1(z, z)$

for a constant $d_1 > 0$ satisfying $d_1 \leqslant 2ad + \epsilon \max((1-\epsilon)^{-1}, 1 + 2ad)$. Inserting this in (13),

$$\epsilon \| z_t \|_M^2 \leqslant 2\epsilon Kt \| z_t \|_M + d_1 \| z \|_M^2 ,$$

and completing the square,

$$\begin{aligned} \epsilon \| z_t \|_M &\leqslant (\epsilon d_1 \| z \|_M^2 + \epsilon^2 K^2 t^2)^{1/2} + \epsilon Kt \\ &\leqslant (\epsilon d_1)^{1/2} \| z \|_M + 2\epsilon Kt. \end{aligned}$$

Substituting this into (16) yields finally

$$\| z \|_M \leqslant 3\epsilon Kt/(a - (\epsilon d_1)^{1/2}). \tag{17}$$

With hypothesis Ib and II we find that (16) and thus (17) hold with a replaced by $a - \epsilon$.

It remains to bound $\epsilon f(s) \equiv 2aiQw_t - Vw$. Choosing (Ia) and (II), it will be sufficient to bound w, Sw, w_t, and Sw_t. Applying the same procedure used to derive (13) and (15) to (1), we obtain the following two equations:

$$\begin{aligned} \epsilon \| w_t \|^2 + (Sw, w) &= C_1 \equiv \epsilon \| g_1 \|^2 + (Sg, g) \\ \epsilon \operatorname{Im}(w_t, w) - a \| w \|^2 &= C_2 \equiv \epsilon \operatorname{Im}(g_1, g) + \| g \|^2. \end{aligned}$$

As in the first part of the proof, these two equations imply

$$\| w \| \leqslant C_2^{1/2} + (\epsilon C_1)^{1/2}/(a - (\epsilon d_1)^{1/2}).$$

Now Sw also satisfies (1) because of the assumptions $g \in D(S^2)$, $g_1 \in D(S)$, so that the above estimate for w implies

$$\| Sw \| \leqslant (C_2')^{1/2} + (\epsilon C_1')^{1/2}/(a - (\epsilon d_1)^{1/2}),$$

where C_1', C_2' are C_1, C_2 with g, g_1 replaced by Sg, Sg_1 respectively.

To bound w_t we multiply (1) on the right and left by w_{tt} and subtract the resulting equations, obtaining

$$\begin{aligned} &a \| w_t \|^2 - \operatorname{Im}(Sw, w_t) = C_3 \equiv a \| g_1 \|^2 + \operatorname{Im}(Sg, g_1), \\ &a \| w_t \|^2 \leqslant C_3 + \| Sw \| \| w_t \|. \end{aligned}$$

Since Sw is bounded, it follows that w_t is bounded. Again applying S to (1), we have immediately that Sw_t is bounded. Inserting these four bounds into (Ia) and (II) establishes that $2aiQw_t - Vw = O(\epsilon)$, which completes the proof of Theorem 2.

Remark. The hypotheses of Theorem 2 are slightly restrictive for the application to relativistic quantum mechanics. In particular, the ϵ in the bounds can be replaced by $C\epsilon$, e.g., $r_1, r_2 < C\epsilon$ in (Ia_2) where $C \ll 1/\epsilon$; this only inserts another multiplicative constant into the bound for $\| z \|$. Also, the requirement that the perturbation Q (due to the electrostatic potential) be negative is not essential, although it simplifies the proof. In fact it follows from the relative boundedness assumptions (Ia_2) and (II) that $|(Qz, z)| \leqslant C\epsilon[|((S - V)z, z)| + (z, z)]$, [3, p. 343, prob. 3.14], where $C \ll 1/\epsilon$. Now $|((S - V)z, z)|$ can be bounded satisfactorily from (13), which permits the term $a(Qz, z)$ to be retained in (15) and estimated without substantial change to (16). Thus the remainder of the proof goes through as before.

For completeness we sketch the proof (based on Kato [3, V, Sect. 5]) that perturbations due to Coulomb type potentials satisfy (Ia_2). First let S be the self-adjoint extension of the Laplace operator ∇^2 and let $u \in D(S)$. Then

$$\left(\int |\hat{u}(k)|\, dk\right)^2 \leqslant \int (|k|^2 + 1)^{-2}\, dk \int (|k|^2 + 1)^2 |u(k)|^2\, dk$$
$$= C_1 \| Su + u \|^2 < \infty.$$

It is a standard result from Fourier analysis that a function whose Fourier transform is integrable is bounded and continuous, with

$$\| u(x) \|_\infty \leqslant C_2 \int |\hat{u}(k)|\, dk \leqslant C_3(\| Su \| + \| u \|).$$

Now if q is a multiplication operator which can be expressed as $q = q_0 + q_1$ where $q_0 \in L^\infty(R^3)$, $q_1 \in L^2(R^3)$, then $qu \in L^2(R^3)$ with

$$\| qu \|_2 \leqslant \| q_0 \|_\infty \| u \|_2 + \| q_1 \|_2 \| u \|_\infty \leqslant C_4(\| Su \| + \| u \|). \tag{18}$$

Finally, if S' is a more general Hamiltonian of the form $S + q'$, where q' satisfies the same hypothes as q, then if follows readily from (18) that

$$\| qu \|_2 \leqslant C_5(\| S'u \| + \| u \|). \tag{19}$$

For Coulomb type potentials the perturbation Q has the form const. $(h/mc^2)/r$, and since $1/r$ has a decomposition as $q_0 + q_1$, it follows from (19) that (Ia_2) holds for an ϵ the order of $\hbar/mc^2$. The perturbation V was introduced to cover the $e^2\phi^2/2mc^2$ term in (4), Section 1 and the ϵd^2 term in (4), Appendix. However, as discussed in the Introduction, $e^2\phi^2/2mc^2$ cannot be shown to satisfy the hypothesis on V if ϕ is a Coulomb type potential.

To tie up loose ends we remove the restriction on the initial data for $y(t)$ in (11). Recall that we have constructed a $y(t)$ solving (4) with data $y(0) = g$, $y_t(0) = i(dg + Th)$, where $T = (S + 2ad)^{1/2}$. The construction of the square root [3, p. 281] shows that $T^{-1} = (S + 2ad)^{-1/2}$ is a bounded linear operator (defined on all of H), so that the equation $i(dg + Th) = g_1$ is solved for arbitrary

g_1 by $h = T^{-1}(g_1 - idg)$, and $\| Th \| \leqslant \| g_1 \| + d \| g \|$. Recalling that $T_1 \equiv d - S/2ai$, we see that (11) coupled with Theorem 2 yields an estimate of the form required in the statement of Theorem 1.

LEMMA 1. *Let*

$$L^\epsilon(t) \equiv \frac{1}{\epsilon^{1/2}} \begin{bmatrix} T & 0 \\ 0 & -T \end{bmatrix} + \frac{1}{\epsilon} \begin{bmatrix} a & -a \\ -a & a \end{bmatrix} + \begin{bmatrix} 0 & d \\ d & 0 \end{bmatrix},$$

and denote the elements of the semi-group $\exp(iL^\epsilon t)$ *by* $S^\epsilon_{ij}(t)$, $i, j = 1, 2$. *Then for* $f \in D(S)$

$$\| [S^\epsilon_{11}(t) - S^\epsilon_{22}(t)] f \| \leqslant \epsilon^{1/2}/(a - \epsilon d) \| Tf \| .$$

This is proved in [7, *p.* 261] *for the case* $d = 0$. *No essential modifications of the proof are required for* $d > 0$.

ACKNOWLEDGMENT

The author wishes to express warm thanks to Professor Reuben Hersh for his advice and constant encouragement. Professor J. D. Finley III provided helpful comments on an earlier draft of this paper. Dr. Jackson Gay helped reduce the obscurities inherent when a mathematician writes about quantum mechanics (T. Kato excepted).

REFERENCES

1. A. FRIEDMAN, Singular perturbations for the Cauchy problem and for boundary value problems, *J. Differential Equations* **5** (1969), 226–261.
2. T. KATO, Some mathematical problems in quantum mechanics, *Suppl. Progr. Theoret. Phys.* **40** (1967) 3–19.
3. T. KATO, "Perturbation Theory for Linear Operators," Springer-Verlag, Berlin, 1966.
4. R. E. O'MALLEY, JR., Topics in singular perturbations, *in* "Advances in Mathematics" (H. Busemann, Ed.), Vol. 2, Fasc. 4, pp. 365–470, Academic Press, New York, 1968,
5. F. RIESZ AND B. SZ.-NAGY, "Functional Analysis," Ungar, New York, 1955.
6. L. I. SCHIFF, "Quantum Mechanics," 3rd ed., McGraw–Hill, New York, 1968.
7. A. Y. SCHOENE, Semi-groups and a class of singular perturbation problems, *Indiana Univ. Math. J.* **20** (1970), 247–263.
8. J. A. SMOLLER, Singular perturbations of Cauchy's problem, *Comm. Pure Appl. Math.* **18** (1965), 665–677.
9. M. ZLAMAL, "On a Singular Perturbation Problem Concerning Hyperbolic Equations," Institute for Fluid Dynamics and Applied Mathematics, Lecture Series, No. 45, Univ. of Maryland, College Park, 1965.