4-29-2016

Nonrelativistic Limit of the Klein-Gordon Equation

Jack W. Dodson

THE NONRELATIVISTIC LIMIT OF THE KLEIN-GORDON EQUATION

by

Jack Dodson

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Major: Mathematical Sciences

The University of Memphis

May 2016

Abstract

In this Master's thesis a simpler proof that the relativistic wavefunction given by the Klein-Gordon equation converges to the wavefunction given by the Schrödinger equation in the nonrelativistic limit is presented. This proof relies on the spectral theorem for selfadjoint operators and Taylor's theorem.

Acknowlegements

I would like to thank my advisors, professors Jerome and Gisele Goldstein, for their help and support throughout the process of writing this thesis. Whenever I needed a question answered or had trouble on some part of my research, I was always able to go to them for help. I would also like to thank professor Hongqiu Chen, who agreed to serve on my final defense committee at short notice.

In addition, I would like to thank Matthew Roberson, a close friend of mine from whom I was passed this project.

Lastly, I would like to thank my family and friends, who supported me through my writing of this thesis.

Contents

Introduction

Relativistic quantum mechanics is an attempt to unite two of the pillars of twentieth century physics, relativity theory and quantum mechanics. Special relativity allows a more accurate explanation of the behavior of matter as it approaches the speed of light, while quantum mechanics deals with phenomenon at very small length scales. For reasons of consistency, we would hope that relativistic physics, including relativistic quantum mechanics, reduces to nonrelativistic physics as the speed of light tends to infinity. Indeed, in the case of quantum mechanics this was proved rigorously by Andrew Schoene in 1979 [2], although it was surely understood by Dirac in 1930 in a heuristic sense. More specifically, Schoene proved that the model equations for relativistic quantum mechanics, the Dirac equation and the Klein-Gordon equation, reduce to the same result as the Schrödinger equation in the appropriate limits. However, his proof was very technically complex, relying heavily on semigroup theory and operator matrices. In this Master's thesis we present a much simpler proof using the spectral theorem and its associated functional calculus. First we will establish some background in the relevent operator theory and physics, before stating our result.

Operator Theory and Spectral Theory

Spectral Theory Preliminaries

Let us develop spectral theory and the spectral functional calculus. Spectral theory arises not only in the mathematical study of differential equations but also throughout many fields of classical and modern physics. Spectral theory extends the ideas of eigenvectors and eigenvalues to the context of operators on infinite dimensional spaces. Our motivation here is to lead up to the spectral theorem, a mechanism allowing us to deal with operators as if they were functions of a real variable.

Before we discuss spectral theory and the spectral functional calculus in detail, we require some introductory notions from operator theory and functional analysis. Intuitively, one can regard operator theory as an infinite dimensional extension of finite dimensional linear algebra, where linear operators can be thought of as "infinite matrices".

Let $\mathcal{H}$ be a complex Hilbert space. Let $A : \mathfrak{D}(A) \to \mathcal{H}$ be a linear operator, where the domain of A, $\mathfrak{D}(A)$, is a linear subspace of $\mathcal{H}$. A is **closed** if for any sequence of vectors $f_n \in \mathfrak{D}(A)$ such that, as $n \to \infty$, $f_n \to f$ and $Af_n \to g$, it follows that $f \in \mathfrak{D}(A)$ and $Af = g$. In other words, the operator is necessarily defined on what it converges to; more precisely, the graph of A in $\mathcal{H} \times \mathcal{H}$ is a closed subspace.

Densely defined operators will also be important to us, as most of the linear operators which arise in physical applications are densely defined. A linear operator $A : \mathfrak{D}(A) \to \mathcal{H}$ is called **densely defined** if $\mathfrak{D}(A)$ is a dense subset of $\mathcal{H}$, i.e.,

$\overline{\mathcal{D}(A)} = \mathcal{H}$. Note that an operator whose domain is the whole of $\mathcal{H}$ is also densely defined because $\mathcal{H}$ is dense in itself.

Next, we want to distinguish between bounded and unbounded operators. This distinction is important because many important operators on Hilbert spaces, as well as many operators in quantum mechanics, are not bounded.

Let $A : \mathfrak{D}(A) \to \mathcal{H}$ be linear. We call the operator A **bounded** if there exists some $0 < M < \infty$ such that for an $f \in \mathfrak{D}(A) \subset \mathcal{H}$,

$$||Af||_{\mathcal{H}} \leq M||f||_{\mathcal{H}}.$$

The Fourier transform on $L^2(\mathbb{R}^n)$, for example, is bounded (and moreover is unitary). Operators which do not satisfy this boundedness property are called **unbounded**. Bounded (respectively, unbounded) linear operators are everywhere continous (respectively, discontinuous). Particularly, the differential operators on L^2 spaces are unbounded operators.

Another theorem we will need is the Riesz representation theorem. The Riesz representation theorem gives us a relationship between elements of a Hilbert space $\mathcal{H}$ and elements of its dual space, $\mathcal{H}^*$, which is the set of all continuous (equivalently, bounded) linear functionals from $\mathcal{H}$ into $\mathbb{R}$ or $\mathbb{C}$, the scalars. The Riesz representation theorem says that if $g \in \mathcal{H}$, then the linear functional ϕ_g, which depends on that fixed g and is defined by

$$\phi_g(f) =< f, g >, \text{ for all } f \in \mathcal{H}$$

is an element of $\mathcal{H}^*$. Moreover, every element of $\mathcal{H}^*$ can be written uniquely in this form. This is a powerful tool, allowing one to write the value of any bounded, linear functional as an inner product of elements of $\mathcal{H}$.

Now we can put some of these notions to use in order to define adjoints, symmetric operators, and selfadjoint operators. Let A be a densely defined linear operator on $\mathcal{H}$. Let $g \in \mathcal{H}$ and suppose there is a constant $K_g > 0$, which depends on g, such that for all $f \in \mathfrak{D}(A)$,

$$| < Af, g > | \leq K_g ||f||.$$

Then we say $g \in \mathfrak{D}(A^*)$ and $A^*g = h$, where, by the Riesz Representation Theorem, h exists, is unique, and

$$\langle Af, g \rangle = \langle f, h \rangle = \langle f, A^*g \rangle$$

for all $f \in \mathfrak{D}(A)$. This defines $\mathfrak{D}(A^*)$ and A^*, which we call the adjoint of A. If A^* is an extension of A, i.e., if

$$\langle Af, g \rangle = \langle f, Ag \rangle$$

for all $f, g \in \mathfrak{D}(A)$ (iff $\mathfrak{D}(A^*) \supset \mathfrak{D}(A)$ and $A^*f = Af$ for all $f \in \mathfrak{D}(A)$), then A is called **symmetric**. The operator A is called **selfadjoint** if it is symmetric and $\mathfrak{D}(A^*) = \mathfrak{D}(A)$, implying $A = A^*$. The above A is assumed to be densely defined.

The Spectral Theorem

In finite dimensional linear algebra, diagonal operators are among the simplest operators one can hope to work with. We want to extend the notion of diagonalizing a linear operator (which is only valid in finite dimensional vector spaces, where linear operators can be represented by finite matrices) to infinite dimensional vector spaces. Such "infinite dimensional diagonal matrices" are examples of **multiplication operators**.

Let (Ω, Σ, μ) be a measure space and let

$$m : \Omega \to \mathbb{R}$$

be Σ-measurable. Define the operator of multiplication by m, M_m, as

$$(M_m g)(x) = m(x)g(x)$$

$$\mathfrak{D}(M_m) = \{g \in L^2(\Omega, \Sigma, \mu) : mg \in L^2(\Omega, \Sigma, \mu)\}.$$

Then M_m is selfadjoint on $L^2(\Omega, \Sigma, \mu)$.[1]

A linear operator A on $\mathcal{H}$ is said to be unitarily equivalent to another linear operator B on a Hilbert space $\mathcal{K}$ if and only if there exists a unitary operator $U : \mathcal{H} \to \mathcal{K}$ such that $Af = U^{-1}BUf$ for all $f \in \mathfrak{D}(A)$. Note that since unitary operators preserve inner products, if an operator is unitarily equivalent to a selfadjoint operator, then it is necessarily selfadjoint.

The spectral theorem is one of the most powerful theorems in operator theory. It states a connection between selfadjoint operators and multiplication operators.

Spectral Theorem for Selfadjoint Operators 1. *Every selfadjoint operator is unitarily equivalent to a multiplication operator by a real valued measurable function on some L^2 space.*

A proof of this important theorem is found in [1]. This allows us to treat selfadjoint operators as if they were real valued functions. Most notably we can have functions

[1]Note that by the generalized Hölder inequality, any m in $L^\infty(\Omega, \Sigma, \mu)$ is sufficent to have mg in L^2, and thus $\mathfrak{D}(M_m) = L^2(\Omega, \Sigma, \mu)$.

of operators. That is, if S is a selfadjoint operator, and

$$S = U^{-1} M_m U$$

and if $f : \mathfrak{D}(f) \subset \mathbb{R} \to \mathbb{C}$ is Borel measurable, we define

$$f(S) = U^{-1} M_{f(m)} U$$

provided that the domain of f contains the essential range of m, so that $f(m)$ makes sense almost everywhere. $f(S)$ is self-adjoint in f if real valued, $f(S)$ is unitary if f takes values from the unit circle of $\mathbb{C}$, and $f(S)$ is normal if f is complex valued.

Basics of Modern Physics

The Special Theory of Relativity

Special relativity grew out of contradicting predictions from the two great paradigms of classical physics, Newton's theory of classical mechanics and Maxwell's theory of electromagnetism. They disagreed over how a moving observer would measure the speed of light. To give an example, if a car moving with uniform velocity on a straight road turned on its headlights, how fast would a person standing by the side of the road measure the light as going? If the speed of the car is s and the speed of light if c, then according to Newtonian Mechanics, we would add their speeds like any other objects. The observer would measure the speed of the light as $s + c$, the speed of the car plus the speed of light. Classical Electromagnetism states that both will measure the speed as c, the speed of light.

In 1905, Albert Einstein published a paper resolving the conflict. This introduced

the special theory of relativity. It is based on two postulates. The first is that there is no preferred frame of reference; any physical laws that are valid in one frame is valid for all frames. The second postulate is that the speed of light in a vaccuum is the same for all observers in uniform motion; it was a confirmation of the prediction of classical electrodynamics. The second postulate was experimentally verified by the Michelson-Morley experiment in 1887, which sparked a line of research that eventually led to general relativity. This was a paradigm shift in the world of physics.

This is called special relativity as it is a specific case of general relativity. It says that as certain gravitational effects go to zero, general relativity reduces to special relativity. More precisely, in general relativity, gravity is explained by the curvature of spacetime.

Quantum Mechanics

Special relativity was not the only paradigm shift that occured at the beginning of the twentieth century. Quantum mechanics began likewise with a contradiction. It was known that light had a wave nature. This was well established experimentally, as well as in agreement with Maxwell's theory of electromagnetism. In the late 19^{th} and early 20^{th} centuries, research was being done into the process known as the photoelectric effect. This phenomenon occurs when light hits a piece of metal and the piece of metal emits an electron into the vaccum. According to the prevailing theory of optics, the amount of electrons should depend on the intensity of the light. Instead, experiments found that it depended only on the frequency of the light shone

upon the metal.

Once again, in 1905, Albert Einstein published a paper presenting a theory that explains the photoelectric effect. He realized that light came not in a continous wave, but in "wave packets" with energy depending on the frequency of the wave packet. This paper earned Einstein his Nobel Prize. Through the work of Einstein, as well as other scientists such as Bohr, Dirac, Planck, Heisenberg and Schrödinger, the quantum theory of matter was developed. Under this theory quantum particles have some properties of classical particles and some properties of waves.

One of the central objects in quantum mechanics is the wavefuction of the system. It is used to predict certain observable quantities, such as position and momentum. The Schrödinger equation is

$$i\hbar\frac{\partial\psi_0}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi_0 + V(x)\psi_0 \tag{1}$$

for $x \in \mathbb{R}^{\mathrm{n}}, t \in \mathbb{R}$. Usually n is 3 times the number of particles in the system.

Here $\psi(x,t)$ represents the wavefunction at time t, a function of x which is a unit vector in $L^2(\mathbb{R}^n)$. $V(x)$ represents the potential, a real valued function of the spatial coordinates. This allows us to make highly accurate predictions about particles and systems of particles.

A wave function is a unit vector in a Hilbert space $\mathcal{H}$. In general the Schrödinger equation is written in the form of $i\hbar\frac{\partial\psi}{\partial t} = H\psi$, where H(which is selfadjoint) is called the Hamiltonian. In (1) this corresponds to the energy of the system and is written

$$H = -\frac{\hbar^2}{2m}\Delta + V(x). \tag{2}$$

A partial interpretation of the wave function is that the position probability den-

sity is $|\psi(x,t)|^2$. Thus

$$\int_r |\psi(x,t)|^2 \mathrm{d}x$$

is the probability of finding the position of a quantum system state ψ in $r \subset \mathbb{R}^n$ at time t. Similarly the momentum probability density is $|\hat{\psi}(\xi,t)|^2$, where $\hat{\psi}$ is the Fourier transform of ψ for a fixed t.

Combining the Two: Relativistic Quantum Mechanics

Quantum mechanics was very accurate for small objects and special relativity was very accurate for things moving quickly (as in, close to the speed of light). The problem was that neither was able to predict accurately small objects moving at a sizable fraction of the speed of light. This was a major problem, as small things tend to move very fast. The first attempt at this instead of using the nonrelativistic energy relation used to derive the Schrödinger equation, used the relativistic energy relation. This led to the Klein-Gordon Equation,

$$\frac{\partial^2 \psi}{\partial t^2} = c^2 \Delta \psi - \frac{m^2 c^4}{\hbar^2} \psi + \frac{2mc^2}{\hbar^2} V(x) \psi, \tag{3}$$

where c is the speed of light.

The Klein-Gordon Equation

Here we will show that the solution to the Klein-Gordon equation reduces to the solution for the Schrödinger equation in the limit when c tends to infinity. First consider the general Schrödinger equation initial value problem:

$$i\hbar \frac{\partial \psi_0}{\partial t} = H\psi_0, \tag{4}$$

$$\psi_0(x,0) = f(x).$$

Here H is given by (2). This yields the unique solution

$$\psi_0 = e^{\frac{-itH}{\hbar}} f(x).$$

This is the solution that we are trying to recover from the Klein-Gordon equation. In the case of (1), H will be $\frac{-\hbar^2}{2m}\Delta + V(x)$, but H can be any other selfadjoint operators as well.

Consider the initial value problem for the Klein-Gordon equation,

$$\frac{\partial^2\psi}{\partial t^2} = c^2\Delta\psi - \frac{m^2c^4}{\hbar^2}\psi + \frac{2mc^2}{\hbar^2}V(x)\psi, \tag{5}$$

$$\psi(x,0) = f(x),$$

$$\frac{\partial\psi}{\partial t}(x,0) = \tilde{g_\epsilon}(x),$$

Where we allow $\tilde{g_\epsilon}(x)$ to depend on c and on f, in a way that we will specify later. We seek a solution of the form $\psi = e^{\frac{-imc^2t}{\hbar}}\psi_\epsilon$, which we can identify with ψ_ϵ, as they have the same position and momentum probability densities at time t, so we can consider them equivalent. More importantly, the vectors $\psi(x,t)$ and $e^{i\theta(t)}\psi(x,t)$ are identified for any real valued function $\theta(t)$ Then (5) reduces to

$$\frac{1}{c^2}\frac{\partial^2\psi_\epsilon}{\partial t^2} - \frac{2mi}{\hbar}\frac{\partial\psi_\epsilon}{\partial t} + (-\Delta + \frac{2m}{\hbar}V(x))\psi_\epsilon = 0, \tag{6}$$

$$\psi_\epsilon(x,0) = f,$$

$$\frac{\partial\psi_\epsilon}{\partial t}(x,0) = g_\epsilon.$$

Where $g_\epsilon = \frac{imc^2}{\hbar}f + \tilde{g_\epsilon}$. Let $\epsilon = \frac{1}{c^2}$, $A = \frac{2m}{\hbar}I$, which we identify with $\frac{2m}{\hbar} > 0$, $B = -\Delta + \frac{2m}{\hbar}V(x)$, so the differential equation becomes

$$\epsilon\frac{\partial^2\psi_\epsilon}{\partial t^2} - iA\frac{\partial\psi_\epsilon}{\partial t} + B\psi_\epsilon = 0, \tag{7}$$

where A and B are selfadjoint ($A = A^*$ and $B = B^*$), and they commute (in the

sense that $[(\lambda I - A)^{-1}, (\mu I - B)^{-1}] = 0$ for all λ, $\mu \in \mathbb{C}\backslash\mathbb{R}$). Since $B = -\Delta + V(x)$ is a nonnegative injective selfadjoint operator on $L^2(\mathbb{R}^n)$ under some restrictions on $V(x)$, we can write $B = S^2$, with $S = S^* \geq 0$, S injective. So the differential equation problem becomes

$$\epsilon\psi_\epsilon''(t) - iA\psi_\epsilon'(t) + S^2\psi_\epsilon(t) = 0 \tag{8}$$

$$\psi_\epsilon(0) = f \tag{9}$$

$$\psi_\epsilon'(0) = g_\epsilon. \tag{10}$$

Since A and $B = S^2$ are commuting selfadjoint operators, the Spectral Theorem tells us that we can represent them by real multiplication operators on the same concrete L^2 space, $L^2(\Omega, \Sigma, \mu)$. This allows us to treat A and $B = S^2$ as real numbers, and thus we can treat our ordinary differential equation in a Hilbert space as a constant coefficient ordinary differential equation with a parameter $\omega \in \Omega$.

The solution to (8) under (9) and (10) is given by

$$\psi_\epsilon(t) = e^{t\Lambda_{+,\epsilon}}\alpha_\epsilon + e^{t\Lambda_{-,\epsilon}}\beta_\epsilon,$$

the "abstract d'Alembert formula" for the unique solution ψ_ϵ. Applying the initial conditions yields

$$\psi_\epsilon(0) = f = \alpha_\epsilon + \beta_\epsilon, \tag{11}$$

and

$$\psi_\epsilon'(0) = g_\epsilon = \Lambda_{+,\epsilon}\alpha_\epsilon + \Lambda_{-,\epsilon}\beta_\epsilon. \tag{12}$$

Now let's look at $\Lambda_{\pm,\epsilon}$ in more detail:

$$\Lambda_{\pm,\epsilon} = \frac{iA \pm \sqrt{-A^2 - 4\epsilon S^2}}{2\epsilon} \tag{13}$$

$$= i\left(\frac{A \pm \sqrt{A^2 + 4\epsilon S^2}}{2\epsilon}\right)$$
$$= i\left(\frac{A \pm A\sqrt{A^{-2}S^2 + 4\epsilon I}}{2\epsilon}\right)$$

Positive selfadjoint operators have many selfadjoint square roots. For instance, I on $L^2(\Omega, \Sigma, \mu)$ can be written as S^2 where for all Γ in $\Sigma, S = M_s$ and $s = \mathbb{1}_\Gamma - \mathbb{1}_{\Omega\setminus\Gamma}$. Let $\tilde{\mathrm{A}} = \tilde{\mathrm{A}}^* \geq 0$ on $\mathcal{H}$;, then there is a unique $S = S^* \geq 0$ such that $\tilde{\mathrm{A}} = S^2$. We define this S to be $S = \tilde{\mathrm{A}}^{\frac{1}{2}} = \sqrt{\tilde{\mathrm{A}}}$.

For $\tilde{\mathrm{A}}$ selfadjoint on $\mathcal{H}$ we write $\tilde{\mathrm{A}}$ as

$$\tilde{\mathrm{A}} = \tilde{\mathrm{A}}_+ - \tilde{\mathrm{A}}_- = \tilde{\mathrm{A}}\mathbb{1}_{[0,\infty)}(\tilde{\mathrm{A}}) + \tilde{\mathrm{A}}\mathbb{1}_{(-\infty,0)}(\tilde{\mathrm{A}}) \tag{14}$$

using the functional calculus. Let $\tilde{\mathrm{A}}_\pm$ be nonnegative and selfadjoint with $\tilde{\mathrm{A}}_+\tilde{\mathrm{A}}_- = \tilde{\mathrm{A}}_-\tilde{\mathrm{A}}_+ = 0$. Define

$$\tilde{\mathrm{A}}^{\frac{1}{2}} = \tilde{\mathrm{A}}_+^{\frac{1}{2}} + i\tilde{\mathrm{A}}_-^{\frac{1}{2}} = \sqrt{\tilde{\mathrm{A}}}. \tag{15}$$

This is the way to define the canonical square root of $\tilde{\mathrm{A}}$.

Recall that if a function $u(\epsilon)$ is in $C^k[0, b]$ then Taylor's theorem allows us to approximate $u(\epsilon)$ for all $\epsilon \in [0, b]$ as

$$u(\epsilon) = u(0) + u'(0)\epsilon + \frac{u''(0)\epsilon^2}{2} \ldots \frac{u^{(k-1)}(0)\epsilon^{k-1}}{(k-1)!} + R_k(\epsilon) \tag{16}$$

and

$$R_k(\epsilon) = \int_0^\epsilon \frac{(s-\epsilon)^{k-1}}{(k-1)!} u^{(k)}(s)\mathrm{d}s$$

But note that the integral is bounded above by $\frac{\epsilon^k}{k!}\|u^{(k)}\|_\infty$ where the norm is taken in the space of continous functions. This means if $W \in C^2([0, b], \mathcal{H})$ for some $b > 0$, then we can use Taylor's theorem to order 1 to write it as[2]

[2]Here $o(\epsilon)$ means a vector in $\mathcal{H}$ such that $\frac{o(\epsilon)}{\epsilon} \to 0$ as $\epsilon \to 0$. This is actually a slight abuse of notation, as $o(\epsilon)$ is actually a set of functions of ϵ, here we are using it to represent a member of that set

$$W(\epsilon) = W(0) + \epsilon W'(\epsilon) + o(\epsilon).$$

or more precisely for $0 < \epsilon < b$,

$$W(\epsilon) = W(0) + \epsilon W'(\epsilon) + R_2(\epsilon),$$

$$R_2(\epsilon) = \int_0^\epsilon \frac{(s-\epsilon)}{2} W''(s)\mathrm{d}s$$

and[3]

$$R_2(\epsilon) < \frac{\epsilon^2}{2}\|W''\|_\infty < C_W\epsilon^2 = O(\epsilon^2) = o(\epsilon)$$

To find our specific solution (8) we must first we must solve for α_ϵ and β_ϵ as functions of f and g. Apply $\Lambda_{-,\epsilon}$ to (11) and subtract (12).

$$\Lambda_{-,\epsilon} f - g_\epsilon = (\Lambda_{-,\epsilon} - \Lambda_{+,\epsilon})\alpha_\epsilon$$

hence

$$\alpha_\epsilon = (\Lambda_{-,\epsilon} - \Lambda_{+,\epsilon})^{-1}(\Lambda_{-,\epsilon} f - g_\epsilon) \tag{17}$$

$$\beta_\epsilon = f - \alpha_\epsilon$$

Using Taylor's formula we see that

$$\Lambda_+ = \frac{i}{2\epsilon}(A + A\sqrt{4\epsilon I + A^{-2}S^2}) = \frac{iA}{2\epsilon}(I + 4A^{-2}\epsilon S^2 + o(\epsilon)) \tag{18}$$

$$\Lambda_- = \frac{i}{2\epsilon}(A - A\sqrt{4\epsilon I + A^{-2}S^2}) = iA(A^{-2}S^2 + \frac{o(\epsilon)}{2\epsilon}) \tag{19}$$

Let h be a fixed vector in $\mathcal{H}$, and let

$$\bar{\alpha}_\epsilon = i\epsilon S^{-1}(A^2S^{-2} + 4\epsilon I)^{-\frac{1}{2}}h. \tag{20}$$

Recall that we can write (18) as

[3] $O(\epsilon)$ means that for every ϵ there exists an M such that $\frac{O(\epsilon)}{\epsilon} < M$ as $\epsilon \to 0$

$$\bar{\alpha}(\epsilon) = \bar{\alpha}(0) + \epsilon\bar{\alpha}'(0) + o(\epsilon)$$

$$\bar{\alpha}'(\epsilon) = iS^{-1}(A^2S^{-2} + 4\epsilon I)^{-\frac{1}{2}}h - 2iS^{-1}(A^2S^{-2} + 4\epsilon I)^{-\frac{3}{2}}h,$$

so $\bar{\alpha}_\epsilon(0) = 0$, and $\bar{\alpha}'_\epsilon(0) = iA^{-1}h$. For this we need $\bar{\alpha}_\epsilon \in C^2([0,b],\mathcal{H})$, which follows if we assume $h \in \mathcal{D}(S^{\frac{3}{2}})$. So

$$\bar{\alpha}'(\epsilon) = i\epsilon A^{-1}h + o(\epsilon, h)$$

where $o(\epsilon, h)$ is a vector in the set $o(\epsilon)$ and is dependent upon h. Thus as $\epsilon \to 0$,

$$\bar{\alpha}(\epsilon) \to 0, \tag{21}$$

$$\bar{\beta}(\epsilon) = f - \bar{\alpha}(\epsilon) \to f. \tag{22}$$

In fact $\bar{\beta} - f = o(\epsilon)$. Recall that by (13)

$$\Lambda_\pm(\epsilon)h = \frac{i}{2\epsilon}(A \pm A\sqrt{A^{-2}S^2 + 4\epsilon I})h = \frac{Ai}{2\epsilon}(I \pm \sqrt{I + 4A^{-2}\epsilon S^2})h.$$

Then Taylor's theorem gives us that

$$\sqrt{1+x} = 1 + \frac{x}{2} + o(x)$$

so

$$\frac{Ai}{2\epsilon}(I + \sqrt{I + 4A^{-2}\epsilon S^2})h = \frac{Ai}{2\epsilon}((I + 4A^{-2}\epsilon S^2)h + o(\epsilon, h)) = \frac{Ai}{2\epsilon}h + iA^{-1}S^2h + o(1, h).$$

So, as $\epsilon \to 0$,

$$\Lambda_-(\epsilon)h = iA^{-1}S^2h + o(1,h) \to iA^{-1}S^2h. \tag{23}$$

On the other hand for any non-zero h,

$$\|\Lambda_+(\epsilon)h\| \to \infty \tag{24}$$

If we specify that $g_\epsilon = \Lambda_{-,\epsilon}f$ (and thus the original initial data $\tilde{g}_\epsilon$ is equal to $(\Lambda_{-,\epsilon} - \frac{im}{\hbar\epsilon})f$), then $\alpha_\epsilon = 0$, and the above norm is irrelevant. Thus the solution to the problem is

$$\psi_\epsilon(t) = e^{t\Lambda_{-,\epsilon}} f,$$

where $f \in \mathcal{D}(S^2)$ and

$$\Lambda_-(\epsilon) = \frac{iA}{2\epsilon}(I - (I + 4A^{-2}\epsilon S^2)^{-\frac{1}{2}})$$

Let $W(\epsilon, t) = \psi_\epsilon(t)$, then

$$W(\epsilon, t) = W(0, t) + \epsilon W'(0, t) + o(\epsilon, f, t)$$

where the $o(\epsilon, f, t)$ term is uniform for compact intervals of t. That is for a fixed $f \in \mathcal{D}$ and $T > 0$

$$\sup_{|t| \le T} (\frac{\|o(\epsilon, f, t)\|}{\epsilon}) \to 0 \text{ as } \epsilon \to 0. \tag{25}$$

Formally

$$\Lambda_-(\epsilon) = \frac{iA}{2\epsilon}(I - (I + 4A^{-2}\epsilon S^2)^{-\frac{1}{2}}) = \frac{iA}{2\epsilon}(I - (I + 2A^{-2}\epsilon S^2) + o(\epsilon) = -iA^{-1}S^2 + o(1)$$

and

$$e^{t\Lambda_{-,\epsilon}} f = e^{-iA^{-1}tS^2} f(1 + o(1)).$$

Let

$$u(t) = e^{-iA^{-1}tS^2} f \tag{26}$$

which is the unique solution of

$$iA\frac{\mathrm{d}u}{\mathrm{d}t} = S^2 u,$$

$$u(0) = f.$$

This is the desired limiting Schrödinger equation. To prove this, we once again use Taylor's Formula

$$W(\epsilon, t) = W(0, t) + \epsilon W'(0, t) + o(\epsilon, f, t),$$

where $w(0,t) = u(t)$ (see (26)) and the little o term satisfies (24), and

$$W'(\epsilon,t) = \frac{\mathrm{d}}{\mathrm{d}\epsilon}(e^{t\Lambda_-(\epsilon)}f) \tag{27}$$

$$= e^{t\Lambda_-(\epsilon)}\frac{\mathrm{d}}{\mathrm{d}\epsilon}(\Lambda_-(\epsilon)f)$$

$$= e^{t\Lambda_-(\epsilon)}(\frac{-iA}{2\epsilon^2}(I-(I+4A^{-2}\epsilon S^2)^{-\frac{1}{2}})f + -\frac{iA}{\epsilon}((I+4A^{-2}\epsilon S^2)^{-\frac{1}{2}})A^{-2}\epsilon S^2 f)$$

We now need an extra term in the taylor expansion. Assuming $F \in C^2([0,b])$ we can write it as

$$F(x) = F(0) + F'(0)x + \frac{F''(0)}{2}x^2 + o(x^2)$$

thus

$$(I-(I+4A^{-2}\epsilon S^2)^{\frac{1}{2}})f = \epsilon 2S^2A^{-2}f - \epsilon^2 4A^{-4}S^4 f + o(\epsilon^2, f)$$

We assume $f \in \mathcal{D}(S^4)$, and thus by (27)

$$\begin{aligned} W'(\epsilon,t) &= e^{t\Lambda_-(\epsilon)}(\frac{-iA}{2\epsilon^2}(\epsilon 2S^2A^{-2}f - \epsilon^2 4A^{-4}S^4 f + o(\epsilon^2,f)) \\ &\quad - \frac{iA}{2\epsilon}(-2A^{-2}S^2(f + 2A^{-2}S^2\epsilon f + o(\epsilon,f)))) \\ &= e^{t\Lambda_-(\epsilon)}4iA^{-3}S^4 f + o(1,f,t) \end{aligned}$$

where the little o term satisfies (25). Because the two terms (besides those containing $S^4 f$) containing $S^2 f$ add up to zero, it follows that

$$\begin{aligned} e^{t\Lambda_-(\epsilon)}f = \psi_\epsilon(t) &= W(\epsilon,t) \\ &= W(0,t) + \epsilon W'(0,t) + o(\epsilon,f,t) \\ &= u(t) + o(1) = e^{-itA^{-1}S^2}f + o(1). \end{aligned}$$

where the error term, satisfying (25) is

$$o(1) = \epsilon e^{t\Lambda_-(\epsilon)}4iA^{-3}S^4 f + o(\epsilon,f,t)$$

We have now proved the following theorem

Theorem 1. *Let ψ be the unique solution of the Klein-Gordon Equation,*

$$\frac{\partial^2\psi}{\partial t^2} = c^2\Delta\psi - \frac{m^2c^4}{\hbar^2}\psi + \frac{2mc^2}{\hbar^2}V(x)\psi, \tag{28}$$

$$\psi(x,0) = f(x),$$

$$\frac{\partial\psi}{\partial t}(x,0) = g(x).$$

where $f, g \in H^4(\mathbb{R}^n)$, the usual Sobolev space, and $g = (\Lambda_{-,\epsilon} - \frac{imc^2}{\hbar})f$. Let $\epsilon = \frac{1}{c^2}$, and let

$$\psi = e^{\frac{imc^2t}{\hbar}}\psi_\epsilon$$

Then ψ_ϵ is the unique solution to

$$\epsilon\frac{\partial^2\psi_\epsilon}{\partial t^2} - iA\frac{\partial\psi_\epsilon}{\partial t} + B\psi_\epsilon = 0, \tag{29}$$

$$\psi_\epsilon(x,0) = f(x),$$

$$\frac{\partial\psi_\epsilon}{\partial t}(x,0) = \frac{im}{\epsilon\hbar}f + g$$

In the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^n)$, where $f, g \in \mathcal{D}(B^2), A = \frac{2m}{\hbar}, B = -\Delta + \frac{2m}{\hbar}V(x)$ and V is a potential so that B is self-adjoint, injective, and $\mathcal{D}(B) = H^2(\mathbb{R}^n)$. Then

$$u(t) = \lim_{\epsilon\to 0}\psi_\epsilon(t).$$

Exists in $\mathcal{H}$, uniformly for bounded intervals of t. And $u(t) = e^{-A^{-1}tB}f$ is the unique solution to the Schrödinger equation

$$iA\frac{\mathrm{d}u}{\mathrm{d}t} = Bu \tag{30}$$

with initial conditions $u(0) = f \in \mathcal{D}(B^2)$. The approximation has the precise estimate given by

$$\|u(t) - \psi_\epsilon(t)\| \leq 4\|A^{-3}B^2f\|\epsilon + o(\epsilon)$$

where the $o(\epsilon)$ term us bounded by $C(T,f)\epsilon^2$ where the constant depends on f and holds for $t \in [-T,T]$.

Here we stated the theorem in its full generality. The usual statement starts from (28) without mentioning initial conditions and then passes to (29), for which initial conditions are denoted by F and G (or by F_ϵ and G_ϵ with $F_\epsilon \to F$ and $G_\epsilon \to G$, then ψ_ϵ converges to u, the solution of (30) with $u(0) = F$, independently of G.

References

[1] Jerome A. Goldstein. *Semigroups of Linear Operators and Applications*. Oxford University Press, 1985.

[2] Andrew Schoene. On the nonrelativistic limit of the Klein-Gordon and Dirac equations. *Journal of Mathematical Analysis and Applications*, 71:36–47, 1979.