Fall 2015

Approximation methods in relativistic eigenvalue perturbation theory

Jonathan Howard Noble

APPROXIMATION METHODS IN RELATIVISTIC EIGENVALUE PERTURBATION THEORY

by

JONATHAN HOWARD NOBLE

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

PHYSICS

2015

Approved by

Dr. Ulrich D. Jentschura, Advisor
Dr. Paul E. Parris
Dr. Barbara N. Hale
Dr. Don H. Madison
Dr. Peter J. Mohr

ABSTRACT

In this dissertation, three questions, concerning approximation methods for the eigenvalues of quantum mechanical systems, are investigated: (i) What is a pseudo–Hermitian Hamiltonian, and how can its eigenvalues be approximated via numerical calculations? This is a fairly broad topic, and the scope of the investigation is narrowed by focusing on a subgroup of pseudo–Hermitian operators, namely, $\mathcal{PT}$–symmetric operators. Within a numerical approach, one projects a $\mathcal{PT}$–symmetric Hamiltonian onto an appropriate basis, and uses a straightforward two–step algorithm to diagonalize the resulting matrix, leading to numerically approximated eigenvalues. (ii) Within an analytic *ansatz*, how can a relativistic Dirac Hamiltonian be decoupled into particle and antiparticle degrees of freedom, in appropriate kinematic limits? One possible answer is the Foldy–Wouthuysen transform; however, there are alternative methods which seem to have some advantages over the time–tested approach. One such method is investigated by applying both the traditional Foldy–Wouthuysen transform and the "chiral" Foldy–Wouthuysen transform to a number of Dirac Hamiltonians, including the central-field Hamiltonian for a gravitationally bound system; namely, the Dirac-(Einstein-)Schwarzschild Hamiltonian, which requires the formalism of general relativity. (iii) Are there are pseudo–Hermitian variants of Dirac Hamiltonians that can be approximated using a decoupling transformation? The tachyonic Dirac Hamiltonian, which describes faster-than-light spin-1/2 particles, is γ^5–Hermitian, i.e., pseudo-Hermitian. Superluminal particles remain faster than light upon a Lorentz transformation, and hence, the Foldy–Wouthuysen program is unsuited for this case. Thus, inspired by the Foldy–Wouthuysen program, a decoupling transform in the ultrarelativistic limit is proposed, which is applicable to both sub- and superluminal particles.

ACKNOWLEDGMENTS

First and foremost, it is my pleasure to thank my dissertation advisor, Dr. Ulrich D. Jentschura, for his constant encouragement, as well as his intensive and comprehensive supervision during the completion of this dissertation. I am amazed at the time and effort he has been willing to put into helping me. He helped make the time I spent here, at the Missouri University of Science and Technology in Rolla (2010–2015), the most productive of my life.

It is also my pleasure to thank my parents for their love, encouragement, and support. Keeping in contact with my family, and friends back in my home state of Maryland, has helped me focus on the physics in Rolla.

Let me also acknowledge the administrative support by the physics department's specialists, Pam Crabtree and Janice Gargus.

This research has been supported, both in the summer months and also during a number of research semesters, by the National Science Foundation, within grants PHY—8555454, PHY–1068547 and PHY–1403973. In times of tight research budgets, this valued support is acknowledged with gratitude.

Finally, I would like to acknowledge the efforts of the dissertation evaluation committee, Dr. P. E. Parris, Dr. B. N. Hale, Dr. D. H. Madison, and Dr. P. J. Mohr in judging this dissertation.

RELATED PUBLICATIONS

This dissertation is based on the work done in a number of publications produced by Jonathan H. Noble and Dr. Ulrich D. Jentschura. The same results are presented in this dissertation, with a greater emphasis on how said results were obtained. These publications are as follows:

Part I

• J. H. Noble, M. Lubasch, and U. D. Jentschura, *Generalized Householder Transformations for the Complex Symmetric Eigenvalue Problem*, Eur. Phys. J. Plus **128**, 93 (2013).

Part II

• U. D. Jentschura, and J. H. Noble, *Nonrelativistic Limit of the Dirac–Schwarzschild Hamiltonian: Gravitational Zitterbewegung and Gravitational Spin–Orbit Coupling*, Phys. Rev. A **88**, 022121 (1013).

• U. D. Jentschura, and J. H. Noble, *Foldy–Wouthuysen Transformation, Scalar Potentials and Gravity*, J. Phys. A **47**, 045402 (2014).

Part III

• J. H. Noble, and U. D. Jentschura, *Ultrarelativistic decoupling transformation for generalized Dirac equations*, Phys. Rev. A **92**, 012101 (2015).

Additionally there are published works, which are related to the work presented in Part II, the results of which we do not examine in this dissertation.

• U. D. Jentschura, J. H. Noble and I. Nándori, *Gravitational Interactions and Fine-Structure Constant*, Bled Workshops in Physics **15**, 115 (2014).

• J. H. Noble, and U. D. Jentschura, *Dirac equations with confining potentials*, Int. J. Mod. Phys. A **30**, 1550002 (2015).

TABLE OF CONTENTS

APPENDICES

LIST OF ILLUSTRATIONS

Figure Page

LIST OF TABLES

ENHANCED ABSTRACT, MOTIVATION, AND INTRODUCTION

It has long been established that in quantum mechanics a Hamiltonian must be Hermitian, or, more precisely speaking, essentially self-adjoint, guaranteeing that the resulting spectrum will be real and that the canonical inner product of the wave–functions is conserved. There are accepted exceptions to this rule. An unstable particle, for example, has a complex energy spectrum where the imaginary part indicates the decay rate of a given energy state. In many cases, a physical system cannot be described any more on the basis of a fully Hermitian Hamiltonian once quantum fluctuations are taken into account. Typically, a Hermitian Hamiltonian describes a closed system (i.e., there are no gain or loss terms); this restriction severely limits the kinds of systems which can be represented by this model. Pseudo–Hermiticity expands on the idea of Hermiticity and by redefining a number of properties; the generalized inner product is still conserved, however the spectrum is no longer guaranteed to be real [1]. The concept of $\mathcal{PT}$-symmetry, as first proposed by Bender and Boettcher [2], has many similarities to pseudo-Hermiticity, including the necessary generalization of the inner product. Most cases studied in the literature are both $\mathcal{PT}$–symmetric and pseudo–Hermitian (examples include [2–14]). As with pseudo–Hermiticity, $\mathcal{PT}$-symmetry allows for the conservation of a generalized inner product. It is relatively simple to show that the complex eigenvalues of a $\mathcal{PT}$–symmetric Hamiltonian come in complex conjugate pairs, while in the case of *exact* $\mathcal{PT}$–symmetry, the eigenvalues are real [15]. The difficulty arises when trying to prove that the $\mathcal{PT}$–symmetry of a given Hamiltonian is exact, as this generally requires the calculation of the corresponding wave–functions [3]. Consequently it was numerical evidence, rather than a formal proof that initially suggested that the eigenvalues of certain

$\mathcal{PT}$–symmetric Hamiltonians are real [2]. It has since been formally proven that specific classes of Hamiltonians are exactly $\mathcal{PT}$–symmetric (i.e., all of their eigenvalues are real) [16–18].

It has been shown that Hermitizing transforms can be applied to an exact $\mathcal{PT}$–symmetric Hamiltonian H, transforming it into a Hermitian Hamiltonian h, i.e., $\mathcal{PT}$–symmetry would then be equivalent to Hermiticity [19]. However, the transforms used turn out to be similarity transforms, which are necessarily not unitary, thus the equivalence is restricted to the eigenvalues [15]. By examining the metric of the imaginary cubic oscillator, $H = p^2 + \mathrm{i}\,x^3$, which is exactly $\mathcal{PT}$–symmetric [16, 17], Siegl and Krejčiřík found that H cannot be similar to a Hermitian Hamiltonian [21]. Furthermore it is possible to interpret a $\mathcal{PT}$-symmetric system as an open system in which the gain and loss are "in equilibrium". All this seems to point to the conclusion that $\mathcal{PT}$-symmetry is an independent concept in its own right. In an attempt to further this point of view, we perform a numerical analysis of the wave–functions of a $\mathcal{PT}$-symmetric Hamiltonian. In examining the properties of these wave–functions, one finds that although they have some similarities to Hermitian wave–functions, the nonvanishing imaginary contributions to the Hamiltonian (the "gain and loss" terms) imply that $\mathcal{PT}$-symmetry is in fact an independent concept. In order to perform these calculations we draw inspiration from the generalized inner product, and apply the concept to Householder matrices. This generalization constitutes the first step in a two step algorithm designed to numerically diagonalize complex symmetric matrices, which arise from the projection of $\mathcal{PT}$–symmetric Hamiltonians onto an approximately complete set of basis states, for example.

Complementing the numerical approach from part I of this dissertation, we next turn our attention to the Foldy–Wouthuysen transformation [22, 23], within an analytic *ansatz* focused on traditional, Hermitian Hamiltonians. We begin our investigation by reviewing the Foldy-Wouthuysen transformation, which is an iterative process, designed to approximate Dirac Hamiltonians in the non-relativistic limit. In reviewing how the transformation is performed, we apply it to a series of Dirac Hamiltonians, including the free Dirac Hamiltonian, the Dirac-Coulomb Hamiltonian, and the Dirac–Einstein–Schwarzschild Hamiltonian. We supplement the analysis by considering Dirac Hamiltonians with scalar potentials and the Dirac Hamiltonian in a non-inertial reference frame.

Because of its iterative and perturbative nature, the Foldy-Wouthuysen transformation can quickly become rather complicated. In many cases, the Foldy-Wouthuysen transformation cannot be carried out exactly (i.e., to all orders in the the momentum operators). Thus, an alternative "chiral Foldy-Wouthuysen" transformation has been proposed [24], which takes advantage of a number of deceptively appealing properties, including a subtle requirement that the chiral operator ($J = \mathrm{i}\,\gamma^5\,\beta$) commutes with the input Hamiltonian, which appear to lead to a simpler method of decoupling the particle and antiparticle degrees of freedom. We apply the chiral Foldy–Wouthuysen transform to the same Hamiltonians to which we applied the traditional Foldy–Wouthuysen transformation. By comparing the results from both methods, we find that the results are fundamentally different. Not only are there discrepancies between some of the prefactors, we additionally find that the chiral method does not conserve the parity of the system, nor the particle–antiparticle symmetry. We are left to conclude that the chiral Foldy–Wouthuysen approximation does not satisfy all consistency requirements for the decoupling of generalized Dirac Hamiltonians.

Finally, in part III of the dissertation, we combine the concept of pseudo–Hermiticity, introduced in part I, with the concept of a decoupling transformation, introduced in part II. Our focus is on fully relativistic, pseudo–Hermitian Hamiltonians within the framework of relativistic quantum mechanics. The so-called tachyonic Dirac equation, introduced by Chodos, Hauser and Kostelecky [25], when written in noncovariant form, transforms into a generalized Dirac Hamiltonian which exactly has the property of being γ^5–Hermitian (pseudo–Hermitian). Regardless of one's personal view as to the existence of tachyons, the Hamiltonians describing them have an interesting underlying structure. Furthermore, if tachyons do in fact exist, then the more we understand their physics, the more likely we are to be able to interpret conceivable experiments in the future.

The tachyonic Dirac Hamiltonian is first and foremost a free-particle Hamiltonian. However, generalizations are possible. Much like generalized subluminal Dirac equations, generalized superluminal Dirac equations describe the interactions of particles and antiparticles with external potentials. As for the subluminal case, the particle and antiparticle degrees of freedom in the Hamiltonians are coupled, making it difficult to interpret their interactions with the external potentials. Unlike the subluminal case, we cannot apply the Foldy–Wouthuysen transformation to a tachyonic Dirac Hamiltonian, as it is used to find the nonrelativistic limit, which is nonsensical in the case of tachyons (namely, tachyons remain superluminal upon a Lorentz transformation [26–29]). Instead, we must perform the opposite transformation and decouple the particle and antiparticle states in the high-energy, ultrarelativistic limit. To that end, we find it advantageous to first transform the Hamiltonian into the Weyl basis before any such transformation can be applied. Once this has been accomplished, we find that there is indeed an exact ultrarelativistic decoupling transformation for both free tardyons and free tachyons. Like the Foldy–Wouthuysen transform, the ultrarelativistic decoupling transform requires a perturbative approach when applied to generalized Dirac Hamiltonians. Again, of particular interest is the case of gravitational coupling, for both sub- and superluminal particles. In both cases we find that there is particle–antiparticle symmetry, meaning that particles and antiparticles are affected by gravity in the same way (e.g., they are both attracted toward a gravitational center). Additionally, starting from the ultrarelativistic limit, we find that the leading-order gravitational effects are identical for both tardyons and tachyons, while higher-order corrections reveal differences between how the two interact with gravity. The somewhat surprising result is that tachyons are attracted by gravity in the ultrarelativistic limit, much like a beam of light which is bent toward the center of a massive gravitational central potential. This is somewhat contrary to the classical result, which states that tachyons are repulsed by gravity [30].

To summarize, once more, for completeness: In part I, we investigate pseudo–Hermiticity and $\mathcal{PT}$–symmetry, and make a case that they are indeed independent concepts, and not a variation on Hermiticity. Additionally, we draw inspiration from the underlying mathematical structures of these Hamiltonians, and describe a matrix diagonalization algorithm designed with $\mathcal{PT}$-symmetric Hamiltonians, which is used to calculate both the eigenvalues as well as the eigenstates (wave–functions). In part II, we look at a number of example of generalized Dirac Hamiltonians, and apply both the traditional, as well as the chiral, Foldy–Wouthuysen transforms. The two methods produce different results, and the pitfalls of the chiral method are discussed. Additionally, the study produces some new results, including the nonrelativistic corrections to the Dirac–Einstein–Schwarzschild Hamiltonian, and the associated transition current. In part III we develop the ultrarelativistic decoupling transformation, in both its exact and perturbative form. We place special emphasis on the gravitationally coupled tardyon and tachyon, and arrive at a somewhat surprising conclusion regarding the latter. Finally conclusions are drawn in part IV.

Part I

Eigenvalues of Pseudo–Hermitian Hamiltonians

1. INTRODUCTION

In this part of the dissertation, we endeavor to answer the question: What is a pseudo–Hermitian Hamiltonian, and how can its eigenvalues be approximated based on numerical calculations? In one sense the first part of this question has already been answered, as Pauli defined pseudo–Hermitian Hamiltonians in 1943 [1]. While we know how pseudo–Hermiticity is defined, the concept covers a wide array of operators, including essentially self–adjoint Hamiltonians. We narrow the scope of our investigation to non–Hermitian $\mathcal{PT}$–symmetric Hamiltonians, as first discussed in [2]. Most of the research on $\mathcal{PT}$–symmetry is centered on Hamiltonians of the form $H = p^2/(2m) + V$ [2–14]. As we shall see, it is straightforward to show that a $\mathcal{PT}$–symmetric Hamiltonian that has the form of H is pseudo–Hermitian under parity.

The reality of the spectra of exact $\mathcal{PT}$–symmetric Hamiltonians has given rise to the idea that they are versions of Hermitian Hamiltonians, and Hermitizing transforms have been proposed, which transform an exact $\mathcal{PT}$–symmetric Hamiltonian into a Hermitian Hamiltonian [19, 31–34]. Under such a transform the spectrum is left unchanged. However, Hermitizing transforms generally require a perturbative calculation, leading to much more complicated, and potentially non–local, Hermitian Hamiltonians [15, 19]. Additionally, the transformation is necessarily non–unitary, and as such the relations between vector–spaces are not conserved [15, 21]. By numerically evaluating an exactly $\mathcal{PT}$–symmetric Hamiltonian, and examining the resulting wave–functions, we work to build an intuitive picture of $\mathcal{PT}$–symmetric wave–functions. In doing so, we note some fundamental differences between the Hermitian and $\mathcal{PT}$–symmetric pictures. Our observations suggest that it would be inconsistent to interpret the $\mathcal{PT}$–symmetric Hamiltonian as a “compact version” of a Hermitian

Hamiltonian. This in turn suggests that $\mathcal{PT}$–symmetry forms a class of Hamiltonians independent of Hermiticity, and any mapping onto, or identification with an, "equivalent" Hermitian Hamiltonian might seem a little contrived.

In order to lend a practical meaning to our investigations, we develop a numerical matrix diagonalization algorithm, which profits from the mathematical structure of the the pseudo–Hermitian Hamiltonians. This generalization turns out to be ideal when working with $\mathcal{PT}$–symmetric Hamiltonians projected onto appropriate basis sets. The resulting algorithm is best used when the entire spectrum of a densely populated complex symmetric matrix is desired. Using high precision arithmetic, we are able to obtain high precision energy approximations of $\mathcal{PT}$–symmetric Hamiltonians.

The organization is as follows: In chapter 2 an overview of the subject area is given. We conduct a basic review of self–adjoint Hamiltonians, and give a basic discussion of pseudo–Hermiticity and $\mathcal{PT}$–symmetry. We additionally investigate the importance of boundary conditions, and briefly discuss Hermitizing transforms. In chapter 3 we build an intuitive picture of the $\mathcal{PT}$–symmetric wave–functions, analogous to the Hermitian case. In chapter 4 we discuss the underlying mathematics of our matrix diagonalization algorithm. In chapter 5 we discuss an `FORTRAN` implementation of the algorithm, which is explicitly found in appendix A. Finally, conclusions are drawn in chapter 6.

2. OVERVIEW OF HERMITICITY AND PSEUDO–HERMITICITY

2.1. HERMITIAN OPERATORS

Given a Hamiltonian H, with wave–function $|\psi\rangle$, the time–dependent Schrödinger equation is given by

$$\mathrm{i}\frac{d}{dt}\,|\psi\rangle = H\;|\psi\rangle\;, \tag{2.1}$$

where H is traditionally assumed to be essentially self adjoint, i.e.,

$$H = H^{+}\,, \tag{2.2}$$

which guarantees that the eigenvalues (i.e., the energies) of H are real. This can easily be shown as follows; first let $H = H^{+}$, then

$$H|\psi\rangle = \lambda|\psi\rangle\,, \quad \langle\psi|H = \langle\psi|\lambda^{*}\,, \tag{2.3}$$

where $\lambda \in \mathbb{C}$, and $|\psi\rangle$ is an eigenvector of H. We then quickly find that

$$\langle\psi|H|\psi\rangle = \langle\psi|\left(H|\psi\rangle\right) = \lambda\langle\psi|\psi\rangle\,, \tag{2.4}$$

and

$$\langle\psi|H|\psi\rangle = \big(\langle\psi|H^{+}\big)\,|\psi\rangle = \lambda^{*}\langle\psi|\psi\rangle\,. \tag{2.5}$$

Comparing equations (2.4) and (2.5) we see $\lambda = \lambda^{*}$, and therefore $\lambda \in \mathbb{R}$.

Furthermore, the self–adjoint nature of H leads to the conservation of the inner product as follows. Let $|\psi\rangle$ and $|\phi\rangle$ be solutions of (2.1). We want to show that

$$\frac{d}{dt}\langle\psi|\phi\rangle = 0\,. \tag{2.6}$$

Using the time–dependent Schrödinger equation (2.1) we quickly find that

$$\frac{d}{dt}|\psi\rangle = |\frac{d}{dt}\psi\rangle = -\mathrm{i}\,H|\psi\rangle\,, \quad \frac{d}{dt}\langle\psi| = \langle\frac{d}{dt}\psi| = \left(\frac{d}{dt}|\psi\rangle\right)^{+} = \langle\psi|\mathrm{i}\,H^{+}\,. \tag{2.7}$$

Therefore

$$\begin{aligned}\frac{d}{dt}\langle\psi|\phi\rangle = \left(\frac{d}{dt}\langle\psi|\right)|\phi\rangle + \langle\psi|\left(\frac{d}{dt}|\phi\rangle\right) = \langle\psi|\mathrm{i}\,H^{+}|\phi\rangle + \langle\psi|\left(-\mathrm{i}\,H|\psi\rangle\right)\\ = \mathrm{i}\,\langle\psi|\left(H^{+} - H\right)|\phi\rangle = 0\,.\end{aligned} \tag{2.8}$$

Equation (2.8) shows that under time evolution, the inner product of the eigenfunctions of a Hermitian operator is conserved (i.e., it is invariant as it moves through time).

A typical nonrelativistic one-particle Hamiltonian can be written as

$$H = \frac{1}{2}\vec{p}^{\,2} + V(\vec{r},t)\,, \tag{2.9}$$

where $\vec{p}$ is the quantum mechanical momentum operator, and V is the potential energy. Under the condition of self–adjointness, it quickly follows that $V(\vec{r},t) = V^{*}(\vec{r},t)$, meaning that the potential is real. If we consider a time independent potential in one dimension (2.9) becomes

$$H = -\frac{1}{2}\frac{\partial^2}{\partial x^2} + V(x,t)\,. \tag{2.10}$$

When applied to the time-independent Schrödinger equation we get

$$E\psi(x) = -\frac{1}{2}\frac{\partial^2}{\partial x^2}\psi(x) + V(x)\psi(x)\,. \tag{2.11}$$

We can then rewrite the equation as

$$\frac{\partial^2}{\partial x^2}\psi(x) = 2(V(x) - E)\psi(x)\,. \tag{2.12}$$

This equation informs on the concavity of the eigenfunctions of Hermitian Hamiltonians, and can be broken down as

$$\begin{aligned}
\psi > 0 \quad &, \quad V(x) > E \quad \Rightarrow \quad \ddot{\psi} > 0 \quad (\text{concave up})\,,\\
\psi > 0 \quad &, \quad V(x) < E \quad \Rightarrow \quad \ddot{\psi} > 0 \quad (\text{concave down})\,,\\
\psi < 0 \quad &, \quad V(x) > E \quad \Rightarrow \quad \ddot{\psi} > 0 \quad (\text{concave up})\,,\\
\psi < 0 \quad &, \quad V(x) < E \quad \Rightarrow \quad \ddot{\psi} > 0 \quad (\text{concave down})\ .
\end{aligned}$$

Thus, when the eigenfunction is in the classically allowed region $(E > V)$ the eigenfunction is concave towards the x axis and when it is in the classically forbidden region $(E < V)$, the eigenfunction is concave away from the x axis. This is know as the concavity condition.

2.2. PSEUDO-HERMITIAN OPERATORS

While Hermitian Hamiltonians provide us with quite a few advantages, one might speculate about physically interesting generalizations of the concept of Hermiticity. In 1943 Pauli introduced pseudo-Hermiticity in which the restrictions imposed on Hamiltonians are relaxed [1]. Pseudo-Hermiticity requires $H = \eta^{-1}\, H^{+}\, \eta$, where η is Hermitian so when $\eta = \mathbb{1}$ the Hamiltonian is Hermitian. In this way, pseudo–Hermiticity expands on the existing framework of Hermiticity. Similarly, in 1998 Bender and Boettcher proposed $\mathcal{PT}$-symmetry [2], where $\mathcal{P}$ is the parity operator ($x \rightarrow -x$) and $\mathcal{T}$ is the time reversal operator ($\mathrm{i} \rightarrow -\mathrm{i}$). By definition a Hamiltonian H is $\mathcal{PT}$–symmetric when $H = \mathcal{PT}\, H\, \mathcal{TP}$. The Hamiltonian $H_3 = \frac{1}{2}p^2 + \frac{1}{2}x^2 + \mathrm{i}Gx^3$ ($G \in \mathbb{R}$), for example, is both $\mathcal{P}$–Hermitian (pseudo–Hermitian with $\eta = \mathcal{P}$) and exactly $\mathcal{PT}$–symmetric. Unlike pseudo-Hermiticity, Hermiticity does not imply $\mathcal{PT}$-symmetry. While there are Hermitian Hamiltonians that are $\mathcal{PT}$-symmetric, there exist Hamiltonians that are self-adjoint, yet do not fulfill the conditions to be $\mathcal{PT}$-symmetric (such as $h_3 = \frac{1}{2}p^2 + \frac{1}{2}x^2 + gx^3$, where $g \in \mathbb{R}$). As such the $\mathcal{PT}$-symmetric class does not expand on an existing class of viable Hamiltonians, but instead constitutes a new class to be examined.

First we examine the properties of pseudo-Hermitian operators. For an operator A to be pseudo-Hermitian it must meet the requirement

$$\eta A = A^{+}\eta\,, \qquad\qquad A = \eta^{-1}A^{+}\eta\,, \tag{2.13}$$

where η is a Hermitian operator itself. In fact, the operator A would then be defined as η–Hermitian. For a pseudo-Hermitian Hamiltonian we need to redefine the inner product. While for a Hermitian Hamiltonian the inner product is $\langle\cdot|\cdot\rangle$, for an η-pseudo-Hermitian Hamiltonian the conserved inner product is defined as

$$\langle\cdot|\cdot\rangle_\eta = \langle\cdot|\eta|\cdot\rangle\,. \tag{2.14}$$

we can show that for such a Hamiltonian the inner product is conserved, i.e.,

$$\frac{d}{dt}\langle\cdot|\cdot\rangle_\eta = 0\,. \tag{2.15}$$

Let H be an η–Hermitian Hamiltonian with eigenvectors $|\psi\rangle$ and $|\phi\rangle$. The time independent Schrödinger is unaffected by the relaxation of the constraints, and equations

(2.7) and (2.8) still hold. Then

$$\frac{d}{dt}\langle\psi|\phi\rangle_\eta = \left(\frac{d}{dt}\langle\psi|\right)\eta|\phi\rangle + \langle\psi|\eta\left(\frac{d}{dt}|\phi\rangle\right) = \langle\psi|\mathrm{i}\,H^{+}\eta|\phi\rangle + \langle\psi|\left(-\mathrm{i}\,\eta H|\psi\rangle\right)$$
$$= \mathrm{i}\,\langle\psi|\left(H^{+}\eta - \eta H\right)|\phi\rangle = 0\,. \tag{2.16}$$

We can now examine the spectrum of H. If we assume that $|\psi\rangle$ is an eigenvector of H, where $H|\psi\rangle = \lambda|\psi\rangle$ and $\lambda \in \mathbb{C}$, we find

$$\langle\psi|\eta H|\psi\rangle = \langle\psi|\eta\left(H|\psi\rangle\right) = \lambda\langle\psi|\psi\rangle_\eta\,, \tag{2.17}$$

and

$$\langle\psi|\eta H|\psi\rangle = \left(\langle\psi|H^{+}\right)\eta|\psi\rangle = \lambda^{*}\langle\psi|\psi\rangle_\eta\,. \tag{2.18}$$

By comparing equations (2.17) and (2.18) we deduce that provided $\langle\psi|\psi\rangle_\eta \neq 0$, i.e., provided we can normalize the state $|\psi\rangle$ in the η norm, then $\lambda = \lambda^{*}$, i.e., the eigenvalues of a pseudo–Hermitian Hamiltonian are real.

Let us now turn our attention to $\mathcal{PT}$–symmetry as defined in [2]. For an operator A to be $\mathcal{PT}$-symmetric, it must fulfill the condition

$$A = \mathcal{PT}\,A\,\mathcal{TP}\,, \tag{2.19}$$

where $\mathcal{T}$ is the time reversal operator ($\mathrm{i} \to -\mathrm{i}$) and $\mathcal{P}$ is the the parity operator ($x \to -x$). Most nonrelativistic Hamiltonians are of the form

$$H = \frac{\vec{p}^{\,2}}{2m} + V\,, \tag{2.20}$$

where the only occurrence of imaginary terms will be in the potential V. When considering such a Hamiltonian we find $\mathcal{T}\,H\,\mathcal{T} = H^{+}$. Then if H is $\mathcal{PT}$–symmetric we find

$$H = \mathcal{PT}\,H\,\mathcal{TP} = \mathcal{P}\,H^{+}\,\mathcal{P}\,, \tag{2.21}$$

which we couple with the knowledge that $\mathcal{P}$ is Hermitian, and $\mathcal{P}^{-1} = \mathcal{P}$, to conclude that H is $\mathcal{P}$–Hermitian. Then provided a $\mathcal{PT}$–symmetric Hamiltonian is of the proper form, it will also be $\mathcal{P}$–Hermitian. However, if we consider the precise definition of the time reversal operator, $\mathcal{T}$ as the operator that takes i to $-$i. Then $\mathcal{T}\,A\,\mathcal{T} = A^{*}$, which coupled with the observation that $\mathcal{P}^{-1} = \mathcal{P}$ leads us to conclude that $\mathcal{PT}$-symmetry means

$$\mathcal{P}\,A = A^{*}\mathcal{P}\,, \tag{2.22}$$

which is strictly $\mathcal{P}$-Hermitian. It is still very much a possibility that A is pseudo–Hermitian is some way. Simply put, $\mathcal{PT}$–symmetry does not imply $\mathcal{P}$–Hermiticity. It is also worth noting that the conserved $\mathcal{PT}$–symmetric scalar product is

$$\langle\psi_n|\psi_m\rangle_{*} = \langle\psi_n|\mathcal{P}|\psi_m\rangle\;, \tag{2.23}$$

which we recognize as the $\mathcal{P}$–Hermitian scalar product.

For completeness sake, let us quickly look at the eigenvalues. Give a $\mathcal{PT}$–symmetric Hamiltonian H, with a wave–function $|\psi\rangle$ and eigenvalue λ such that

$$H\,|\psi\rangle = \lambda\,|\psi\rangle\;, \tag{2.24}$$

one can easily show that $|\phi\rangle = \mathcal{PT}\,|\psi\rangle$ is wave–function of H with eigenvalue λ^*,

$$H\,|\phi\rangle = (\mathcal{PT}\,H\,\mathcal{TP})\,H\mathcal{PT}\,|\psi\rangle = \mathcal{PT}\,H\,|\psi\rangle = \mathcal{PT}\,\lambda\,|\psi\rangle = \lambda^*\,\mathcal{PT}\,|\psi\rangle = \lambda^*\;|\phi\rangle\;. \tag{2.25}$$

Thus the eigenvalues of H come in complex conjugate pairs. If the $\mathcal{PT}$–symmetry is exact, then $|\psi\rangle$ is also an eigenfunction of $\mathcal{PT}$, and the eigenvalues will be real. For example, if $\mathcal{PT}\,|\psi\rangle = c\,|\psi\rangle$ then (2.25) quickly yields

$$H\,|\psi\rangle = \lambda^*\,|\psi\rangle\;, \tag{2.26}$$

and we conclude that $\lambda = \lambda^*$, meaning that λ is real.

Recent years have seen pseudo–Hermiticity and $\mathcal{PT}$–symmetry gain a foothold in quantum mechanics and field theories, including the following four areas. (i) Bender, Jones and collaborators [35–37] have revisited several theoretical quantum field models, that for one reason or another were deemed problematic under the restrictions imposed by Hermiticity, using $\mathcal{PT}$-symmetry; they were able to show that some of the “problems” are remedied under this interpretation. Notably the *ghost* state in the Lee model was shown to have a positive norm when reinterpreted using $\mathcal{PT}$-symmetry [35]. (ii) A standard way to create cosmological models with phantom energy is to use a scalar field with negative kinetic energy. Unfortunately this method is unstable. Andrianov *et al.* [38] have studied cosmological models coupling two fields, one of which has a complex potential while both have a positive kinetic term. This model is described by a $\mathcal{PT}$-symmetric Lagrangian, and is free of the instability that plagues the other model. Furthermore, this model may help explain a number of phenomenological paradoxa in the evolution of the Universe from the big bang to the “big rip.” (iii) Canonically, the index of refraction is used to describe the propagation of light through a medium. When said medium is opaque, light is absorbed and the index of refraction gains an imaginary term, making the index of refraction complex. As such the evolution equations of certain waveguides are $\mathcal{PT}$-symmetric Schrödinger equations. One such example is a double channel waveguide, where one channel has a loss and the other a gain [39]. Rather than being invariant in time, the equation is invariant along the direction of propagation. $\mathcal{PT}$-symmetric photonics honeycomb lattices have also been studied [40–42]. By introducing an alternating gain-loss structure and a specific deformation a $\mathcal{PT}$-symmetric lattice is created. If the deformation is not applied, then the $\mathcal{PT}$-symmetry is broken and wave propagation in such a lattice is related to tachyonic dispersion relations [40]. (iv) The "tachyonic Dirac Hamiltonian" proposed by Chodos, Hauser and Kostelecky [25] has recently been identified as a pseudo-Hermitian Hamiltonian [43]; one should be stressed that current experimental data neither excludes nor confirms neutrino propagation exceeding the speed of light [44, 45].

2.3. IMPORTANCE OF BOUNDARY CONDITIONS

As mentioned at the beginning of chapter 2.1, eigenvalues of Hamilton operators are used to mathematically describe the eigenenergies of quantum mechanical systems. As such, the energies can be discrete or continuous, as well as finite or infinite. The spectrum of a Hamiltonian is the set of all possible eigenvalues of said Hamiltonian. For example, we can consider the Hamiltonian for the quantum harmonic oscillator ($H_0 = \frac{1}{2}\,p^{\,2} + \frac{1}{2}\,x^2$), the solution to which is very well known. The spectrum of H_0 is $\{n + \frac{1}{2}\,|\,n \in \mathbb{N}\}$ (we have set $\hbar = \omega = m = 1$). Similarly, the spectrum of a free particle Hamiltonian, $H = p^{\,2}/(2m)$, is comprised of all positive real numbers. Implicit with this very brief discussion of the spectrum are the boundary conditions imposed on the eigenfunctions of the Hamiltonians. In this case we are requiring that the eigenfunctions do not diverge as x goes to $\pm\infty$, as well as being either orthonormal or Dirac orthonormal. Eigenfunctions with a discrete spectrum will be orthonormal, while eigenfunctions with a continuous spectrum will be Dirac orthonormal. Throughout this section we are going to investigate the importance of the boundary conditions and the resulting spectrums. We begin our investigation by looking at a free particle in one dimension, i.e.,

$$H = \frac{p^2}{2m} = -\frac{1}{2m}\frac{\partial^2}{\partial x^2}\,. \tag{2.27}$$

Let $\psi(x)$ be an eigenfunction of H with eigenvalue λ, then

$$H\psi(x) = -\frac{1}{2m}\frac{\partial^2}{\partial x^2}\psi(x) = \lambda\psi(x)\,, \tag{2.28}$$

which has the solution

$$\psi(x) = \psi_k(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}kx}\,, \tag{2.29}$$

where $k \in \mathbb{R}$. We can now find the eigenvalues of H

$$H\psi(x) = -\frac{1}{2m}\frac{\partial^2}{\partial x^2}\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}kx}\right) = -\frac{1}{2m}(-k^2)\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}kx}\right) = \frac{k^2}{2m}\psi(x)\,. \tag{2.30}$$

As the spectrum is continuous, we require the eigenfunctions to be Dirac orthonormal:

$$\int_{-\infty}^{\infty} dx\,\psi_k^*(x)\,\psi_{k'}(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^2\int_{-\infty}^{\infty} dx\,\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{e}^{\mathrm{i}k'x} = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\,\mathrm{e}^{\mathrm{i}(k'-k)x} = \delta(k'-k)\,. \tag{2.31}$$

Thus we have a basis of eigenfunctions that are Dirac orthonormal, as well as a positive spectrum of eigenvalues for the free particle. This is all dependent on the implicit boundary condition that the inner product is integrated along the real axis ($\int_{-\infty}^{\infty} dx\psi_k^*(x)\psi_{k'}(x) = \delta(k'-k)$). If instead we require that the inner product is integrated along the imaginary axis ($\int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} dx \psi_k^*(x)\psi_{k'}(x)$, notice that the limits of integration are now *imaginary*), the eigenfunctions are then

$$\phi(x) = \phi_k(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{kx}\,. \tag{2.32}$$

Solving for the eigenvalues we find

$$H\phi(x) = -\frac{1}{2m}\frac{\partial^2}{\partial x^2}\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{kx}\right) = -\frac{1}{2m}(k^2)\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{kx}\right) = -\frac{k^2}{2m}\phi(x)\,. \tag{2.33}$$

Leaving us with eigenvalues $-\frac{k^2}{2m}$, where $k \in \mathbb{R}$, implying that the eigenvalues are negative. If we then check the inner product we find

$$\int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} dx\, \phi_k^*(x)\, \phi_{k'}(x) = \mathrm{i}\left(\frac{1}{\sqrt{2\pi}}\right)^2 \int_{\infty}^{-\infty} du\, \mathrm{e}^{-\mathrm{i}ku}\, \mathrm{e}^{\mathrm{i}ki'u} = -\mathrm{i}\delta(k'-k)\,. \tag{2.34}$$

Thus by changing the boundary conditions on the free particle we have real eigenfunctions that are orthogonal, but no longer Dirac normalizable along the real axis (the inner product gives an imaginary result), whose energies are negative.

Notice that $\phi_k(\mathrm{i}x) = \psi_k(x)$ for the free particle. Trivially, the same can be shown for the harmonic oscillator ($H_0(x) = \frac{1}{2}p^2 + \frac{1}{2}x^2$). i.e.,

$$\phi_n(\mathrm{i}x) = \psi_n(x)\,. \tag{2.35}$$

It is well known that $H_0(x)\psi_n(x) = (n+\frac{1}{2})\psi_n(x)$ (i.e., $\epsilon_n^{(0)} = n+\frac{1}{2}$), while it is less well known what $\widetilde{\epsilon}_n^{(0)}$ is (where $H_0(x)\phi_n(x) = \widetilde{\epsilon}_n^{(0)}\phi_n(x)$). We begin by noting that if we let $x = \mathrm{i}\chi$ then the boundary conditions on $\phi_n(x)$ are that ϕ_n must be square integrable as $\chi \to \pm\infty$. Thus

$$\begin{aligned} H_0(x)\phi_n(x) &= H_0(\mathrm{i}\chi)\phi_n(\mathrm{i}\chi) = \left(-\frac{1}{2}\frac{\partial^2}{\partial(\mathrm{i}\chi)^2} + \frac{1}{2}(\mathrm{i}\chi)^2\right)\psi_n(\chi) \\ &= \left(\frac{1}{2}\frac{\partial^2}{\partial\chi^2} - \frac{1}{2}\chi^2\right)\psi_n(\chi) = -\left(-\frac{1}{2}\frac{\partial^2}{\partial\chi^2} + \frac{1}{2}\chi^2\right)\psi_n(\chi) \\ &= -H_0(\chi)\psi_n(\chi) = -\epsilon_n^{(0)}\phi_n(\mathrm{i}\chi) = -\epsilon_n^{(0)}\phi_n(x)\,. \end{aligned} \tag{2.36}$$

Thus $\widetilde{\epsilon}_n^{(0)} = -\epsilon_n^{(0)} = -n - \frac{1}{2}$. So, as with the free particle, the energies of H_0 can be positive or negative depending on the boundary conditions. For the sake of completion, we should make sure that the eigenfunctions associated with the imaginary boundary conditions are normalized as well. We already know that

$$\int_{-\infty}^{\infty} dx\,\psi_n^*(x)\psi_{n'}(x) = \delta_{n,n'}\,. \tag{2.37}$$

By again letting $x = \mathrm{i}\chi$ we find

$$\int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} dx\,\phi_n^*(x)\phi_{n'}(x) = \int_{\infty}^{-\infty} d(\mathrm{i}\chi)\,\phi_n^*(\mathrm{i}\chi)\phi_{n'}(\mathrm{i}\chi) = -\mathrm{i}\int_{-\infty}^{\infty} d\chi\,\psi_n^*(\chi)\psi_{n'}(\chi) = -\mathrm{i}\delta_{n,n'}\,. \tag{2.38}$$

We might have expected this result based on the above mentioned considerations for the free particle.

Up to this point, we examined how to change the boundary conditions from purely real to purely imaginary ($x \to \mathrm{i}\chi$, where χ is real). What if instead we rotate the boundary conditions into the complex plane by some angle θ (a procedure known as "complex scaling")? Let us consider the Hamiltonian,

$$h_3 = \frac{1}{2}p^{\,2} + \frac{1}{2}x^2 + g\,x^3\,, \quad g \in \mathbb{R}_+\,, \tag{2.39}$$

which is clearly Hermitian, and as such will have a real energy spectrum. The potential of h_3 ($V(X) = \frac{1}{2}x^2 + g\,x^3$) presents a problem: While $V(x) \to \infty$ as $x \to \infty$, $V(x) \to -\infty$ as $x \to -\infty$, telling us that the eigenvalues of h_3 with real eigenenergies that do not diverge $x \to \infty$, however when $x \to -\infty$ the wave–functions diverge. However, there is an established method to find the complex resonance energies of the Hamiltonian h_3, which allows for tunneling effects (the particle reaches $x = \infty$ in a finite time, classically. To find the complex resonance eigenvalues of h_3 (and in general), we complex scale the Hamiltonian, i.e.,

$$\begin{aligned} x &\to x\,\mathrm{e}^{i\,\theta}\,, \quad \text{where} \quad 0 < \theta < \frac{\pi}{5}\,, \\ h_3 &\to \frac{1}{2}\mathrm{e}^{-2\mathrm{i}\theta}p^{\,2} + \frac{1}{2}\mathrm{e}^{2\mathrm{i}\theta}x^2 + g\,\mathrm{e}^{3\mathrm{i}\theta}x^3\,. \end{aligned} \tag{2.40}$$

We then construct and diagonalize the corresponding matrix, using the harmonic oscillator wave–functions as a basis. This results in complex energies of the form

$$E_n = \mathrm{Re}(E_n) - \mathrm{i}\frac{\Gamma_n(g)}{2}\,, \tag{2.41}$$

where $\Gamma_n(g)$ is the decay width of the eigenstates. More details on anharmonic oscillators, including generalized quantization conditions, have been described in [5, 46, 47].

2.4. HERMITIZING TRANSFORMS

In some cases, where one is only interested in the energy levels, a $\mathcal{PT}$–symmetric Hamiltonian may be considered a compact version of a more complicated Hermitian Hamiltonian, and to this end, there is a transform which relates the two Hamiltonians [34, 49]. Let h be a Hermitian Hamiltonian and H be a $\mathcal{PT}$-symmetric

Hamiltonian, then the conjecture suggests that there is some positive-definite, Hermitian operator ρ such that

$$H = \rho^{-1}\, h\, \rho\,. \tag{2.42}$$

Notice that ρ takes a Hermitian Hamiltonian to a non–Hermitian Hamiltonian. Thus, ρ is necessarily not a unitary transform (which would otherwise conserve Hermiticity). As such, it should come as no surprise that when one considers the wave–functions, as we do in chapter 3, the $\mathcal{PT}$–symmetric Hamiltonians cannot be similar to any Hermitian Hamiltonians. This assertion is confirmed by [21], in which it is shown that the wave–functions of the $\mathcal{PT}$–symmetric Hamiltonian $H = p^{\,2} + \mathrm{i}\, x^3$ do not from a Riesz basis. Still, Hermitizing transforms are still an interesting concept, and the back–transformation is

$$h = \rho\, H\, \rho^{-1}\,. \tag{2.43}$$

Since ρ is positive definite and Hermitian, the Hermitian adjoint of H is given by

$$H^{+} = \rho^{+} h^{+} (\rho^{-1})^{+} = \rho\, h\, \rho^{-1}\,. \tag{2.44}$$

Working under the assumption that $\mathcal{T} H \mathcal{T} = H^{+}$, then

$$H = \rho^{-1} h\, \rho = \rho^{-1}(\rho^{-1}\rho)\, h\, (\rho^{-1}\rho)\rho = \rho^{-2}(\rho\, h\, \rho^{-1})\rho^{2} = \rho^{-2} H^{+} \rho^{2} = \rho^{-2} \mathcal{T} H \mathcal{T} \rho^{2}\,. \tag{2.45}$$

We know, from the assumption that H is $\mathcal{PT}$-symmetric that $\mathcal{PT} H\, \mathcal{TP} = H$, which leads us to the conclusion that a proper choice for the Hermitizing operator ρ might be $\rho^2 = \mathcal{P}$. However $\mathcal{P}$ is not positive definite, and neither is $\sqrt{\mathcal{P}}$, meaning we have the wrong operator for ρ, and we need to try a different approach.

In this approach we will attempt to relate the inner products associated with h and H. Firstly, let the eigenvectors of h and H be $|\phi\rangle$ and $|\psi\rangle$ respectively. Then, we can show that they have the same eigenenergies and $|\phi\rangle = \rho|\psi\rangle$. Let $H|\psi\rangle = E|\psi\rangle$, then

$$h|\phi\rangle = \left(\rho H \rho^{-1}\right)(\rho|\psi\rangle) = \rho H \rho^{-1}\rho|\psi\rangle = \rho H|\psi\rangle = E\rho|\psi\rangle = E|\phi\rangle \,. \tag{2.46}$$

We know that the inner product of the eigenvectors of h are positive-definite, however the inner product of the eigenvectors of H are not necessarily positive-definite. Then $\langle\phi|\phi'\rangle \geq 0$, while the same can not be said for $\langle\psi|\psi'\rangle$. Furthermore, we need to use the pseudo–Hermitian inner product as discussed in chapter 2.2. We can modify the inner product in the following way:

$$\langle\psi|\mathcal{P}|\psi'\rangle \rightarrow \langle\psi|\mathcal{CP}|\psi'\rangle \,, \tag{2.47}$$

where the operator $\mathcal{C}$ insures the "new" inner product is positive–definite (note that this $\mathcal{C}$ is not the charge conjugation operator). Now,

$$\langle\phi|\phi'\rangle = \langle\psi|(\rho^{-1})^{+}\rho^{-1}|\psi\rangle = \langle\psi|\rho^{-1}\rho^{-1}|\psi\rangle = \langle\psi|\rho^{-2}|\psi\rangle \,. \tag{2.48}$$

We then let

$$\rho = \mathrm{e}^{-\frac{1}{2}Q} \,, \tag{2.49}$$

from which we deduce that

$$\mathcal{CP} = \mathrm{e}^{Q} \,, \tag{2.50}$$

and thus

$$\mathcal{C} = \mathrm{e}^{Q}\mathcal{P} \quad , \mathcal{C}^{-1} = \mathcal{P}\mathrm{e}^{-Q} \,. \tag{2.51}$$

Determining exactly what $\mathcal{C}$ is can be quite involved, as the $\mathcal{C}$ operator is not well known [31,50]. Since $\mathcal{C}$ is unitary and commutes with both H and $\mathcal{PT}$ we can rather trivially extend $\mathcal{PT}$-symmetry to $\mathcal{CPT}$-symmetry in which

$$\mathcal{C}\,\mathcal{PT}\,H\,\mathcal{TP}\,\mathcal{C}^{-1} = H \,, \tag{2.52}$$

and the inner product is

$$\langle\psi|\psi'\rangle_{\mathcal{CP}} = \langle\psi|\mathcal{CP}|\psi'\rangle \,. \tag{2.53}$$

This adds an advantage in that the inner product is now positive-definite.

In principle any $\mathcal{PT}$-symmetric Hamiltonian can be mapped onto a Hermitian Hamiltonian. But why bother? The $\mathcal{PT}$-symmetric Hamiltonians are far less complicated and therefore more practical. Beyond that, the rotations ρ don't necessarily conserve parity (i.e., $[\rho, \mathcal{P}] \neq 0$) and as we will discuss in chapter 10.8, unless parity is conserved, the results can become meaningless. In short, while the Hermitizing transforms constitute an interesting theoretical connection between pseudo–Hermiticity and pure Hermiticity, in practice they are an unnecessary exercise that complicates an otherwise reasonable representation of a given system.

3. PSEUDO–HERMITICITY AS AN INDEPENDENT CONCEPT

3.1. OVERVIEW AND ORIENTATION

In the previous chapters, we have discussed pseudo–Hermitian, and $\mathcal{PT}$–symmetric, Hamiltonians, and argued that they constitute independent classes of time derivative operators in quantum mechanics, rather than alternative versions of Hermitian Hamiltonians, connected to the "original" Hermitian operator via Hermitizing transforms. In this chapter we work to further this cause by studying the structure of the wave–functions of a $\mathcal{PT}$–symmetric Hamiltonian (which also happens to be $\mathcal{P}$–Hermitian).

Despite the prolific literature on $\mathcal{PT}$-symmetric quantum mechanics [1–14], where the spectrum of $\mathcal{PT}$–symmetric Hamiltonians has been analyzed in detail, the properties of the wave–functions corresponding to the eigenstates of the Hamiltonians are generally overlooked (with the exception of [21], which shows that the wave–functions are complete for $H = p^2 + \mathrm{i}\,x^3$). This is all the more surprising because a number of interesting field-theoretical model theories and a streamlined description of phenomenologically important so-called $\mathcal{PT}$-symmetric wave guides rely on $\mathcal{PT}$-symmetric quantum mechanics and field theory (as discussed in chapter 2.2). Moreover, the concept of a $\mathcal{PT}$-symmetric Hamiltonian has recently been instrumental in finding a generalization of the so-called Bender–Wu formulas [51–53] to odd anharmonic oscillators [5, 12]. An intuitive understanding of the physics involved in $\mathcal{PT}$-symmetric models is hard to obtain without looking at the wave–functions.

In this context, one may well ask the following question: The Hamiltonian H_3

$$H_3 = \frac{1}{2}p^2 + \frac{1}{2}x^2 + \mathrm{i}\,G\,x^3\,, \tag{3.1}$$

involves a manifestly complex potential,

$$V(x) = \frac{1}{2}\,x^2 + \mathrm{i}\,G\,x^3 = |V(x)|\,\mathrm{e}^{\mathrm{i}\,\arg(V(x))}\,, \tag{3.2}$$

whose complex modulus tends to infinity as $x \to \pm\infty$. For ordinary (purely real) potentials, intuition suggests that the "bulk" of the probability density of the eigenstate wave–function should be concentrated in the "classically allowed" region, i.e., in the region where the eigenenergy E is greater than the potential, $E > V(x)$ (with $V(x) \in \mathbb{R}$). Here we endeavor to generalize this concept to $\mathcal{PT}$-symmetric quantum mechanics. For a manifestly complex potential, the condition $E > V(x)$ (with $V(x) \in \mathbb{C}$) does not make any sense because the complex numbers are not ordered. It is thus unclear how the concept of a "classically allowed region" should be generalized to the complex domain. We also observe that the eigenstate wave–functions of a $\mathcal{PT}$-symmetric Hamiltonian do not need to be eigenstates of parity, because the parity operator does not necessarily commute with the $\mathcal{PT}$-symmetric Hamiltonian. These observations raise a number of obvious pertinent questions which we attempt to answer.

Based on the work presented in part III of [54], we proceed by recalling a few basic facts about eigenvalue perturbation theory in chapter 3.2, analyzing the parity of eigenstates in chapter 3.3, and continue with a visualization of the manifestly complex $\mathcal{PT}$-symmetric eigenstates in chapter 3.4. Finally, some conclusions are drawn in chapter 3.5.

3.2. ASYMPTOTICS OF IMAGINARY CUBIC PERTURBATION

In order to fix ideas, we would first like to recall a few basic facts about eigenvalue perturbation theory and the imaginary cubic perturbation. The Hamiltonian given in equation (3.1) can easily be split into an unperturbed part $H^{(0)} = -\frac{1}{2}\,\partial_x^2 + \frac{1}{2}\,x^2$

and a perturbed part $H^{(1)} = \mathrm{i}\, G\, x^3$. Starting from the unperturbed harmonic oscillator eigenvalues $H^{(0)}|n\rangle = \epsilon_n^{(0)}|n\rangle$, where $\epsilon_n^{(0)} = n + \frac{1}{2}$, one can develop perturbation theory, either using the classical Rayleigh–Schrödinger approach [55] or using a complex contour integration of the logarithm of the wave–function [46, 47], which transforms the Schrödinger equation into a Riccati differential equation. The first non-vanishing term is of second order,

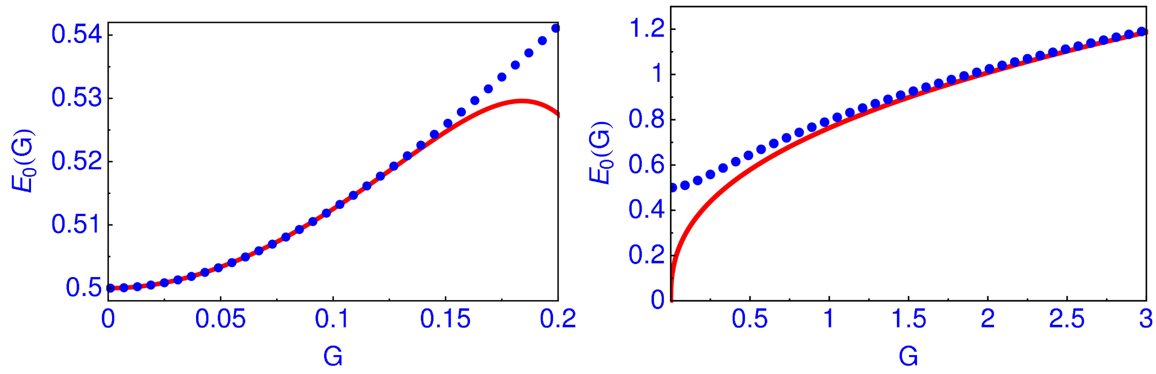
$$\Delta\epsilon_n^{(2)} = -\sum_{m \neq n} \frac{\langle m|H^{(1)}|n\rangle\langle n|H^{(1)}|m\rangle}{\epsilon_m^{(0)} - \epsilon_n^{(0)}} = G^2 \sum_{m \neq n} \frac{|\langle m|x^3|n\rangle|^2}{m - n}$$
$$= \frac{G^2}{8}\left(30n^2 + 30n + 11\right) + \mathcal{O}(G^4)\,. \tag{3.3}$$

Unlike second-order perturbations involving a Hermitian operator, the second-order term here is positive and shifts the ground-state energy level upward. Through fourth order, the result reads as

$$\epsilon_n = n + \frac{1}{2} + \frac{G^2}{8}\left(30n^2 + 30n + 11\right) \tag{3.4}$$
$$- \frac{15}{32}\, G^4\, \left(94n^3 + 141n^2 + 109n + 21\right) + \mathcal{O}(G^6))\,.$$

For the ground state ($n = 0$), this expression evaluates to $1/2 + 11\, G^2/8 - 465\, G^4/32$, while going to eighth order we obtain $1/2 + 11\, G^2/8 - 465\, G^4 + 39709\, G^6/128 - 19250805\, G^8/2048$ which is plotted against numerical values of the ground-state energy, as a function of G, in figure 3.1(a). The positive curvature of the ground-state energy in the weak-coupling regime is clearly visible.

We can also look at the strong-coupling asymptotic as $G \to \infty$, for which we employ a "poor man's scaling" in which we rotate $x \to G^{-\frac{2}{5}}x$. Under this scaling, H_3



(a) (b)

Figure 3.1: The ground-state energy of the anharmonic oscillator Hamiltonian with imaginary cubic perturbation, as given in equation (3.1), is plotted in the perturbative (weak-coupling) regime in figure (a). The solid curve represents the results from eighth-order perturbation theory. In figure (b), the same eigenvalue is plotted in the strong-coupling regime $G > 0$. The small dots represent numerical values.

is replaced by the following Hamiltonian H_3', which has the same spectrum as H_3,

$$\begin{aligned} H_3' &= G^{\frac{2}{5}}\left(-\frac{1}{2}\partial_x^2 + \mathrm{i}x^3 + \frac{1}{2}G^{-\frac{2}{5}}x^2\right) \\ &\to G^{\frac{2}{5}}\left(-\frac{1}{2}\partial_x^2 + \mathrm{i}x^3\right), \qquad G \to \infty \end{aligned} \tag{3.5}$$

The $\mathcal{PT}$-symmetric Hamiltonian $H_3' = G^{2/5}(-\frac{1}{2}\partial_x^2 + \mathrm{i}x^3)$ has its own set of eigenvalues, which we label as ϵ_n'. Thus, the large-coupling asymptotics for the energies ϵ_n' of H_3' reads as

$$\epsilon_n' \sim G^{\frac{2}{5}}\,\epsilon_n^{(0)}\,, \tag{3.6}$$

for large G, where the $\epsilon_n^{(0)}$ are the energies of the Hamiltonian $H_3^{(0)}$, with

$$H_3^{(0)} = -\frac{1}{2}\partial_x^2 + \mathrm{i}x^3 \to 2^{-\frac{3}{5}}\left(-\partial_x^2 + \mathrm{i}x^3\right)\,. \tag{3.7}$$

In the last step, we have done the scaling transformation $x \to 2^{-\frac{1}{5}}\,x$. This transformation allows us to connect the strong-coupling asymptotics with the literature, notably, with references [2] and [56, 57]. In particular, we have $\epsilon_n^{(0)} = 2^{-\frac{3}{5}}\,\widetilde{\epsilon}_n$ where the $\widetilde{\epsilon}_n$ are the energies of the Hamiltonian $p^2 + \mathrm{i}x^3$.

3.3. PARITY OF EIGENSTATES

We have determined, to high numerical accuracy, the manifestly complex wave–functions of the ground state and the first two excited eigenstates of the imaginary cubic oscillator. The results are displayed in figure 3.2. The parity operator $\mathcal{P}$ does not commute with the Hamiltonian H_3, and the eigenstates of H_3 are not eigenstates of parity. Furthermore, because the potential is manifestly complex, we cannot choose the wave–functions to be purely real. However, numerical evidence drawn from figure 3.2 suggests that individually, both real as well as imaginary part of the wave–function are eigenstates of parity. Indeed, we can formally split the Hamiltonian H_3 into a "real part" and a "imaginary part" as follows,

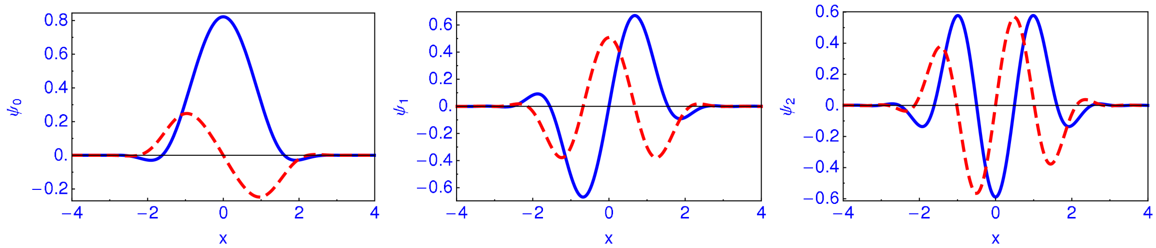
$$\mathrm{Re}\, H_3 = -\tfrac{1}{2}\,\partial_x^2 + \tfrac{1}{2}x^2\,, \quad \mathrm{Im}\, H_3 = \mathrm{i}\, G\, x^3\,. \tag{3.8}$$

If we also split the eigenstate wave–function $\psi_n(x)$ into real and imaginary parts,

$$\psi_n(x) = \mathrm{Re}\,\psi_n(x) + \mathrm{i}\,\mathrm{Im}\,\psi_n(x)\,, \tag{3.9}$$

then it is rather easy to show that if $\mathrm{Re}\,\psi_n(x)$ is even under parity and $\psi_n(x)$ is an eigenstate of H_3 with real eigenvalue of ϵ_n, then $\mathrm{Im}\,\psi_n(x)$ has to be parity-odd, and vice versa. This is accomplished by first constructing the eigenstates as a linear combination of the eigenstates of the harmonic oscillator, $|m\rangle$, i.e.,

$$|\psi_n\rangle = \sum_m a_m\ |m\rangle\ . \tag{3.10}$$



(a) (b) (c)

Figure 3.2: Figure (a) displays the complex ground-state wave–function of the imaginary cubic Hamiltonian (3.1) for $G = 1.0$. The real part is plotted using solid lines, and the dashed lines plot the imaginary part. The real part is even, the imaginary part is odd under parity. For the first excited state (still, $G = 1.0$), the real part is odd, while the imaginary part is even under parity (see figure (b)). The second excited state (figure (c)) has an even real part, while its imaginary part is odd.

One then finds that

$$a_m = \mathrm{i}\, C_n \left(\alpha_m\, a_{m+3} + \beta_m\, a_{m+1} + \beta_{m-1}\, a_{n-1} + \alpha_{m-3}\, a_{m-3}\right) \,, \tag{3.11}$$

where

$$C_m = \frac{G}{2\sqrt{2}\left(\epsilon_n - m - \frac{1}{2}\right)} \,, \qquad \alpha_m = \sqrt{(m+1)\,(m+2)\,(m+3)}\,, \tag{3.12a}$$

$$\beta_m = (m+1)^{3/2} \,. \tag{3.12b}$$

Since the spectrum of H_3 is real $C_m, \alpha_m, \beta_m \in \mathbb{R}$. Considering our results for the prefactors, along with the fact that $|m\rangle$ is even under parity when m is even, and odd under parity when m is odd, we find that by assuming $\mathrm{Re}\,\psi_n(x)$ is even under parity, it follows that $a_{2l} \in \mathbb{R}$ and $a_{2l+1} \in \mathbb{I}$ for all $l \in \mathbb{N}$, i.e., if $\mathrm{Re}\,\psi_n(x)$ is even under parity, then $\mathrm{Im}\,\psi_n(x)$ is odd under parity. Similarly, if $\mathrm{Re}\,\psi_n(x)$ is odd under parity and $\psi_n(x)$ has a real eigenvalue of ϵ_n, then $\mathrm{Im}\,\psi_n(x)$ has to be even under parity. In references [58], it has been observed that since $\mathcal{PT}$ commutes with the Hamiltonian, the eigenfunctions of H_3 also have to be eigenfunctions of the $\mathcal{PT}$ operator. Numerical evidence suggests that the appropriate eigenvalues are

$$\mathcal{PT}\,\psi_n(x) = \psi_n^*(x) = (-1)^n\,\psi(x) \tag{3.13}$$

(see equation (5) of reference [58]). In the space of eigenfunctions ψ_n, the conserved scalar product (2.23) then becomes

$$\begin{aligned}\langle\psi_n|\psi_m\rangle_* &= \int \mathrm{d}x\,\psi_n^*(x)\mathcal{P}\psi_m(x) = \int \mathrm{d}x\;(\mathcal{PT}\psi_n)\,(x)\,\psi_m(x)\\ &= (-1)^n \int \mathrm{d}x\,\psi_n(x)\,\psi_m(x)\,,\end{aligned} \tag{3.14}$$

where the last line is *without* complex conjugation. This latter formula shows that the conserved scalar product for the $\mathcal{PT}$-symmetric imaginary cubic perturbation is equal (up to a prefactor) to the generalized inner product for complex-scaled Hamiltonians as defined in equation (2.4.2) of reference [48], which avoids the complex conjugation of the first argument. Identifying the $\mathcal{PT}$-symmetric scalar product as a generalization of the generalized inner product for complex-scaled Hamiltonians (which otherwise give rise to resonances), we stress once more the connection of the imaginary cubic perturbation $\mathrm{i}\,G\,x^3$ to the real cubic perturbation $g\,x^3$, which gives rise to manifestly complex resonance energies [5, 12].

3.4. VISUALIZATION OF $\mathcal{PT}$–SYMMETRIC EIGENSTATES

From the investigations [5, 9, 12], we know that a dispersion relation connects the energy levels of the imaginary cubic perturbation $\mathrm{i}\,G\,x^3$ to the "real" cubic perturbation $g\,x^3$. Furthermore, the eigenenergies of the imaginary cubic potential are real, whereas the resonance and antiresonance energies of the real cubic perturbation are complex. There is no direct and obvious visualization available for the imaginary cubic perturbation. One may well ask in which sense the imaginary cubic perturbation "confines" the eigenstate wave–functions to a "classically allowed" region of space, it is difficult if not impossible to characterize this "classically allowed" region because the set of complex numbers $\mathbb{C}$ is not ordered.

An intuitive understanding can be obtained if we interpret the potential in terms of a complex modulus and a phase, which according to equation (3.2) reads as

$$V(x) = \frac{1}{2}\,x^2 + \mathrm{i}\,G\,x^3 = |V(x)|\,\mathrm{e}^{\mathrm{i}\,\arg(V(x))}\,, \tag{3.15a}$$

$$-\pi \leq \arg(V(x)) < \pi\,. \tag{3.15b}$$

If we then plot the modulus of the complex potential and its complex phase, a "confining" shape is obtained, which is modulated by a complex phase in the range $[-\pi, \pi]$. This is represented in figure 3.3.

The eigenstate wave–function are plotted in figures 3.4 and 3.5, with the idea of writing the complex eigenstate wave–functions as

$$\psi_n(x) = |\psi_n(x)|\,\mathrm{e}^{\mathrm{i}\,\arg(\psi_n(x))}\,. \tag{3.16}$$

It is known that for Hermitian operators, the ground-state wave–function always has maximum symmetry. Furthermore, for both the harmonic oscillator as well as for the "stable" quartic perturbation, the number of zeros of the wave–function is equal to the principal quantum number. This is illustrated in figures 3.4(a) and (b). It has been observed previously [56] that the resonance state wave–function has no zeros when considered as a complex variable (see figure 4 of reference [56]). In figure 3.4(c), the plot of the wave–function square $|\psi_n(x)|^2$ suggests that the same statement holds for the eigenstate wave–functions of the imaginary cubic perturbation: they carry no

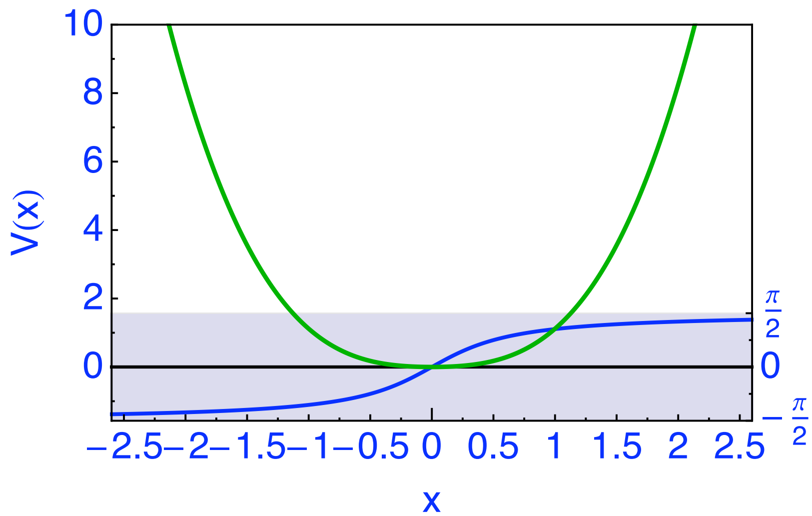
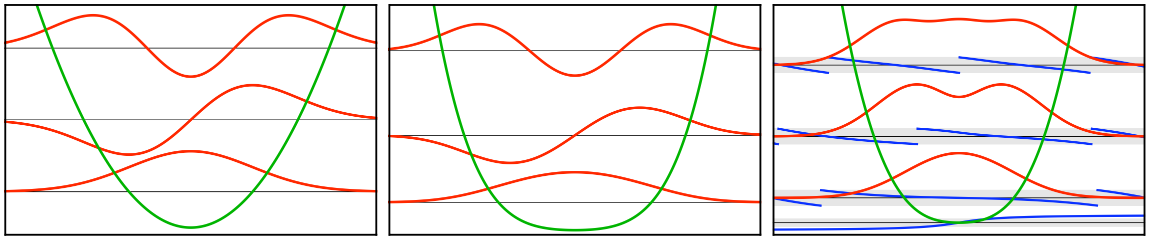


Figure 3.3: Visualization of the harmonic+cubic potential given in equation (3.2). The modulus of the potential $\left(|V(x)| = \sqrt{x^4/4 + G^2\,x^6}\right)$ leads to a confining mechanism for $x \to \pm\infty$. The value of G in the plot is $G = 1.0$. The shaded area displays the complex phase of the potential and covers the interval $(-\pi/2, \pi/2)$.

complex zeros. We have used a numerical value of $G = 1$ in the plot; the abscissa as plotted covers a range of $0 < V(x) < V_0 = 0 < V(x) < 7$ for the potential. The wave–function squares are plotted in arbitrary units; in practical calculations, one normalizes to $\int \mathrm{d}x |\psi_n(x)|^2 = 1$. The complex phase of the wave–function, which covers the shaded areas in figure 3.4(c), revolves in the complex plane and transits a number of Riemann sheets, i.e., it jumps from $-\pi$ to $+\pi$ several times in our interval $-3 < x < 3$. Indeed, we observe that the wave–functions of the imaginary cubic potential have no complex zeros, while the real and imaginary parts, individually, have a number of zeros. This is evident from figure 3.2. The question then is, how many. In answer to this question, we refer to the complex phase of the wave–functions as plotted in the shaded areas of figure 3.4(c).



(a) (b) (c)

Figure 3.4: In figure (a), we plot the wave–function of the ground and the first two excited states of the harmonic oscillator. A "stable" quartic perturbation perturbs the potential in figure (b) and leads to a "confinement" of the wave–function to the classically allowed region $E > V(x)$. The eigenenergies of the imaginary cubic perturbation are real and allow us to plot the complex $\mathcal{PT}$-symmetric eigenfunctions as in figure (c). The complex phase of the wave–function is displayed in the shaded region, as in figure 3.3 for the potential. The red curves denote the wave–functions (figures (a) and (b)), while the blue curves display the complex phase of the wave–functions. It is perhaps useful to note that we use $g = 1$ and $G = 1$ for the plot, while stressing that the main purpose of the plot is to illustrate the qualitative behavior of the wave–functions of the "stable" perturbations (positive quartic and imaginary cubic), and the concrete value of G is irrelevant for this illustration and would clutter the figure. See also the following figure 3.5.

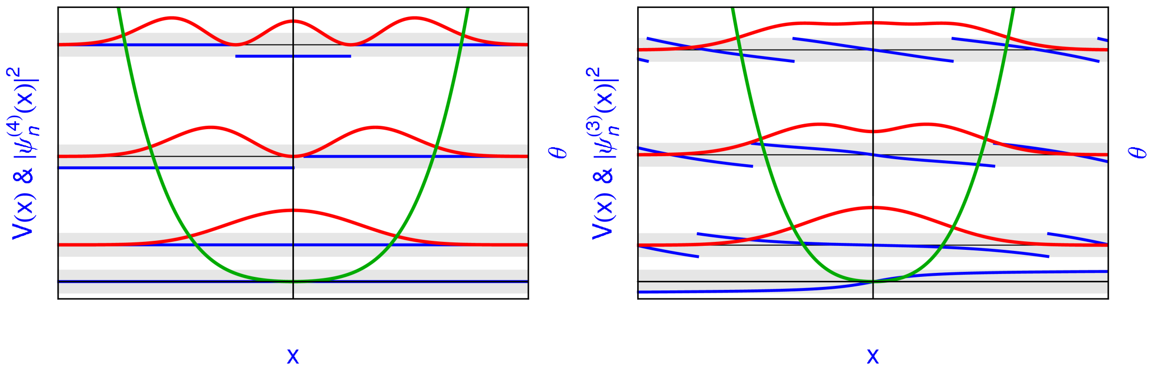
Using a WKB analysis, the wave–functions can be approximated as

$$\psi_n(x) = \frac{1}{(2m\,[V(x) - E_n])^{1/4}} \left[\gamma\,\exp\left(+\int^x dx\,\sqrt{2m\,[V(x) - E_n]}\right) + \delta\,\exp\left(-\int^x dx\,\sqrt{2m\,[V(x) - E_n]}\right)\right], \tag{3.17}$$

where γ and δ are arbitrary constants [59]. Here we are only interested in the phase as $|x| \to \infty$, in which case

$$\int^x dx\,\sqrt{2m[V(x) - E_n]} \to \sqrt{2\mathrm{i}\,G\,m}\int^x dx\,x^{3/2} = C\,(1 + \mathrm{i})x^{5/2}\,, \tag{3.18}$$

where $C = \sqrt{G\,m}$ is a positive, real constant. By considering the imaginary part, we have the asymptotic behavior of our phase, save for a prefactor of ± 1, which we determine based on our numerical analysis. Both our numerical results (see figure 3.6),



(a) $g = 1.0$ (b) $G = 1.0$

Figure 3.5: In figure (a) we plot the probability density $\rho = |\psi(x)|^2$ of the quartic oscillator's ground state and first two excited states. Along with which we plotted the complex phase of the wave–functions which is chosen to be $-\pi$ when the wave–function is negative and zero when positive. In figure (b) we investigate the cubic oscillator. The complex phase of the wave–functions is normalized to zero at the origin by an appropriate scaling factor (multiple of the imaginary unit). Notice that the qualitative features ("humps") of the quartic oscillator are still present in the complex ("$\mathcal{PT}$-symmetric") domain, however the zeros of the wave–functions are "washed out" and become local minima.

as well as a WKB analysis show that the phase of the wave–function has to behave, asymptotically, as $\arg(\psi_n(x)) \sim -x^{5/2}$ for $x \to +\infty$ and as $\arg(\psi_n(x)) \sim x^{5/2}$ for $x \to -\infty$. Every time the complex phase $\arg(\psi_n(x))$ (modulo π) attains zero, the imaginary part of the wave–function vanishes, and whenever $\arg(\psi_n(x)) \bmod \pi = \pi/2$, the real part of the wave–function vanishes. So, we conclude that, even for an infinitesimally small coupling G in the imaginary cubic perturbation $\mathrm{i}\, G\, x^3$, the real and imaginary parts of the wave–function, individually, have an infinite number of zeros. This is somewhat surprising. E.g., the above considerations imply, among other things, that the number of complex zeros of the first-excited-state wave–function of the imaginary cubic perturbation is a discontinuous function of G. Namely, for $G = 0$, we have one complex zero (because the Hamiltonian is equal to the harmonic oscillator), while for non-vanishing G, the total number of complex zeros of

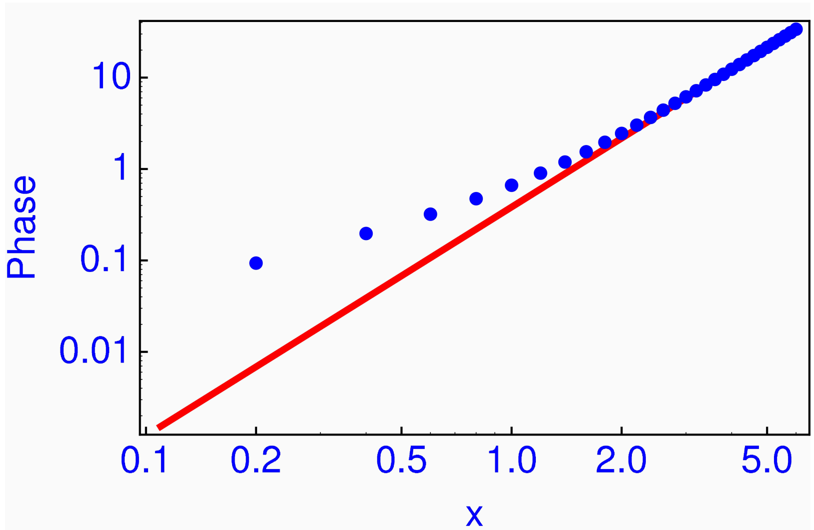


Figure 3.6: Here we plot the negative of the accumulated phase of the ground state wave–function from $x = 0$ to $x = 6$. The solid line is the large coupling $\theta(x) = -3.837\, x^{5/2}$, while the points denote the numerical value of the complex phase.

the wave–function vanishes, while both real and imaginary parts of the wave–function (for non-vanishing G) have an infinite number of zeros.

3.5. CONCLUDING REMARKS

In studying the wave–functions of a $\mathcal{PT}$–symmetric Hamiltonian, we have developed a number of arguments supporting the conclusion that $\mathcal{PT}$–symmetry should be viewed as an independent concept: (i) the physical interpretation of a $\mathcal{PT}$–symmetric Hamiltonian is a system in which the gain and loss terms are "in equilibrium," where a Hermitian Hamiltonian describes a closed system. (ii) While there is an overlap between Hermitian and $\mathcal{PT}$–symmetric Hamiltonians, neither forms a subset of the other. (iii) The wave–functions of the Hamiltonian we examined, which is both $\mathcal{PT}$–symmetric and pseudo–Hermitian, cannot be chosen as real (i.e., they must be complex). Furthermore, the real and imaginary parts of the wave–functions have properties which cannot be reconciled with intuitive concepts we have when considering Hermitian Hamiltonians. (iv) The counter argument, namely Hermitizing transforms, require a non–unitary similarity transform, which does transform a $\mathcal{PT}$–symmetric Hamiltonian into a Hermitian Hamiltonian which shares its eigenvalues with the original Hamiltonian [34, 49], however neither parity, nor the metric are conserved [20, 21]. With these considerations in mind, it becomes clearer that $\mathcal{PT}$–symmetry is not an extension of Hermiticity, but an independent concept in its own right.

Our considerations, reported in figures 3.4, 3.5 and 3.7, suggest that the eigenstate wave–functions of $\mathcal{PT}$-symmetric Hamiltonians are not as "exotic" as one might otherwise imagine. An interesting observation, described in chapter 3.4, is as follows: In a $\mathcal{PT}$-symmetric case, the number of zeros of the wave–function cannot be used to enumerate the eigenstates. The wave–functions have no complex zeros; yet, quite contrarily, both the real as well as the imaginary part have an infinite number of zeros, individually. The modulus of the potential, which tends to infinity as $|x| \to \infty$, is responsible for the confinement of the wave–function of the imaginary cubic Hamiltonian to a "classically allowed region" in much the same way as one would expect from Hermitian quantum dynamics. However, the numerical evidence suggests that one may be able to enumerate the $\mathcal{PT}$–symmetric wave–functions by considering the local minima of the modulus of the wave–function in this region. As remarked near the end of chapter 3.3, the field of $\mathcal{PT}$-symmetric Hamiltonians gives rise to a (still unanswered) number of theoretical questions, even if the original idea of pseudo-Hermiticity has been formulated more than 60 years ago [1].

It has been stressed in the literature that the scalar product $\langle\psi|\phi\rangle_*$ as defined in equation (2.23) is not positive definite. This has been used as an argument against the viability of $\mathcal{PT}$-symmetric Hamiltonians for the description of natural phenomena. However, one may counter argue that the same problem persists with regard to

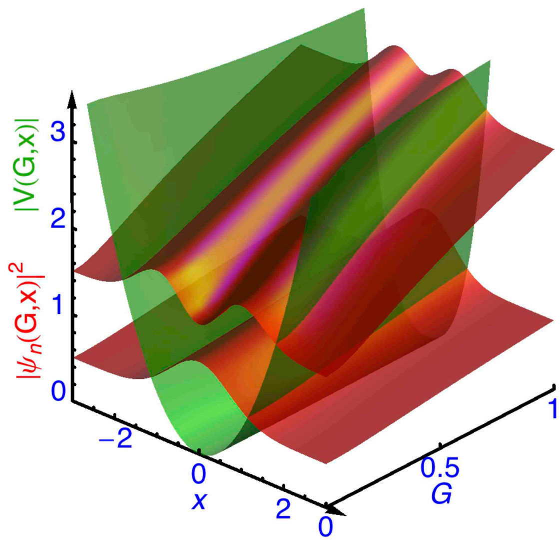


Figure 3.7: Illustration of the confinement mechanism for the imaginary cubic potential described by the Hamiltonian (3.1), for the ground and the first excited state. The bulk of the modulus square of the wave–function is centered in the "allowed" region where the (real rather than complex) energy $E > V(x) = V(G, x) = |W(G, x)|$. The potential is plotted in green, the complex moduli of the wave–functions are plotted in red. The ground state wave–function has a modulus square $|\psi_0(x)|^2 = |\psi_0(G, x)|^2$ as a function of G and x. As G increases, the bound-state energy (which is equal to the base line of the wave–function curve at any given value of G) increases, and the modulus of the potential forms a more narrow trough to which the ground-state wave–function is confined. The same is true for the first excited state. The central minimum of the complex modulus square of the first-excited state wave–function is clearly visible.

the Klein-Gordon equation where the time-like component of the current can become negative (see chapter 2 of reference [60]). For the Klein-Gordon equation, which describes a charged scalar field like (a component of) the Higgs field, one therefore has to reinterpret the zero component of the conserved Noether current as a charge density, not a probability density. Analogously, in the context of neutrino physics [44], an interpretation of a zero component of a conserved Noether current in a $\mathcal{PT}$-symmetric theory in terms of a weak interaction density has recently been proposed. No physical system is known which is directly described by an iGx^3 interaction, so the question of how to interpret the conserved density $\rho_*(x) = \psi(x)^+ \mathcal{P} \psi(x) = \psi(x)^+ \psi(-x)$ is perhaps not as immediate as in the other cases. It has found an application within physics in the context of generalized Bender–Wu formulas for odd anharmonic oscillators [5,12]. Still, it is reassuring to observe that quantum theories do not necessarily have to rely on a conserved positive-definite probability density; the Klein–Gordon theory is an example, where a re-interpretation of the probability as a charge density is accepted within the particle physics community and is an integral part of the accepted description of a fundamental spinless particle within the standard model, namely, the Higgs particle.

4. PSEUDO–HERMITICITY AND MATRIX DIAGONALIZATION

4.1. ORIENTATION

A number of powerful matrix diagonalization algorithms are known from the literature and have been implemented, i.e., in the `LAPACK` library [61]. Why should there be yet another effort at constructing numerical matrix diagonalization algorithms? The answer is that the numerical calculations reported in chapter 3 of this thesis were actually performed using an algorithm for matrix diagonalization, which is inspired by the physical structure of the pseudo–Hermitian quantum dynamics, notably the complex inner product (see equations (2.23) and (3.14)). These avoid the complex conjugation of the first argument.

In quantum mechanics, one often projects the Hamiltonian onto a (necessarily somewhat incomplete) finite subset of basis vectors, say, composed of eigenstates of the harmonic oscillator (chapter 7 of [59]). This semi–complete basis is then used to find approximations to energies and wave–functions of the system. It turns out that when they are projected onto an appropriate set of basis states, certain pseudo–Hermitian, or $\mathcal{PT}$–symmetric, Hamiltonians naturally give rise to complex symmetric matrices. In general, matrix diagonalization algorithms can be specialized based on the properties of the input matrix, as well as the subset of eigenvalues and eigenvectors one is interested in. For example, if only a specific eigenvalue of a sparsely populated matrix is desired, then “shooting” techniques such as the Arnoldi method, or variants thereof [62, 63], can be applied. For our purposes, we need to specialize in the diagonalization of complex symmetric matrices. Furthermore, we want the algorithm to be easily scalable in terms of numerical precision, requiring that the algorithm not be overly complicated. Arguably, the QL and QR algorithms constitute the conceptually simplest approaches to matrix diagonalization, as they iterate similarity transforms based on the decomposition of the input matrix into orthogonal ($\underline{Q}$) and left triangular ($\underline{L}$) or right triangular ($\underline{R}$) matrices [64,65], which eventually converges to a diagonal matrix. It has been shown that, in the event that the input matrix is an $(n \times n)$ tridiagonal matrix, the rate of convergence is λ_i/λ_{i+1}, for an ordered set of eigenvalues $|\lambda_1| < |\lambda_2| < ... < |\lambda_n|$ [66,67].

Our proposed (and tested) algorithm combines a generalization of the QL–factorization technique with an initial highly efficient tridiagonalization step, employing generalized Householder reflections. We refer to this algorithm as "Householder–based tridiagonalization followed by generalized QL decompositions with an implicit shift," or `HTDQLS` for short. The first step of the algorithm is to iteratively transform the input matrix $\underline{A}$ into tridiagonal form, utilizing a series of generalized Householder reflections. This step can be summarized as $\underline{T} = \underline{Z}^{-1}\underline{A}\,\underline{Z}$, where $\underline{Z} = \underline{H}_{n-1}\underline{H}_{n-2}\ldots\underline{H}_2$ is an orthogonal ($\underline{Z}^T = \underline{Z}^{-1}$) product of $n-2$ generalized Householder reflections. After calculating the tridiagonal matrix $\underline{T}$, we employ a generalization of the "chasing the bulge" strategy (see section 8.13 of [69]). In each iteration of this step, the algorithm calculates a guess, σ for a specific eigenvalue λ of $\underline{T}$, and calculates the QL decomposition for the tridiagonal matrix $\underline{T} - \sigma\,\mathbb{1}_{n\times n} = \underline{Q}\underline{L}$, then $\underline{T}' = \underline{L}\underline{Q} + \sigma\,\mathbb{1}_{n\times n} = \underline{Q}^{-1}\underline{T}\,\underline{Q}$. Eventually, this procedure will lead to the diagonalization of $\underline{T}$, up to machine precision, while preserving the tridiagonal structure of $\underline{T}$ in each step. Practically, this procedure manifests itself as the creation of a "bulge" (an off tridiagonal element), introduced by the initial rotation. The remaining rotations of each iteration are Givens rotations, which "chase", and eventually annihilate, the "bulge." This procedure relies on the super-/sub–diagonal elements going to zero (to machine precision) in order, starting at the top left corner, and working down to the bottom right corner. In the event that an element is zeroed before its turn, then "deflation" techniques are used to subdivide the input matrix into two smaller matrices, each of which must be diagonalized separately. For large matrices it may become necessary to apply the deflation procedure recursively.

Here we go into greater detail than in Part II of [54] in describing the steps taken by the algorithm. We begin with a discussion of the tridiagonalization step in chapter 4.2, which includes a discussion of an alternative diagonalization method. In chapter 4.3 we discuss the diagonalization step, including a discussion of the deflation procedure. In chapter 4.4 we briefly go over how one might implement the QR version of the algorithm. Finally in chapter 4.5 we include numerical reference data. In chapter 5 we include a brief discussion of the implementation of the algorithm, while an explicit `FORTRAN` implementation can be found in appendix A.

4.2. TRIDIAGONALIZATION

4.2.1. Householder Reflections and Hermitian Matrices. Before we discuss the generalization of Householder reflections let us review the traditional (not generalized) Householder matrices. By definition the Householder reflection is given as [70]

$$\underline{\underline{H}}_{\underline{v}}\underline{x} \equiv \underline{x} - \frac{2}{\langle \underline{v}, \underline{v} \rangle}\langle \underline{v}, \underline{x} \rangle \underline{v}, \tag{4.1}$$

and can be rewritten as

$$\underline{\underline{H}}_{\underline{v}}\underline{x} = \underline{x} - 2\langle \underline{u}, \underline{x} \rangle \underline{u}\,, \tag{4.2}$$

where

$$\underline{u} = \frac{\underline{v}}{|\underline{v}|} \quad \text{and} \quad |\underline{x}| \equiv \sqrt{\langle \underline{x}, \underline{x} \rangle}\,. \tag{4.3}$$

By definition, the scalar product is $\langle \underline{x}, \underline{y} \rangle = \underline{x}^{+}\underline{y}$, and $\underline{x}^{+} = \underline{x}^{T*} = \underline{x}^{*T}$. From here we find that

$$\underline{\underline{H}}_{\underline{v}} \equiv \mathbb{1} - 2\underline{u} \otimes \underline{u}^{+}\,, \tag{4.4}$$

where $\mathbb{1}$ is the identity matrix, and $\otimes$ is the tensor (dyadic) product. We can now show that $\underline{H}_{\underline{v}}$ is Hermitian,

$$H_{\underline{v}ij} = \delta_{ij} - 2\, u_i\, u_j^* = \delta_{ji} - 2\, u_j^*\, u_i = (\delta_{ji} - 2\, u_j\, u_i^*)^* = H_{\underline{v}ji}^*\,, \tag{4.5}$$

i.e.,

$$\underline{H}_{\underline{v}} = \underline{H}_{\underline{v}}^+\,, \tag{4.6}$$

and subsequently that $\underline{H}_{\underline{v}}$ is unitary

$$\underline{H}_{\underline{v}}^2 = \underline{H}_{\underline{v}}\underline{H}_{\underline{v}}^+ = \underline{H}_{\underline{v}}^+\underline{H}_{\underline{v}} = \left(\mathbb{1} - 2\underline{u} \otimes \underline{u}^+\right)^2 = \mathbb{1} - 4\underline{u} \otimes \underline{u}^+ + 4\,\underline{u} \otimes \left[(\underline{u}^+\underline{u})\underline{u}^+\right]\,. \tag{4.7}$$

Now

$$\underline{u}^+\underline{u} = \frac{\underline{v}^+\underline{v}}{|\underline{v}|^2} = \frac{\langle \underline{v}, \underline{v}\rangle}{\langle \underline{v}, \underline{v}\rangle} = 1\,, \tag{4.8}$$

and by plugging this result into (4.7) we quickly find

$$\underline{H}_{\underline{v}}^2 = \mathbb{1}\,. \tag{4.9}$$

Now if we set $\underline{v} = \underline{y} + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n$ (where $\underline{x}^+\hat{\mathbf{e}}_n = x_n^*$ and $\hat{\mathbf{e}}^+\underline{x} = x_n$, where x_n is the n^{th} element of $\underline{x}$), where $y_n = |y_n|\mathrm{e}^{\mathrm{i}\theta}$, then

$$\begin{aligned}
\underline{H}_{\underline{v}}\underline{y} &= \underline{y} - 2\,\langle \underline{u}, \underline{y}\rangle \underline{u} = \underline{y} - \frac{2}{|\underline{v}|^2}\,\underline{v}^+\underline{y}\,\underline{v} = \underline{y} - \frac{2\left(\underline{y} + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n\right)^+\underline{y}}{\left(\underline{y} + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n\right)^+\left(\underline{y} + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n\right)}\underline{v} \\
&= \underline{y} - \frac{2\left(\underline{y}^+ + \mathrm{e}^{-\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n^+\right)\underline{y}}{\left(\underline{y}^+ + \mathrm{e}^{-\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n^+\right)\left(\underline{y} + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n\right)}\underline{v} = \underline{y} - \frac{2\left(|\underline{y}|^2 + \mathrm{e}^{-\mathrm{i}\theta}|\underline{y}|y_n\right)}{|\underline{y}|^2 + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}|y_n + \mathrm{e}^{-\mathrm{i}\theta}|\underline{y}|y_n^* + |\underline{y}|^2}\underline{v} \\
&= \underline{y} - \frac{2\left(|\underline{y}|^2 + \mathrm{e}^{-\mathrm{i}\theta}|\underline{y}||y_n|\mathrm{e}^{\mathrm{i}\theta}\right)}{2|\underline{y}|^2 + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}||y_n|\mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta}|\underline{y}||y_n|\mathrm{e}^{-\mathrm{i}\theta}}\underline{v} = \underline{y} - \frac{2\left(|\underline{y}|^2 + |\underline{y}||y_n|\right)}{2|\underline{y}|^2 + 2|\underline{y}||y_n|}\underline{v} = \underline{y} - \underline{v} \\
&= -\,|\underline{y}|\mathrm{e}^{\mathrm{i}\theta}\hat{\mathbf{e}}_n\,.
\end{aligned} \tag{4.10}$$

If $\underline{y}$ is real, as it is when we tridiagonalize real symmetric matrices, then we immediately have $\theta = 0$, and

$$\underline{H}_{\underline{v}}\underline{y} = -|\underline{y}|\hat{\mathbf{e}}_n \,. \tag{4.11}$$

These properties (equations (4.10) and (4.11)) are essential when tridiagonalizing a Hermitian (or real symmetric) matrix [66, 69]. As such, we want the generalization of the Householder matrices to have a similar property when applied to a complex symmetric (non–Hermitian) matrix. Implicit within this discussion is that for Householder reflections operating on a Hermitian matrix, in addition to (4.10), we also want $\underline{y}^+\underline{H}_{\underline{v}}$ to reduce so a single element row matrix, which we can easily show it does:

$$\underline{y}\,\underline{H}_{\underline{v}} = \left(\underline{H}_{\underline{v}}\underline{y}\right)^+ = -|\underline{y}|\mathrm{e}^{-\mathrm{i}\theta}\hat{\mathbf{e}}_n^+ \,. \tag{4.12}$$

When considering a complex symmetric matrix, we need $\underline{H}_{\underline{v}}\underline{y}$ and $\underline{y}^T\underline{H}_{\underline{v}}$ to both reduce to single element row matrices (as we will see in chapter 4.2.4). If we use the usual definition, we will again find that the former does reduce to the desired form (4.10), however

$$\underline{y}^T\underline{H}_{\underline{v}} = \underline{y}^T\underline{H}_{\underline{v}}^+ = (\underline{H}_{\underline{v}}\underline{y}^*)^+ \,, \tag{4.13}$$

and

$$\begin{aligned}
\underline{H}_{\underline{v}}\underline{y}^* &= \underline{y}^* - 2\,\langle\underline{u},\underline{y}^*\rangle\underline{u} = \underline{y}^* - \frac{2}{|\underline{v}|^2}\,\underline{v}^+\underline{y}\,\underline{v} = \underline{y}^* - \frac{2\left(\underline{y} + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n\right)^+\underline{y}^*}{\left(\underline{y} + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n\right)^+\left(\underline{y} + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n\right)}\underline{v} \\
&= \underline{y}^* - \frac{2\left(\underline{y}^+ + \mathrm{e}^{-\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n^+\right)\underline{y}^*}{\left(\underline{y}^+ + \mathrm{e}^{-\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n^+\right)\left(\underline{y} + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}|\hat{\mathbf{e}}_n\right)}\underline{v} \\
&= \underline{y}^* - \frac{2\left(|\underline{y}|^2 + \mathrm{e}^{-\mathrm{i}\theta}|\underline{y}||y_n|\mathrm{e}^{-\mathrm{i}\theta}\right)}{2|\underline{y}|^2 + \mathrm{e}^{\mathrm{i}\theta}|\underline{y}||y_n|\mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta}|\underline{y}||y_n|\mathrm{e}^{-\mathrm{i}\theta}}\underline{v} = \underline{y}^* - \frac{2\left(|\underline{y}|^2 + |\underline{y}||y_n|\mathrm{e}^{-2\mathrm{i}\theta}\right)}{2|\underline{y}|^2 + 2|\underline{y}||y_n|}\underline{v}\,.
\end{aligned} \tag{4.14}$$

Thus $\underline{y}^T \underline{H}_{\underline{v}}$ does not reduce to a single element row matrix and we cannot use Householder reflections (as defined in the usual way) to tridiagonalize complex symmetric matrices.

4.2.2. Generalized Householder Reflections. Drawing inspiration from the indefinite inner product [see equations (2.23) and (3.14)] which doesn't have any complex conjugation, we define the indefinite scalar product, which avoids complex conjugation, as

$$\langle \underline{x}, \underline{y} \rangle_* = \underline{x}^T \underline{y} \,, \tag{4.15}$$

for complex matrices. We now modify our definition of the Householder reflections accordingly, and define the generalized Householder reflection as

$$\underline{H}_{\underline{v}} \underline{x} \equiv \underline{x} - \frac{2}{\langle \underline{v}, \underline{v} \rangle_*} \langle \underline{v}, \underline{x} \rangle_* \underline{v} \,. \tag{4.16}$$

From this result we quickly find that

$$\underline{H}_{\underline{v}} = \mathbb{1} - 2\underline{u} \otimes \underline{u}^T \,, \tag{4.17}$$

where $\otimes$ is the tensor (dyadic) product, and

$$\underline{u} = \frac{\underline{v}}{|\underline{v}|_*} \quad \text{and} \quad |\underline{x}|_* = \sqrt{\langle \underline{x}, \underline{x} \rangle_*} \,. \tag{4.18}$$

This definition is in agreement with the first (unnumbered) equation in [73] and equation (1) in [74], the latter of which uses an unnecessary normalization of the vector $\underline{v}$. We can now show that $\underline{H}_{\underline{v}}$ is symmetric,

$$H_{\underline{v}ij} = \delta_{ij} - 2\, u_i\, u_j = \delta_{ji} - 2\, u_j\, u_i = H_{\underline{v}ji} \,, \tag{4.19}$$

i.e.,

$$\underline{H}_{\underline{v}} = \underline{H}_{\underline{v}}^{T} \,, \tag{4.20}$$

and subsequently that $\underline{H}_{\underline{v}}$ is a square root of unity

$$\begin{aligned} \underline{H}_{\underline{v}}^{2} &= \underline{H}_{\underline{v}}\underline{H}_{\underline{v}}^{T} = \underline{H}_{\underline{v}}^{T}\underline{H}_{\underline{v}} = \left(\mathbb{1} - 2\underline{u} \otimes \underline{u}^{T}\right)^{2} = \mathbb{1} - 4\underline{u} \otimes \underline{u}^{T} + 4\,\underline{u} \otimes \left[(\underline{u}^{T}\underline{u})\underline{u}^{T}\right] \\ &= \mathbb{1} - 4\underline{u} \otimes \underline{u}^{T} + 4\,\underline{u} \otimes \underline{u}^{T} = \mathbb{1}\,. \end{aligned} \tag{4.21}$$

Now if we set

$$\underline{v} = \underline{y} + |\underline{y}|_{*}\hat{\mathbf{e}}_{n} \,, \tag{4.22}$$

then

$$\begin{aligned} \underline{H}_{\underline{v}}\underline{y} &= \underline{y} - 2\,\langle \underline{u}, \underline{y}\rangle \underline{u} = \underline{y} - \frac{2}{|\underline{v}|^{2}}\,\underline{v}^{T}\underline{y}\,\underline{v} = \underline{y} - \frac{2\left(\underline{y} + |\underline{y}|_{*}\hat{\mathbf{e}}_{n}\right)^{T}\underline{y}}{\left(\underline{y} + |\underline{y}|_{*}\hat{\mathbf{e}}_{n}\right)^{T}\left(\underline{y} + |\underline{y}|_{*}\hat{\mathbf{e}}_{n}\right)}\underline{v} \\ &= \underline{y} - \frac{2\left(\underline{y}^{T} + |\underline{y}|_{*}\hat{\mathbf{e}}_{n}^{+}\right)\underline{y}}{\left(\underline{y}^{T} + |\underline{y}|_{*}\hat{\mathbf{e}}_{n}^{T}\right)\left(\underline{y} + |\underline{y}|_{*}\hat{\mathbf{e}}_{n}\right)}\underline{v} = \underline{y} - \frac{2\left(|\underline{y}|_{*}^{2} + |\underline{y}|_{*}y_{n}\right)}{2|\underline{y}|_{*}^{2}2|\underline{y}|_{*}y_{n}}\underline{v} \\ &= \underline{y} - \underline{v} = -|\underline{y}|_{*}\hat{\mathbf{e}}_{n} \,, \end{aligned} \tag{4.23}$$

and

$$\underline{y}^{T}\underline{H}_{\underline{v}} = \underline{y}^{T}\underline{H}_{\underline{v}}^{T} = (\underline{H}_{\underline{v}}\underline{y})^{T} = -|\underline{y}|_{*}\hat{\mathbf{e}}_{n}^{T} \,. \tag{4.24}$$

As we will see in the following section, by using the indefinite scalar product it has become quite simple to tridiagonalize complex symmetric matrices.

(Remark: It should be noted that not all complex symmetric matrices are diagonalizable, such as

$$\begin{pmatrix} \mathrm{i} & 1 \\ 1 & -\mathrm{i} \end{pmatrix} ,$$

whose determinant is 0. Our additional implicit assumption is that the matrix we are working on is diagonalizable.)

4.2.3. Matrix Diagonalization using solely Householder Reflections.

While the central theme of this chapter is the description of an algorithm designed to diagonalize complex symmetric (not Hermitian) matrices in two steps (tridiagonalization followed by diagonalization), there is an alternative approach, namely plain QL and QR decomposition. We take the opportunity to briefly describe the underlying process, while again specializing in complex symmetric (non-Hermitian) matrices. Both the plain QL and QR decompositions exclusively use generalized Householder reflections (see chapter 4.2.2) to diagonalize a matrix, and are performed using one step, in which the following procedure is iterated until the input matrix has been diagonalized to machine accuracy. Versions of the algorithms have been implemented using `FORTRAN` and are included in appendix B. We begin by describing the plain QL decomposition for an $n \times n$ complex symmetric (not Hermitian) matrix $\underline{A}$. In the first step, we set $\underline{y}_n$ equal to the final column of $\underline{A}$,

$$\underline{y}_n = \begin{pmatrix} A_{1n} \\ \vdots \\ A_{nn} \end{pmatrix}, \tag{4.25}$$

from which we construct $\underline{v}_n$ as prescribed by (4.22) and then set $\underline{H}_n = \underline{H}_{\underline{v}_n}$ which is defined in (4.16). Then $\underline{L}_n = \underline{H}_n \underline{A}$ will have the form

$$\underline{L}_n = \left(\begin{array}{ccc|c} (\underline{L}_n)_{1\,1} & \dots & (\underline{L}_n)_{1\,n-1} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ (\underline{L}_n)_{n-1\,1} & \dots & (\underline{L}_n)_{n-1\,n-1} & 0 \\ (\underline{L}_n)_{n\,1} & \dots & (\underline{L}_n)_{n\,n-1} & -|\underline{y}_n|_* \end{array}\right).$$

We now proceed by setting $\underline{y}_{n-1}$ equal to the first $(n-1)$ elements of the second to last column of $\underline{L}_n$. The column matrix $\underline{v}_{n-1}$ is then calculated, followed by $\underline{H}_{\underline{v}_{n-1}}$.

We then set

$$\underline{H}_{n-1} = \begin{pmatrix} \underline{H}_{\underline{v}_{n-1}} & \\ & \mathbb{1}_{1\times 1} \end{pmatrix}, \tag{4.26}$$

and find $\underline{L}_{n-1} = \underline{H}_{\underline{v}_{n-1}} \underline{H}_{\underline{v}_n} \underline{A}$ which will be a left triangular matrix in the last two columns,

$$\underline{L}_{n-1} = \left(\begin{array}{ccc|cc} (\underline{L}_n)_{1\,1} & \cdots & (\underline{L}_n)_{1\,n-2} & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ (\underline{L}_n)_{n-1\,1} & \cdots & (\underline{L}_n)_{n-1\,n-2} & -|\underline{y}_{n-1}| & 0 \\ (\underline{L}_n)_{n\,1} & \cdots & (\underline{L}_{n-1})_{n\,n-2} & (\underline{L}_{n-1})_{n\,n-1} & -|\underline{y}_n|_* \end{array}\right). \tag{4.27}$$

We repeat this procedure, setting $\underline{y}_i$ equal to the first i elements of the last column of $\underline{L}_{i+1}$ that is not in the left triangular form. We then calculate $\underline{v}_i$, from which the generalized Householder reflection, $\underline{H}_{\underline{v}_i}$, is constructed. Then, by the application of the Householder reflection $\underline{H}_i$ =diag$(\underline{H}_{\underline{v}_i}, \mathbb{1}_{(n-i)\times(n-i)})$, the matrix $\underline{L}_i$ is generated. The index i runs from $i = n$ down to $i = 2$. After a total of $(n-1)$ Householder reflections, we will have

$$\underline{L} = \underline{L}_2 = \underline{H}\,\underline{A}\,, \qquad \underline{H} = \underline{H}_2\,\underline{H}_3 \ldots \underline{H}_n\,, \tag{4.28a}$$

$$\underline{A} = \underline{Q}\,\underline{L}\,, \qquad \underline{Q} = \underline{H}^{-1} = \underline{H}^{\mathrm{T}} = \underline{H}_n^{\mathrm{T}} \ldots \underline{H}_2^{\mathrm{T}}\,, \tag{4.28b}$$

where $\underline{L}$ is a left triangular matrix. We can now rotate $\underline{A}$ into $\underline{A}'$,

$$\underline{A}' = \underline{Q}^{\mathrm{T}}\,\underline{A}\,\underline{Q} = \underline{L}\,\underline{Q}\,. \tag{4.29}$$

In general $\underline{A}'$ will not be a diagonal matrix, and as such we need to repeat the decomposition for $\underline{A}'$, following the procedure described above. Once the decomposition is performed, we find

$$\underline{A}' = \underline{Q}'\,\underline{L}'\,, \quad \underline{A}'' = \underline{L}'\,\underline{Q}'\,, \tag{4.30}$$

and after k QL decompositions,

$$\underline{A}^{(k)} = \underline{Q}^{(k)}\,\underline{L}^{(k)}\,, \quad \underline{A}^{(k+1)} = \underline{L}^{(k)}\,\underline{Q}^{(k)}\,. \tag{4.31}$$

Similarly, we can use the generalized Householder reflections to perform a QR decomposition of $\underline{A}$. For the QR decomposition we begin by setting $\underline{y}_1$ equal to the first column of $\underline{A}$,

$$\underline{y}_1 = \begin{pmatrix} A_{11} \\ \vdots \\ A_{nn} \end{pmatrix}, \tag{4.32}$$

from which calculate $\underline{v}_1 = \underline{y}_1 + |\underline{y}|_* \hat{e}_1$, and construct the corresponding Householder reflection, $\underline{H}_{\underline{v}_1}$. Then $\underline{R}_1 = \underline{H}_1\,\underline{A}$ will be

$$\underline{R}_1 = \left(\begin{array}{c|ccc} -|\underline{y}_1|_* & (\underline{R}_1)_{1\,2} & \dots & (\underline{R}_1)_{1\,n} \\ 0 & \vdots & \ddots & \vdots \\ \vdots & (\underline{R}_1)_{n-1\,2} & \dots & (\underline{R}_1)_{n-1\,n} \\ 0 & (\underline{R}_1)_{n\,2} & \dots & (\underline{R}_1)_{n\,n} \end{array}\right). \tag{4.33}$$

We then choose $\underline{y}_2$ to be the last $n-1$ elements of $\underline{R}_1$, find $\underline{v}_2$ and $\underline{H}_{\underline{v}_2}$ and set

$$\underline{H}_2 = \begin{pmatrix} \mathbb{1}_{1\times 1} & \\ & \underline{H}_{\underline{v}_2} \end{pmatrix}, \tag{4.34}$$

and continue the process as we did for the QL case, only this time i runs from $i = 1$ to $i = n-1$ and we are finding $\underline{R}_i$. After the $(n-1)$ generalized Householder reflections are performed we find

$$\underline{R} = \underline{R}_{n-1} = \underline{H}\,\underline{A}\,, \qquad \underline{H} = \underline{H}_{n-1}\,\underline{H}_{n-2}\dots\underline{H}_1\,, \tag{4.35a}$$

$$\underline{A} = \underline{Q}\,\underline{R}\,, \qquad \underline{Q} = \underline{H}^{-1} = \underline{H}^{\mathrm{T}} = \underline{H}_1^{\mathrm{T}}\dots\underline{H}_{n-1}^{\mathrm{T}}\,, \tag{4.35b}$$

where $\underline{R}$ is a right triangular matrix. After k QR decompositions we have

$$\underline{A}^{(k)} = \underline{Q}^{(k)}\,\underline{R}^{(k)}\,, \quad \underline{A}^{(k+1)} = \underline{R}^{(k)}\,\underline{Q}^{(k)}\,. \tag{4.36}$$

For both cases (QL and QR), after a sufficient number of iterations, m, is performed $\underline{A}^{(m)}$ will be a diagonal matrix.

The "plain QL" (`PQL`) and "plain QR" (`PQR`) algorithms described above are not the most efficient way to find the eigenvalues and eigenvectors of a complex symmetric (non-Hermitian) matrix, yet they are included here because of their versatility. These simple routines are very easy to implement, as well as being easily scalable. Finally they provide for surprisingly robust algorithms. Again, more on this can be found in appendix B.

4.2.4. Procedure: First Step of the `HTDQLS` Algorithm. With the preparations given above, it is possible to now switch to the description of the first step of our proposed matrix algorithm, `HTDQLS` (see figure 4.1). Concerning the first step (tridiagonalization), we should add that in principle, it is possible to use a variety of methods to bring complex symmetric matrices into tridiagonal form. For instance, Cullum and Willoughby have shown that it is possible to use the Lanczos method to tridiagonalize complex symmetric matrices [71, 72], yet we have chosen to employ generalized Householder reflections to accomplish the same goal (see equation (4.16) of chapter 4.2.2). These other methods are primarily useful when a subset of eigenvalues (e.g., those of largest magnitude) are to be determined. Here, we are specifically

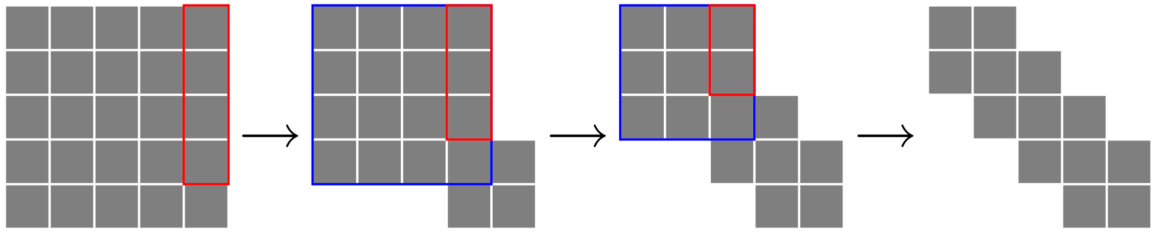


Figure 4.1: These figures represent the tridiagonalization of a 5×5 complex symmetric matrix, generated by $A_{ij} = \frac{\mathrm{i}^{i+j+1}}{i+j+1}$. In the first step we choose our y_4 column matrix, as the first 4 elements of the 5^{th} column of A (outlined in red). We then set $v_4 = y_4 + \sqrt{\langle y_4, y_4 \rangle_*}\,\hat{e}_4$, construct H_{v_4} and finally construct our rotation matrix H_4. We then rotate the matrix into $A' = H_4\,A\,H_4$. In doing so we have eliminated the off-tridiagonal elements in the final column and row. We then repeat the process, however instead of using the entire matrix to construct H_3 we focus only on the part of the matrix that is still not tridiagonal (outlined in blue). We can now define y_3 as the first 3 elements of the 2^{nd} to last column in A', which we use to create v_3 then H_{v_3} and finally H_3. Then $A'' = H_3\,A'\,H_3$. We repeat the process a final time, giving us a tridiagonal matrix, A''', where, in this case, $A''' = H_2 H_3 H_4\,A\,H_5 H_3 H_2$.

concerned with the full tridiagonalization of the input matrix, and therefore choose a generalization of the method of Householder transformations. This method contrasts the `PQL` algorithm, as it requires a total of $(n-2)$ generalized Householder reflections to tridiagonalize, and subsequently diagonalize, the input matrix, while the `PQL` algorithm requires $(n-1)$ generalized Householder reflections per iteration.

While the concept of using the generalized Householder reflections to tridiagonalize a complex symmetric matrix has been mentioned in reference [73, 74], the implementation of the precise calculation procedure is not always made clear. In reference [74], because of a lack of true complex arithmetic, the complex symmetric matrix is separated into real and imaginary parts, causing each step to require two Householder reflections, as well as an additional unitary transform. By contrast, we here use an algorithm with a single generalized Householder reflection in each step. In the following, we endeavor to clarify the procedure utilized by our algorithm. While the procedure described here is similar to the `PQL` procedure described in chapter 4.2.3, there are differences between the two. Moreover the procedure described in the following is designed to tridiagonalize the input matrix, rather than to diagonalize it (as in chapter 4.2.3). Despite the similarities, it is worth going through the tridiagonalization process in detail.

The tridiagonalization of a complex symmetric matrix of rank n can now be performed using $n-2$ generalized Householder reflections. Simply because it provides for clearer notation, we index the steps (in order) as $i = n-1$ to $i = 2$ in steps of 1. Let $\underline{A}$ be the matrix we want to tridiagonalize, then in the first step we choose $\underline{y}_{n-1}$ to be the first $n-1$ elements of the last column of $\underline{A}$,

$$\underline{y}_{n-1} = \begin{pmatrix} A_{1n} \\ A_{2n} \\ \vdots \\ A_{n-1n} \end{pmatrix} . \tag{4.37}$$

By defining $\underline{B}_{n-1}$ as an $n-1 \times n-1$ matrix, where $B_{ij} = A_{ij}$, we can then write $\underline{A}$ as

$$\underline{A} = \left(\begin{array}{c|c} \underline{B}_{n-1} & \underline{y}_{n-1} \\ \hline \underline{y}^T_{n-1} & A_{nn} \end{array} \right) . \tag{4.38}$$

We then calculate $\underline{v}_{n-1}$ as

$$\underline{v}_{n-1} = \underline{y}_{n-1} + |\underline{y}_{n-1}|_* \hat{\mathbf{e}}_{n-1} . \tag{4.39}$$

We now construct $\underline{H}_{\underline{v}_{n-1}}$, which will be a Householder matrix of rank $n-1$. We then define $\underline{H}_{n-1}$ as

$$\underline{H}_{n-1} = \begin{pmatrix} \underline{H}_{\underline{v}_{n-1}} & 0 \\ 0 & \mathbb{1}_{1\times 1} \end{pmatrix}. \tag{4.40}$$

Then

$$\underline{H}_{n-1}\,\underline{A} = \left(\begin{array}{c|c} \underline{H}_{\underline{v}_{n-1}}\underline{B}_{n-1} & \underline{H}_{\underline{v}_{n-1}}\underline{y}_{n-1} \\ \hline \underline{y}^T_{n-1} & A_{nn} \end{array}\right), \tag{4.41}$$

and thusly

$$\underline{H}_{n-1}\,\underline{A}\,\underline{H}_{n-1} = \left(\begin{array}{c|c} \underline{H}_{\underline{v}_{n-1}}\underline{B}_{n-1}\underline{H}_{\underline{v}_{n-1}} & \underline{H}_{\underline{v}_{n-1}}\underline{y}_{n-1} \\ \hline \underline{y}^T_{n-1}\underline{H}_{\underline{v}_{n-1}} & A_{nn} \end{array}\right). \tag{4.42}$$

Using equations (4.23) and (4.24) this reduces to

$$\underline{A}' = \left(\begin{array}{ccc|c} & & & 0 \\ & \underline{B}'_{n-1} & & \vdots \\ & & & |\underline{y}_{n-1}|_* \\ \hline 0 & \cdots & |\underline{y}_{n-1}|_* & A_{nn} \end{array}\right), \tag{4.43}$$

where

$$\underline{A}' = \underline{H}_{n-1}\,\underline{A}\,\underline{H}_{n-1}, \quad \underline{B}'_{n-1} = \underline{H}_{\underline{v}_{n-1}}\underline{B}_{n-1}\underline{H}_{\underline{v}_{n-1}}. \tag{4.44}$$

For the second step we choose $\underline{y}_{n-2}$ to be the first $n-2$ elements of the second to last column of $\underline{A}'$ (which is the same as the last column of $\underline{B}'_{n-1}$),

$$\underline{y}_{n-2} = \begin{pmatrix} A'_{1n-1} \\ A'_{2n-1} \\ \vdots \\ A'_{n-2n-1} \end{pmatrix} = \begin{pmatrix} B'_{1n-1} \\ B'_{2n-1} \\ \vdots \\ B'_{n-2n-1} \end{pmatrix} . \tag{4.45}$$

We then calculate $\underline{v}_{n-2}$, construct $\underline{H}_{\underline{v}_{n-2}}$, which will be a Householder matrix of rank $n-2$, and define $\underline{H}_{n-2}$ as

$$\underline{H}_{n-2} = \begin{pmatrix} \underline{H}_{\underline{v}_{n-2}} & 0 \\ 0 & \mathbb{1}_{2\times 2} \end{pmatrix} . \tag{4.46}$$

Then

$$\underline{A}'' = \underline{H}_{n-2}\,\underline{A}'\,\underline{H}_{n-2}\,, \tag{4.47}$$

and

$$\underline{A}'' = \left(\begin{array}{ccc|c|c} & & & 0 & 0 \\ & \underline{B}'_{m-1} & & \vdots & \vdots \\ & & & |\underline{y}_{n-2}|_* & 0 \\ \hline 0 & \cdots & |\underline{y}_{n-2}|_* & A'_{mm} & |\underline{y}_{n-1}|_* \\ \hline 0 & \cdots & 0 & |\underline{y}_{n-1}|_* & A_{nn} \end{array}\right) , \tag{4.48}$$

where $m = n-1$. We then repeat the process, until we have a tridiagonal matrix. As previously mentioned, this will take a total of $n-2$ rotations.

What makes this process so convenient is that at no point do we have to actually calculate an entire matrix, $\underline{A}'$ (or any of the subsequent matrices). Instead, due to the properties of the Householder reflections we already know what the last row and column of the new matrix will be, and only have to consider

$$\underline{B}' = \underline{H}_{\underline{v}}\,\underline{B}\,\underline{H}_{\underline{v}} = \left(\mathbb{1} - \frac{2}{|\underline{v}^2|_*}\underline{v}\otimes\underline{v}^T\right)\,\underline{B}\,\left(\mathbb{1} - \frac{2}{|\underline{v}|_*^4}\underline{v}\otimes\underline{v}^T\right)$$
$$= \underline{B} - \frac{2}{|\underline{v}|_*^2}(\underline{v}\otimes\underline{v}^T\,\underline{B} + \underline{B}\,\underline{v}\otimes\underline{v}^T) + \frac{4}{|\underline{v}|_*^4}\,\underline{v}\otimes\underline{v}^T\,\underline{B}\,\underline{v}\otimes\underline{v}^T\,. \tag{4.49}$$

Introducing $p = \frac{1}{2}|\underline{v}|_*^2$ this becomes

$$\underline{B}' = \underline{B} - \frac{1}{p^2}(\underline{v}\otimes\underline{v}^T\,\underline{B} + \underline{B}\,\underline{v}\otimes\underline{v}^T) + \frac{1}{p^2}\,\underline{v}\otimes\underline{v}^T\,\underline{B}\,\underline{v}\otimes\underline{v}^T$$
$$= \underline{B} - \underline{v}\otimes\left(\frac{\underline{B}\underline{v}}{p}\right)^T - \left(\frac{\underline{B}\underline{v}}{p}\right)\otimes\underline{v}^T + \underline{v}\otimes\frac{\underline{v}^T}{p}\left(\frac{\underline{B}\underline{v}}{p}\right)\otimes\underline{v}^T\,. \tag{4.50}$$

We now define $\underline{u} = \frac{\underline{B}\underline{v}}{p}$ and $q = \frac{\underline{v}^T\underline{u}}{2p}$, allowing us to rewrite $\underline{B}'$ as

$$\underline{B}' = \underline{B} - \underline{v}\otimes\underline{u}^T - \underline{u}\otimes\underline{v}^T + 2\underline{v}\otimes\left[\left(\frac{\underline{v}^T\underline{u}}{2p}\right)\underline{v}^T\right] = \underline{B} - \underline{v}\otimes\underline{u}^T - \underline{u}\otimes\underline{v}^T + 2q\underline{v}\otimes\underline{v}^T$$
$$= \underline{B} - \underline{v}\otimes(\underline{u} - q\underline{v})^T - (\underline{u} - q\underline{v})\otimes\underline{v}^T\,, \tag{4.51}$$

and finally we define $\underline{w} = \underline{u} - q\underline{v}$, from which we get

$$\underline{B}' = \underline{B} - \underline{v}\otimes\underline{w}^T - \underline{w}\otimes\underline{v}^T\,. \tag{4.52}$$

Armed with this result we can now calculate each step by first choosing $\underline{y}$ and $\underline{B}$. We then calculate $|\underline{y}|_*$ and $\underline{v}$ then p, $\underline{u}$, q and $\underline{w}$. We can then find $\underline{B}'$ and finally construct $\underline{A}'$.

After we complete $n-2$ iterations of this procedure, $\underline{A}$ will be in tridiagonal form, i.e.,

$$\underline{T} = \underline{Z}^{-1}\,\underline{A}\,\underline{Z}\,, \tag{4.53}$$

where

$$\underline{Z} = \underline{H}_{n-1}\underline{H}_{n-2}\ldots\underline{H}_2\,, \tag{4.54}$$

and

$$\underline{Z}^{-1} = \underline{H}_2\underline{H}_3 \ldots \underline{H}_{n-1} \,. \tag{4.55}$$

4.3. DIAGONALIZATION

4.3.1. Implicit Shift. Cullum and Willoughby have treated the QL decomposition in reference [71, 75], as well as provided an algorithm in reference [72]. Here we provide a more illustrative discussion on the procedure implemented in our algorithm.

Now that we have managed to reduce the starting matrix to a tridiagonal form, we can begin working on the diagonalization of $\underline{T}$. Each iteration of this QL decomposition manifests itself as an implicitly shifted initial rotation, followed by a series of generalized Givens rotations. We zero out the super-/sub–diagonal elements, proceeding from the top left corner to the bottom right corner by iterating the transformation

$$\underline{T}^{(k)} - \sigma_k \, \mathbb{1}_{n\times n} = \underline{Q}^{(k)} \, \underline{L}^{(k)} \,, \tag{4.56a}$$

$$\underline{T}^{(k+1)} = \underline{L}^{(k)}\underline{Q}^{(k)} + \sigma_k \, \mathbb{1}_{n\times n} = \left(\underline{Q}^{(k)}\right)^T \underline{T}^{(k)}\underline{Q}^{(k)} \,. \tag{4.56b}$$

Thus after each iteration the matrix returns to tridiagonal form. If we do not include the shift ($\sigma_i = 0$), then the super-/sub–diagonal elements we are focusing on will converge like

$$T^{(k)}_{i\,i+1} \propto \left(\frac{\lambda_i}{\lambda_{i+1}}\right)^k \,. \tag{4.57}$$

If instead we consider a non–zero guess ($\sigma_k \neq 0$), and λ_i are the eigenvalues of $\underline{T}$ with eigenvector $\underline{x}_i$, then

$$(\underline{T}^{(k)} - \sigma_k \mathbb{1}_{n\times n})\underline{x}_i = (\lambda_i - \sigma_k) x_i \,, \tag{4.58}$$

and our rate of convergence becomes,

$$T^{(k)}_{i\,i+1} \propto \left(\frac{\lambda_i - \sigma_k}{\lambda_{i+1} - \sigma_k}\right)^k . \tag{4.59}$$

Thus the closer our guess, σ_k, is to λ_i, the faster the convergence [66, 67].

To calculate our shift, let us consider the tridiagonal input matrix, which is of the form

$$\underline{T} = \begin{pmatrix} D_1 & E_1 & & & \\ E_1 & D_2 & E_2 & & \\ & E_2 & \ddots & \ddots & \\ & & \ddots & D_{n-1} & E_{n-1} \\ & & & E_{n-1} & D_n \end{pmatrix} . \tag{4.60}$$

The usual choice for a shift is the Wilkinson shift [66, 67], which is obtained by calculating the eigenvalues of the 2×2 submatrix containing the elements we wish to zero (E_i), i.e.,

$$\begin{pmatrix} D_i & E_i \\ E_i & D_{i+1} \end{pmatrix} . \tag{4.61}$$

The eigenvalues of this matrix are

$$\sigma_k = D^{(k)}_i + E^{(k)}_i \left(\frac{D^{(k)}_{i+1} - D^{(k)}_i}{2E^{(k)}_i} \pm \sqrt{\left(\frac{D^{(k)}_{i+1} - D^{(k)}_i}{2E^{(k)}_i}\right)^2 + 1} \right) . \tag{4.62}$$

The shift is chosen by first finding the difference between the possible shifts and D_i, i.e.,

$$\delta_k^+ = |\sigma_k^+ - D_i| = \left| E_i \left(\frac{D_{i+1} - D_i}{2\,E_i} + \sqrt{\left(\frac{D_{i+1} - D_i}{2\,E_i}\right)^2 + 1} \right) \right| , \tag{4.63a}$$

$$\delta_k^- = |\sigma_k^- - D_i| = \left| E_i \left(\frac{D_{i+1} - D_i}{2\,E_i} - \sqrt{\left(\frac{D_{i+1} - D_i}{2\,E_i}\right)^2 + 1} \right) \right| . \tag{4.63b}$$

From here, the shift is chosen as

$$\sigma_k = \begin{cases} \sigma_k^+ & \text{for} \quad \delta_k^+ \leq \delta_k^- \\ \sigma_k^- & \text{for} \quad \delta_k^- < \delta_k^+ \end{cases} . \tag{4.64}$$

For larger matrices we observe that a better choice for the shift may be obtained by considering the 3×3 submatrix

$$\begin{pmatrix} D_i & E_i & 0 \\ E_i & D_{i+1} & E_{i+1} \\ 0 & E_{i+1} & D_{i+1} \end{pmatrix} . \tag{4.65}$$

It seems as thought this “cubic” shift over compensates, and the time spent finding the better shifts cancels out the increase in speed that results from finding them. The extension from the Wilkinson shift to using larger sub-matrices is also utilized in [76], however in this case a $(k \times k)$-bulge is created, and k shifts are needed to perform the calculation, as exemplified in (2.1) of [76]. Thus, the use of the eigenvalues of the trailing $(k \times k)$-matrix is a natural extension of their multi-shift program. Here, rather than using a larger “bulge”, we use a larger sub-matrix in order to obtain a (hopefully) better “guess” for the eigenvalue of the matrix toward which we are iterating. In extensive tests of the algorithm we found that different shifts seem to be optimal for different classes of input matrices. The "cubic" shift (based on the (3×3)–submatrix) seems to be better suited for ill conditioned matrices, while the Wilkinson shift performs better for banded matrices. For the case of well conditioned matrices the difference in performance between the two is negligible. Due to the convergence gained by using different shifts, the `FORTRAN` code included in appendix A includes a variable `SHIFTMODE`, with possible values 0 (no shift, $\sigma_k = 0$), 1 ($\sigma_k = D_i$), 2 (Wilkinson shift), and 3 (cubic shift). Generally the elimination of the shift (`SHIFTMODE`=0) is computationally disadvantageous.

We will never explicitly calculate the shifted matrix (but instead use it, as the name suggests) implicitly. This works by using the shift only to calculate the initial rotation, and then subtracting the shift out. We shift back, and the original eigenvalues are recovered. As such, the only actual calculation we have to do for the shift is for the C and S in the initial rotation of each iteration, as we will see in (4.66).

4.3.2. Procedure: Second Step of the `HTDQLS` Algorithm. The diagonalization of the symmetric tridiagonal matrix is done using a combination of initial rotations with an implicit shift and a series of Givens rotations (illistrated in figure 4.2). In the beginning we are trying to zero the super-/sub–diagonal element E_1, and one starts by calculating the shift σ_1, and constructing the initial rotation matrix $\underline{R}$ as follows,

$$\underline{R} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & C & S \\ & & & -S & C \end{pmatrix},$$

$$C = \frac{D_n - \sigma}{\sqrt{(D_n - \sigma)^2 + E_{n-1}^2}}, \qquad S = \frac{E_{n-1}}{\sqrt{(D_n - \sigma)^2 + E_{n-1}^2}}, \tag{4.66}$$

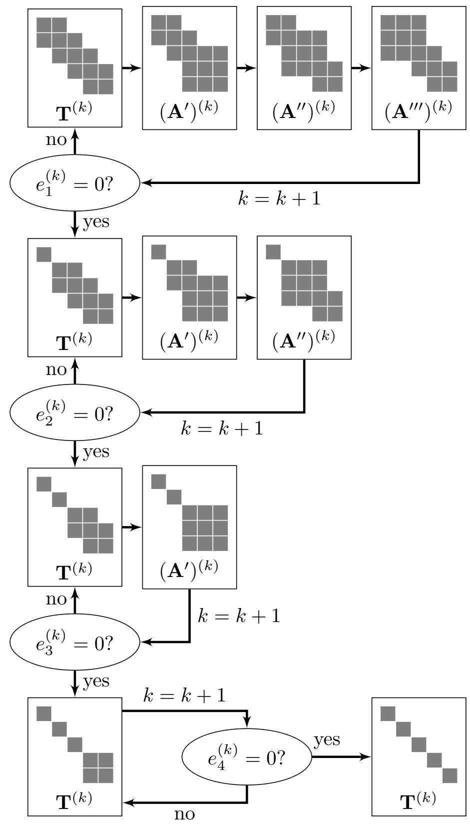


Figure 4.2: This flowchart illustrates the progression of the diagonalization step for a 5×5–matrix. For each step a "bulge" is introduced, and then chased out. Following this, the appropriate off diagonal elements are checked for convergence. Not include in this chart is the check for premature zeroes, and the applied solution.

where $C^2+S^2 = 1$ and $\underline{R}^T\ \underline{R} = \mathbb{1}_{n\times n}$. The initial step, which consists in the calculation of $\underline{T}' = \underline{R}^T\ \underline{T}\ \underline{R}$, creates an off tridiagonal element without eliminating any elements, constitutes the first step in the chasing the bulge program. We notice that the initial rotation matrix has a form similar to that of either a Jacobi or a Givens rotation. Jacobi and Givens rotations are defined on page 100 of reference [69] in terms of the elements they eliminate from a matrix; essentially, a Jacobi rotation eliminates the same matrix element that was used in the construction of the rotation matrix, whereas a Givens rotation eliminates a different element. The rotation $\underline{R}$ creates rather than eliminates a matrix element, creating a "bulge". This first rotation is neither a Jacobi nor a Givens rotation. Both here as well as in the `FORTRAN` implementation, we are using the notation that $\underline{D}$ is an n-dimensional vector containing the diagonal elements of our matrix and $\underline{E}$ is an $(n-1)$-dimensional vector (with $n-1$ elements) containing the sub-diagonal elements of our matrix. Then the first transformed matrix $\underline{T}' = \underline{R}^T\,\underline{T}\,\underline{R}$ is of the form

$$\underline{T}' = \begin{pmatrix} \ddots & \ddots & & & \\ \ddots & D'_{n-3} & E'_{n-3} & & \\ & E'_{n-3} & D'_{n-2} & E'_{n-2} & F' \\ & & E'_{n-2} & D'_{n-1} & E'_{n-1} \\ & & F' & E'_{n-1} & D'_n \end{pmatrix}, \tag{4.67}$$

with an obvious "bulge" as the off tridiagonal elements $T'_{n-2\,n}$ and $T'_{n\,n-2}$ are not equal to zero. The updated elements are

$$D'_n = C^2\,D_n + 2\,C\,S\,E_{n-1} + S^2\,D_{n-1}\,, \tag{4.68a}$$

$$E'_{n-1} = (C^2 - S^2)\,E_{n-1} + C\,S(D_{n-1} - D_n)\,, \tag{4.68b}$$

$$F' = T'_{nn-2} = T'_{n-2n} = S\,E_{n-2}\,, \tag{4.68c}$$

$$D'_{n-1} = C^2\,D_{n-1} + S^2\,D_n - 2C\,S\,E_{n-1}\,, \tag{4.68d}$$

$$E'_{n-2} = C\,E_{n-2}\,, \tag{4.68e}$$

while the remainder of the elements remain unchanged (the rest of the elements of $\underline{T}'$ are the same as in $\underline{T}$). To annihilate the off tridiagonal elements, we use a Givens rotation [77], which is constructed as

$$\underline{G}_{n-2} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & C & S & \\ & & & -S & C & \\ & & & & & 1 \end{pmatrix}, \tag{4.69}$$

$$C = \frac{E'_{n-1}}{\sqrt{E'^2_{n-1} + T'^2_{nn-2}}}, \qquad S = \frac{T'_{nn-2}}{\sqrt{E'^2_{n-1} + T'^2_{nn-2}}}. \tag{4.70}$$

Notice that $C^2 + S^2 = 1$, while for Hermitian matrices the condition would be $|C|^2 + |S|^2 = 1$. We have generalized the Givens rotations in order to preserve the complex symmetric structure of the matrix. Applying this rotation to our matrix will eliminate the "bulge," however a new "bulge" will be created one element up along the off tridiagonal ($T''_{n-1\,n-3} = T''_{n-3\,n-1} \neq 0$). The altered values of $\underline{T}'' = \underline{G}^T_{n-2}\underline{T}'\,\underline{G}_{n-2}$ are then

$$E''_{n-1} = C\,E'_{n-1} + S\,A'_{nn-2}\,, \tag{4.71a}$$

$$D''_{n-1} = C^2 D'_{n-1} + 2CSE'_{n-2} + S^2 D'_{n-2}\,, \tag{4.71b}$$

$$E''_{n-2} = (C^2 - S^2)E'_{n-2} + CS(D'_{n-2} - D'_{n-1})\,, \tag{4.71c}$$

$$F'' = T''_{n-1n-3} = T''_{n-3n-1} = SE'_{n-3}\,, \tag{4.71d}$$

$$D''_{n-2} = C^2\,D'_{n-2} + S^2\,D'_{n-1} - 2\,C\,SE'_{n-2}\,, \tag{4.71e}$$

$$E''_{n-3} = C\,E'_{n-3}\,. \tag{4.71f}$$

We can generalize these results, with the off tridiagonal element defined as F for each step, we find that in step i, where i runs from $n-1$ to the element which we are diagonalizing, the general Givens rotation are

$$C = \frac{E_{i+1}}{\sqrt{E_{i+1}^2 + F^2}}\,, \qquad S = \frac{F}{\sqrt{E_{i+1}^2 + F^2}}\,, \tag{4.72}$$

while the updated elements are

$$D_{i+1} = C^2 D_{i+1} + 2CSE_i + S^2 D_i\,, \qquad E_{i+1} = \sqrt{E_{i+1}^2 + F^2}\,, \tag{4.73a}$$

$$D_i = C^2 D_i - 2CSE_i + S^2 D_{i+1}\,, \tag{4.73b}$$

$$E_i = (C^2 - S^2)E_i + CS(D_i - D_{i+1})\,, \tag{4.73c}$$

$$E_{i-1} = CE_{i-1}\,, \qquad F = SE_{i-1}\,. \tag{4.73d}$$

For the initial rotation ($i = n - 1$), we note that $E_{i+1} = E_n$ is not really an element of the matrix, and we therefore set it equal to zero in the scheme defined in (4.73).

After the creation of the bulge, a total of $(n - 2)$ Givens rotations are required to eliminate the bulge, and return the matrix to tridiagonal form. Practically speaking, one does not have to recalculate the entire matrix in each step, but instead must only calculate S, C, and the 6 updated elements. After a sufficient number of iterations, the matrix will be diagonalized to machine precision. Extensive testing of the algorithm informs us that in typical cases, less than 30 iterations of this QL procedure are required to reach machine accuracy for a desired eigenvalue.

4.3.3. Deflation and Partitioning: Reducing the Matrix Size. In principle, one might think that the above procedure should constitute a generally applicable algorithm, which diagonalizes any general diagonalizable complex symmetric input matrix. However, a pitfall must be avoided. Namely, if one encounters a zero (to machine accuracy) in an off-diagonal element, within the second step described in chapters 4.3.1 and 4.3.2 then deflation becomes necessary. Put differently, when chasing the bulge as described in chapter 4.3.2, one strives to calculate the eigenvalues of the tridiagonal matrix $\underline{T}$ from the upper left to the lower right, i.e., one subsequently zeros (to machine accuracy) the elements E_i with i running from 1 to $n-1$. In the sense of equation (4.73), one iterates in ascending transformation orders k in order to zero the element E_i in the matrix $\underline{T}^{(k)}$. Let us assume that in this process an element E_j, with $j > i$, accidentally becomes equal to zero, within machine accuracy, before E_i is zeroed. This constitutes an early, or "premature," zero which requires special treatment. Namely, if we were to continue the the recursive algorithm of chapter 4.3.2 without any changes, then the bulge would always be annihilated prior to the point where it would affect E_i, due to the premature zero, in any subsequent iteration. In fact, reducing the effective size of the matrix is a known technique for speeding up algorithms [78].

In order to overcome the lock-up, we divide, or "partition" the matrix **T** into two smaller matrices,

$$\underline{T} = \begin{pmatrix} \underline{T}_1 & 0 \\ 0 & \underline{T}_2 \end{pmatrix}, \tag{4.74}$$

where $\underline{T}_1$ and $\underline{T}_2$ are tridiagonal matrices, with columns and rows running over the indices $i = 1, \ldots, j-1$ for $\underline{T}_1$ and $i = j, \ldots, n$ for $\underline{T}_2$. We assume that $\underline{Q}_1$ and $\underline{Q}_2$ diagonalize the matrices $\underline{T}_1$ and $\underline{T}_2$,

$$\underline{Q}_1^T \underline{T}_1 \, \underline{Q}_1 = \underline{D}_1 \,, \quad \underline{Q}_2^T \underline{T}_2 \, \underline{Q}_2 = \underline{D}_2 \,, \tag{4.75}$$

where $\underline{Q}_1$ and $\underline{Q}_2$ are the similarity transforms and $\underline{D}_1$ and $\underline{D}_2$ are the corresponding diagonal matrices of $\underline{T}_1$ and $\underline{T}_2$, respectively. We can then almost trivially construct the orthogonal transformation

$$\underline{Q} = \begin{pmatrix} \underline{Q}_1 & 0 \\ 0 & \underline{Q}_2 \end{pmatrix}, \tag{4.76}$$

for which

$$\underline{Q}^T \underline{T}\,\underline{Q} = \begin{pmatrix} \underline{Q}_1^T & 0 \\ 0 & \underline{Q}_2^T \end{pmatrix} \begin{pmatrix} \underline{T}_1 & 0 \\ 0 & \underline{T}_2 \end{pmatrix} \begin{pmatrix} \underline{Q}_1 & 0 \\ 0 & \underline{Q}_2 \end{pmatrix} = \begin{pmatrix} \underline{D}_1 & 0 \\ 0 & \underline{D}_2 \end{pmatrix} . \tag{4.77}$$

The tridiagonal matrices $\underline{T}_1$ and $\underline{T}_2$ are smaller in size than $\underline{T}$. One needs to invoke the iterated, implicitly shifted QL decomposition on both of them, individually. As such we have "deflated" the matrix $\underline{T}$ into two smaller matrices. Quite surprisingly, this problem is rather scarcely treated in the literature. It is discussed very briefly in section 7.11 of reference [69]. There are further unpublished notes that address the issue, and the solution is referred to as "deflation" in section 11.4 of reference [79] and near the end of section 3.6.2 of reference [80]. In section 4.7 of [81], the same procedure is called "partitioning".

4.4. COMPLEMENTARY QR ALGORITHM

We now discuss an alternative formulation for the second step of the `HTDQLS` algorithm, in which we implement an iterative QR procedure rather then the QL decompositions. Where the QL decomposition an input matrix into an orthogonal matrix $\underline{Q}$ ($\underline{Q}^T\,\underline{Q} = \mathbb{1}$) and a left triangular matrix $\underline{L}$, the complimentary QR algorithm uses the same tools to decompose the input matrix into an orthogonal matrix $\underline{Q}$ and a right triangular matrix $\underline{R}$. When applied in the same manner as the QL decompositions are in the second step of the algorithm, the QR procedure generates a "bulge" which is chased from the top left corner, out through the bottom right corner (while the QL implementation chases the "bulge" from the bottom right up to the top left). By definition, QL decomposition is given as

$$\underline{A} = \underline{Q}\,\underline{L}\,, \quad \underline{A}' = \underline{L}\,\underline{Q}\,, \tag{4.78}$$

where $\underline{L}$ is a left triangular matrix (i.e., only the elements below and including the diagonal elements are non-zero). By contrast, QR decomposition is given as

$$\underline{A} = \underline{Q}\,\underline{R}\,, \quad \underline{A}' = \underline{A}\,\underline{Q}\,, \tag{4.79}$$

where $\underline{R}$ is a right triangular matrix (i.e., only the elements above and including the diagonal elements are non-zero).

While it has not been explicitly shown, the diagonalization step in chapter 4.3 is an application of QL decomposition. Once we have a tridiagonal matrix, we have a choice to either use a QL or a QR decomposition. It should be noted that the way in which the matrix is tridiagonalized also speaks to which decomposition we are going to use, in this case the matrix was tridiagonalized from the bottom right corner up to the upper left. This implicitly tells us that the next step should be diagonalization based on QL decompositions. If instead we wanted to use a QR procedure we should have tridiagonalized the matrix (still using Householder reflections) from the upper left down to the lower right. Regardless, the option of which decomposition to use for the diagonalization step is still present. We chose to use a QL decomposition, and this application manifestes in us chasing the bulge from the lower right out through the upper left. Had a QR decomposition been used, then the process would still involve chasing the bulge, however it would have originated in the upper left, and been chased out through the lower right. Since these techniques are so similar, and work in much the same way, it is possible to rewrite the algorithm, following the same theory, so that a QR decomposition is implemented. The method of tridiagonalization is quite trivial, however there are a few changes to the diagonalization step.

As with the QL implementation, we use an implicit shift when implementing the QR version. Here, instead of considering the $(m \times m)$–submatrix $(m = 0, 1, 2, 3,$ depending on `SHIFTMODE`), in the top left corner, we use the $(m \times m)$–submatrix in the bottom right corner. This is a manifestation of the fact that the QR algorithm works to zero the super-/sub–diagonal elements E_i starting with E_{n-1}, and working up to E_1. As with the QL case, the shift closest to the element we wish to converge is chosen. We then proceed to create the bulge using a the initial rotation, constructed as

$$\underline{R} = \begin{pmatrix} C & S & & & \\ -S & C & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}, \tag{4.80}$$

$$C = \frac{D_1 - \sigma_n}{\sqrt{(D_1 - \sigma_n)^2 + E_1^2}}, \qquad S = -\frac{E_1}{\sqrt{(D_1 - \sigma_n)^2 + E_1^2}}. \tag{4.81}$$

The updated elements of $\underline{T}' = \underline{R}^T \, \underline{T} \, \underline{R}$ are then

$$D'_1 = C^2 D_1 - 2\,C\,S\,E_1 + S^2 D_2\,, \qquad D'_2 = C^2 D_2 + 2C\,S\,E_1 + S^2 D_1\,, \tag{4.82a}$$

$$E'_1 = (C^2 - S^2)\,E_1 + C\,S\,(D_1 - D_2)\,, \qquad E'_2 = C\,E_2\,, \tag{4.82b}$$

$$F' = T'_{1\,3} = T'_{3\,1} = -S\,E_2\,. \tag{4.82c}$$

This is then followed by $(n-2)$ generalized Givens rotations, given by $\underline{G}_i$ where i runs from 2 up to $(n-1)$, and the corresponding updated elements are given as

$$C = \frac{E_i}{\sqrt{E_i^2 + F'^2}}, \qquad S = -\frac{F'}{\sqrt{E_i^2 + F'^2}}, \tag{4.83a}$$

$$D'_i = C^2 D_i - 2\,C\,S\,E_i + S^2 D_{i+1}\,, \qquad D'_{i+1} = C^2 D_{i+1} + 2\,C\,S\,E_i + S^2 D_i\,, \tag{4.83b}$$

$$E'_{i-1} = \sqrt{E_{i-1}^2 + F'^2}\,, \qquad E'_i = (C^2 - S^2)\,E_i + C\,S\,(D_i - D_{i+1})\,, \tag{4.83c}$$

$$E'_{i+1} = C\,E_{i+1}\,, \qquad F'' = -S\,E_{i-1}\,. \tag{4.83d}$$

The process is then iterated until convergence is achieved, much like we did for the QL decomposition. The difference being that the order in which the eigenvalues converge is reversed. A `FORTRAN` implementation of the QR version of the algorithm is provided in appendix A.2.

4.5. NUMERICAL REFERENCE DATA

Complex symmetric matrices arise naturally in physics. These occurrences include, but are not limited to, the projection of a $\mathcal{PT}$-symmetric Hamiltonian onto an appropriate set of basis states as well as complex scaled Hermitian Hamiltonians. Let us examine both the real and imaginary cubic perturbations to the harmonic oscillator,

$$H_3 = \frac{1}{2}\,p^2 + \frac{1}{2}\,x^2 + \mathrm{i}\,G\,x^3\,, \tag{4.84}$$
$$h_3 = \frac{1}{2}\,p^2 + \frac{1}{2}\,x^2 + g\,x^3\,, \quad x \to x\,\mathrm{e}^{\mathrm{i}\,\theta}\,, \quad 0 < \theta < \frac{\pi}{5}\,, \tag{4.85}$$

as well as both the real and imaginary quintic perturbations,

$$H_5 = \frac{1}{2}\,p^2 + \frac{1}{2}\,x^2 + \mathrm{i}\,G\,x^5\,, \tag{4.86}$$
$$h_5 = \frac{1}{2}\,p^2 + \frac{1}{2}\,x^2 + g\,x^5\,, \quad x \to x\,\mathrm{e}^{\mathrm{i}\,\theta}\,, \quad 0 < \theta < \frac{\pi}{7}\,, \tag{4.87}$$

where $\hbar = 1$. These Hamiltonians have been extensively studied [2,5,8,12,53,56,57], and as such they are ideally suited as a testbed to generate the reference data found in tables 4.1 and 4.2. The algorithm is best suited for fully populated complex symmetric matrices, however, when using the harmonic oscillator wave–functions as our basis states, these matrices will be sparsely populated. The algorithm is, of course, still able to diagonalize these matrices, but may not be the most efficient method. Algorithms that take advantage of the band structure of these matrices

Table 4.1: Example ground and first excited state energies for H_3 and h_3, as defined in equations (4.84) and (4.85) respectively, with different values of G and g.

G	$E_0^{(3)}(G)$
0.8	0.740948971482359671409952387680562989649218672786322029506972
1.0	0.797342607508906189039080960791013163097244534480331157578578
1.2	0.849097066890258015437917408284257146062837108501875662860849

G	$E_1^{(3)}(G)$
0.8	2.559093658684295834337630756394908959050331276996157589230549
1.0	2.773524985195379715405817000015530142310848902829685205722959
1.2	2.967273593442652066085730346704529728796925557957963243225577

g	$\epsilon_0^{(3)}(g)$
0.8	0.561066208979404775116928166431422687738464300780198739535914
	-0.358599844691200673512575409270525983934995081882520287419416 i
1.0	0.612888433307754624258817501988651413733339788307182942066181
	-0.408592666932267283159498868767160516270974834438403999097532 i
1.2	0.659471416719299127897719134154497156025222371045692402135959
	-0.450150034262365046307565768244376605581927184833450590658757 i

g	$\epsilon_1^{(3)}(g)$
0.8	1.991456698898661194884384549965325120089956168185455548446729
	-1.369705736282645527841528155126063494834869438727230322025554 i
1.0	2.180413837536348771230161963541741131247172136835058974459041
	-1.526207655693032510006853946967495624445906099848804410355220 i
1.2	2.347898333307082484602271828699097353118353897863695684660492
	-1.659906360584923744548090528514636695168085721095250043307566 i

tend to be more efficient. Fortunately the point of this is to provide reference data generated by the `HTDQLS` algorithm, and as such the emphasis should be placed on obtaining data which can be collaborated, rather then choosing matrices for which this algorithm is particularly suited.

The eigenvalues, $E_i^{(3)}$, of H_3 will be functions of G (with $E_i^{(3)} \equiv E_i^{(3)}(G)$) while the eigenvalues, $\epsilon_i^{(3)}$, of h_3 will be eigenvalues of the coupling g (with $\epsilon_i^{(3)} \equiv \epsilon_i^{(3)}(G)$). Similarly the eigenvalues of H_5 and h_5 will be functions of G and g, and as such

Table 4.2: Example ground and first excited state energies for H_5 and h_5, as defined in equations (4.86) and (4.87) respectively, with different values of G and g.

G	$E_0^{(5)}(G)$
0.8	0.753891246215897860993377553234208475583085701533373
1.0	0.79140175777864076155860859702017488502261796732061
1.2	0.82455852408502236750428497431837653777269857364975

G	$E_1^{(5)}(G)$
0.8	2.72584719948185664152979718639908174280319603142609
1.0	2.87580993584575639268780020627266878199726701715050
1.2	3.00702088732252853807330439861318505358286029883287

g	$\epsilon_0^{(5)}(g)$
0.8	0.67450329439700818291896353707548915544787045415294
	-0.24497261241861497579575125097610853374744892054504 i
1.0	0.70875699952222315206082074451175354669392841899262
	-0.26762943090641375087559443630536242086978935856853 i
1.2	0.73900997993150365467103706748425891486602648607993
	-0.28682232235861442117427312071149714709711905229129 i

g	$\epsilon_1^{(5)}(g)$
0.8	2.45493787189842310938036792862778581446876990238564
	-0.98294115378930202944681413329505028737592346575922 i
1.0	2.59036713329602333960788878304385306013391798261856
	-1.06050350702232561922594400015297159643910084409003 i
1.2	2.70878087811349467420450292754718908864089691541160
	-1.12701898384173378366059369869202592386004389371268 i

we denote them as $E_i^{(5)}(G)$ and $\epsilon_i^{(5)}(g)$, respectively. By projecting the Hamiltonians onto the first few thousand eigenstates of the harmonic oscillator, and using a multi-precision implementation [82–85] of the algorithm, we obtain the lowest two eigenvalues of the Hamiltonians for $G = g = 0.8, 1.0, 1.2$. These values are given in tables 4.1 and 4.2. Every digit given is significant and the accuracy is estimated based on the apparent convergence of the numerical data as the size of the matrix is increased.

We can exploit the fact that the `HTDQLS` algorithm is best suited for densely populated matrices by using a non–orthogonal basis. We choose a non–orthogonal basis spanned by the functions

$$\psi_m(x) = \exp(-a\, m\, x^2)\,, \quad m = 1, \ldots, \frac{n}{2}\,, \tag{4.88a}$$

$$\psi_{m'}(x) = x\, \exp(-a\, m'\, x^2)\,, \quad m' = \frac{n}{2} + 1, \ldots, n\,, \tag{4.88b}$$

where n is the (even integer) total number of basis functions and m, m' serve as counters. This defines basis functions $\psi_m(x)$ with $m = 1, \ldots, n$ which have even parity for $1 \leq m \leq n/2$ and odd parity for $n/2 < m \leq n$. We also note that a is a real, positive number. We then find that the $(n \times n)$-overlap matrix $\underline{S}$ and the Hamiltonian matrix $\underline{H}$, have the elements

$$S_{ij} = \int_{-\infty}^{\infty} dx\, \psi_i(x)\, \psi_j(x) = \langle \psi_i | \psi_j \rangle_*\,, \tag{4.89}$$

$$H_{ij} = \int_{-\infty}^{\infty} dx\, \psi_i(x)\, H_3\, \psi_j(x) = \langle \psi_i | H | \psi_j \rangle_*\,, \tag{4.90}$$

where H is a $\mathcal{PT}$–symmetric Hamiltonian, and the inner product is denoted as in equation (3.14). On the basis of the HTDQLS algorithm, we first calculate the square root of the overlap matrix,

$$\underline{S} = \underline{Q}\,\underline{D}\,\underline{Q}^{\mathrm{T}}\,. \qquad \underline{M} = \underline{Q}\,\sqrt{\underline{D}}\,\underline{Q}^{\mathrm{T}}\,, \qquad \underline{S} = \underline{M}^2\,. \tag{4.91}$$

The square root of the diagonal matrix $\underline{D}$ is easily calculated. We now make the following ansatz for an eigenvector, expressed in the non-orthogonal basis,

$$|\psi\rangle = \sum_j c_j\, |\psi_j\rangle\,. \tag{4.92}$$

The eigenvalue problem within the basis, $\sum_j c_j \, H \, |\psi_j\rangle = E \sum_j c_j \, |\psi_j\rangle$, can then be formulated as

$$\sum_j \langle \psi_i | H | \psi_j \rangle \, c_j = \sum_j E \, \langle \psi_i | \psi_j \rangle \, c_j \, . \tag{4.93}$$

equivalently, with the coefficient vector $\underline{c}$,

$$\underline{H} \, \underline{c} = E \, \underline{S} \, \underline{c} \, . \tag{4.94}$$

We then define

$$\underline{d} = \underline{M} \, \underline{c} \, , \tag{4.95}$$

yielding

$$\underline{M}^{-1} \, \underline{H} \, \underline{M}^{-1} \, \underline{d} = E \, \underline{d} \, . \tag{4.96}$$

A diagonalization of the effective Hamiltonian matrix

$$\underline{H}_{\text{eff}} = \underline{M}^{-1} \, \underline{H} \, \underline{M}^{-1} \tag{4.97}$$

then leads to the approximate energies of the input Hamiltonian H. Implementing this technique, as well as extended arithmetic precision (Bailey's `MPFUN` [82–85]), on the imaginary cubic anharmonic oscillator (4.84), we find the ground state energy to be

$$\begin{aligned} E_0^{(3)}(G = 0.8) = 0.\; & 74094\,89714\,82359\,67140\,99523\,87680\,56298\,96492\,18672\,78632 \\ & 20295\,06972\,65779\,86489\,95262\,29285\,78562\,62734\,77203\,42411 \, . \end{aligned} \tag{4.98}$$

This 100-decimal reference value was calculate uses a matrix representation of H_3 of relatively modest size (700×700).

5. A FORTRAN IMPLEMENTATION OF THE ALGORITHM

5.1. GENERAL REMARKS

While we have presented the underlying theory of the HTDQLS algorithm in the previous chapter, and further reading will detail an explicit FORTRAN implementation (see appendix A), one may ask if a new matrix diagonalization routine is strictly necessary. For example, the ZGEEVX algorithm in LAPACK [61] utilizes the tools provided by LAPACK to diagonalize complex (not necessarily symmetric) matrices. Furthermore, ZGEEVX is a very robust algorithm, and as such is rather complicated. This level of complexity makes ZGEEVX a "black box," by which we mean that it is not easily scalable in terms of numerical precision. Our algorithm on the other hand is relatively simple due to its narrow scope, and is easily scalable in terms of numerical precision. In fact, the explicit FORTRAN implementation included in appendix A is written using COMPLEX*32 precision, which already exceeds the accuracy utilized in LAPACK. As seen in chapter 4.5 the precision can be increased further by utilizing a multi–precision package, such as Bailey's MPFUN90 [82–85]. We further find, that for typical applications (matrices around rank 500), that the HTDQLS algorithm tends to be faster than the publicly accessible ZGEEVX algorithm. With these considerations in mind, especially the ease of scalability, it becomes clear why there is a need for the HTDQLS algorithm.

5.2. SPECIFIC ALGORITHMS

Here we discuss the implementation of the algorithm as discussed in chapter 4, while an explicit FORTRAN implementation can be found in appendix A. All the steps are presented, save for the initial setup. It is up to the user to create a shell for the program, that in some way defines the matrix to be diagonalized. As described in chapter 4, the algorithm first employs generalized Householder reflections to tridiagonalize the matrix, and then employs a generalization of the QL algorithm with an implicit shift. As such the routine is called `HTDQLS`, which stands for Householder–based Tridiagonalization followed by generalized QL decompositions with an implicit shift.

The algorithm is implemented using separate subroutines for the tridiagonalization and diagonalization steps, along with several other supporting subroutines and a master subroutine. There are two versions of three of these routines, used to either calculate solely the eigenvalues or to calculate both the eigenvalues and eigenvectors of a given input matrix. These routines are denoted by either a "`1`" (for the eigenvalue implementation) or a "`2`" (for the eigenvalue and eigenvector implementation) at the end of the subroutine's name. The master subroutine directs the flow of the algorithm so that the desired option is implemented. Here we briefly describe all the subroutines.

The master subroutine `HTDQLS(JOBZ, N, A, D, Z, SORTFLAG, SHIFTMODE)` is used to call the other subroutines, as well as determine the order in which they are used. If `JOBZ='N'`, then only the eigenvalues are calculated, while if `JOBZ='V'` then the eigenvalues and eigenvectors are calculated. The rank of the input matrix and the input matrix itself are denoted by `N` and `A` respectively. Upon the completion of the program, the eigenvalues are stored in '`D`' and if the eigenvectors were calculated they are stored in '`Z`', where the ith column of `A` is the eigenvector of the input matrix corresponding to the ith eigenvalue stored in `D(i)`. Upon completion of the routine, `A` retains its original values, and may be used to check the results, which is straightforward when the eigenvectors are calculated. The boolean variable `SORTFLAG` determines if the eigenvalues (and corresponding eigenvectors) are sorted according to the real part of the eigenvalues, and the integer `SHIFTMODE` can be set to 0, 1, 2, or 3 depending on whether zero–shift, linear, quadratic, or cubic (respectively) mode is desired.

The routines `HTD1(N, A, D, E)` (similarity transforms are not stored) and `HTD2(N, A, D, E)` (similarity transforms are stored) implement the tridiagonalization step of the program. Each routine takes the input matrix `A` of rank `N` and tridiagonalizes it as prescribed in chapter 4.2 The tridiagonal matrix is then stored in `D` and `E`. `HTD2` stores the similarity transform in `A`.

The routines `QLS1(N, D, E, SHIFTMODE)` (does not store the similarity transforms) and `QLS2(N, D, E, Z,SHIFTMODE)` (does store the similarity transforms in `Z`) diagonalize the tridiagonal input matrix stored in the vectors `D` and `E`. The calculated eigenvalues are stored in `D`. When a premature zero occurs, the routines perform the deflation step automatically. `QLS2` stores the similarity transforms in `Z`.

The routine `SHIFT(N, K, V, D, E, S, SHIFTMODE)` calculates the implicit shift based on the values of `SHIFTMODE`, as prescribed in section 4.3.1. The input tridiagonal matrix of rank `N` is stored in `D` and `E`, while `K` and `V` are used to determine the elements used to calculate the possible shifts. After the calculations are complete, `SHIFT` chooses the shift whose value is closest to that of the diagonal element which we wish to converge. The shift is output on `S`, and returned to the appropriate version of `QLS`. Finally, the routines `SORT1(N, D)` and `SORT2(N, D, A)` sort the `N` eigenvalues stored in `D` into ascending order of the real part. `SORT2` additionally sorts the eigenvectors to match the position of the associated eigenvalue.

5.3. COMPUTATIONAL PERFORMANCE OF THE ALGORITHM

5.3.1. Numerical Accuracy. In order to gauge the numerical accuracy of the `HTDQLS` algorithm, we turn to a complex rotated version of the harmonic oscillator

Hamiltonian,

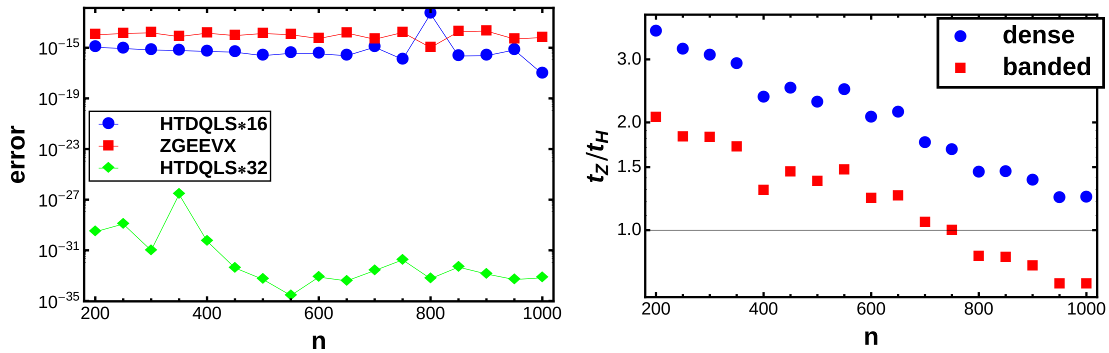
$$H_0 = -\frac{1}{2}\,\partial_x^2 + \frac{1}{2}\,x^2\,, \quad x \to x\,\mathrm{e}^{\mathrm{i}\,\theta}\,, \quad \partial_x \to \partial_x\,\mathrm{e}^{-\mathrm{i}\,\theta}\,, \quad \theta = \frac{\pi}{16}\,. \tag{5.1}$$

The ground-state eigenvalue of the harmonic oscillator is unaffected by the complex scaling and reads as $\lambda_0 = \frac{1}{2}$. On the other hand, using a projection of the complex rotated H_0 onto a suitable basis, we can generate complex symmetric matrices in which at least the first eigenvalue is known, namely, λ_0. A measure of the numerical accuracy of the method is given as follows,

$$\mathrm{err} = \frac{|D_1 - \lambda_0|}{\lambda_0}\,, \tag{5.2}$$

where err is the numerical error, and D_1 is the ground-state eigenvalue as found by the corresponding algorithm. The goal is to compare `COMPLEX*16` versions (roughly 16 significant decimals) of `HTDQLS` to `ZGEEVX`, which is a `LAPACK` routine [61] that diagonalizes complex matrices. (The latter does not specialize in complex *symmetric* matrices but is a more general solver.) Aside from a single outlier at $n = 800$, we found that the `HTDQLS` algorithm is generally an order of magnitude more accurate than the `LAPACK` routine `ZGEEVX` (see figure 5.1). In typical cases, we find that the final numerical loss of our method in reproducing known eigenvalues of Hamiltonians does not exceed 4–5 decimals, consistent with the outlier in figure 5.1. For comparison, we also plot in figure 5.1 the numerical accuracy obtained using a `COMPLEX*32` version of `HTDQLS`; such a high-precision version is not available for `ZGEEVX`.

5.3.2. Speed. In order to test the computational efficiency of `HTDQLS`, we again compare the `LAPACK` routine `ZGEEVX` with a `COMPLEX*16` version of `HTDQLS`. This is done with the help of two types of matrices, the first being composed of random complex numbers, leading to densely populated, complex symmetric matrices, while the second type of matrices are generated using the harmonic oscillator Hamiltonian



(a) Numerical Accuracy

(b) Speed

Figure 5.1: In comparing the `HTDQLS` algorithm with the `LAPACK` routine `ZGEEVX`, the relative numerical accuracy of the ground state energy of the complex rotated harmonic oscillator H_0 given in equation (5.1) is plotted as a function of the size of the matrices (see figure (a)). In figure (b), the average ratio of the runtimes, t_Z/t_H (where t_Z is the runtime of `ZGEEVX` and t_H is the runtime of `HTDQLS`), is plotted against the rank of the matrices. Two types of matrices were used, densely populated and banded ones. Further details are in the text.

with an imaginary cubic perturbation (see equation (3.1)), with random values of G, resulting in banded, complex symmetric matrices. We then average 150 trials for each rank (200 to 1000) and find the ratio of the run times (see figure 5.1b). For smaller matrices we found that `HTDQLS` runs quite a bit faster, but as the size of the matrices increases `ZGEEVX`'s performance improves. By rank 750 `ZGEEVX` performs faster (albeit slightly) than `HTDQLS` for the banded matrices. For the densely populated matrices on the other hand, `HTDQLS` is faster for all the matrices we tested.

6. (PARTIAL) CONCLUSIONS

In this part of the dissertation we considered Hermitian, pseudo–Hermitian, and $\mathcal{PT}$–symmetric Hamiltonians. From the definition of pseudo–Hermiticity, it is immediately clear that the set of Hermitian operators is a subset of pseudo–Hermitian operators. There is certainly some overlap between $\mathcal{PT}$–symmetry and Hermiticity. The harmonic oscillator, for example, is both $\mathcal{PT}$–symmetric and Hermitian. Neither the set of $\mathcal{PT}$–symmetric operators, nor the set of Hermitian operators, is a subset of the other. This is clearly demonstrated by considering two example cases; the imaginary cubic anharmonic oscillator (3.1) is $\mathcal{PT}$–symmetric, but not Hermitian, while the real cubic anharmonic oscillator (2.39) is Hermitian, but not $\mathcal{PT}$–symmetric. We are left to consider the relationship between $\mathcal{PT}$–symmetry and pseudo–Hermiticity. If we consider the superluminal Dirac–Hamiltonian $\vec{\alpha}\cdot\vec{p}+\beta\,\gamma^5\,m$ [25], we find that it is not $\mathcal{PT}$–symmetric (see chapter 14), but it has been identified as γ^5–Hermitian [43]. We can then conclude that the set of pseudo–Hermitian operators are not a subset of $\mathcal{PT}$–symmetric operators. In chapter 2.2 we found that when considering Hamiltonians of the form $H = \vec{p}^{\,2}/(2m) + V$, one finds that if H is $\mathcal{PT}$–symmetric then H is $\mathcal{P}$–Hermitian. The obvious identification would then be that all $\mathcal{PT}$–symmetric operators are $\mathcal{P}$–Hermitian. This "obvious identification" turns out to be incorrect. First let us consider the somewhat trivial example of the x momentum operator/Hamiltonian, $H_p = p = -\mathrm{i}\,\partial_x$, which we can easily show is $\mathcal{PT}$–symmetric ($\mathcal{PT}\,H_p\,\mathcal{TP} = -\mathcal{P}\,p\,\mathcal{P} = p = H_p$), but not $\mathcal{P}$–Hermitian ($\mathcal{P}^{-1}\,H_p^{+}\,\mathcal{P} = \mathcal{P}\,p\,\mathcal{P} = -p \neq H_p$). However the momentum operator/Hamiltonian is pseudo–Hermitian, as the momentum operator is known to be Hermitian (1–Hermitian). We can also consider the cases of $H_{\mathrm{A}} = p + x^2 + \mathrm{i}\,x^3$ and $H_{\mathrm{B}} = p + \mathrm{i}\,x^3$, for which it is trivial to show that both are $\mathcal{PT}$–symmetric, but not $\mathcal{P}$–Hermitian. However, they are both trivially $\mathcal{M}$–Hermitian, where $\mathcal{M}$ is the mirror operator about the p axis in phase space ($x \to -x$, $p \to p$). When working with typical Hamiltonians ($H = \vec{p}^2/2m + V$), which much of the literature focuses on [2–14], and as we focused on in this part of the dissertation, one can make the obvious identification that $\mathcal{PT}$–symmetric Hamiltonians are $\mathcal{P}$–Hermitian, but one must be careful as this is not always true for more general $\mathcal{PT}$–symmetric Hamiltonians. In any case, it is clear that the concepts of pseudo–Hermiticity and $\mathcal{PT}$–symmetry are related, and constitute viable alternatives to Hermiticity.

Like Hermitian Hamiltonians, exactly $\mathcal{PT}$–symmetric Hamiltonians have a real spectra. It is this shared property that has lead to the development of Hermitizing transforms, which transforms a $\mathcal{PT}$–symmetric Hamiltonian into a Hermitian Hamiltonian utilizing a similarity transform, which is by necessity non–unitary. The existence of such a procedure has lead to the conclusion that $\mathcal{PT}$–symmetry is equivalent to Hermiticity [19]. These transforms are perturbative by nature, and generally lead to a much more complicated, potentially non–local, Hermitian Hamiltonian [15, 19]. Due to the fact that the Hermitizing transform is not unitary, the relation between vector–spaces are not conserved [15, 21]. Additionally, we note that these transforms do not conserve parity. Furthermore, by considering the wave–functions of the of a $\mathcal{PT}$–symmetric Hamiltonian, we find a number of inconsistencies when compared to the characteristics of Hermitian wave–functions. Finally, Hermitian Hamiltonians describe closed systems, while $\mathcal{PT}$–symmetric Hamiltonians are special cases of open systems, in which the gain and loss terms are in equilibrium. Under these considerations, one is left to conclude that while Hermitizing transforms do conserve the spectrum of the initial Hamiltonian, $\mathcal{PT}$–symmetry and Hermiticity are not equivalent.

The wave–functions of Hermitian Hamiltonians have a number of nice characteristics, including nodes which can be used to enumerate the wave–functions, as well as the fact that they are governed by the concavity condition. $\mathcal{PT}$–symmetric wave–functions do not share either of these characteristics. In fact, they do not have any complex zeroes, but instead have an infinite number of both real and imaginary zeroes. As for the concavity condition, $\mathcal{PT}$–symmetric wave–functions have complex potentials, and as such no version of the concavity condition can be claimed.

Despite the obvious differences, the $\mathcal{PT}$–symmetric wave–functions are not as "alien" as one might initially suspect. The modulus of the potential does confine the wave–functions to the "classically allowed region," much like the Hermitian counterpart. Furthermore, where we would expect to find nodes in the Hermitian case, we find local minima in the $\mathcal{PT}$–symmetric case, thus providing a potential solution to the question of how to enumerate the wave–functions.

Finally, we used the inspiration given to us from the study of the pseudo–Hermitian Hamiltonians, in order to delineate a matrix diagonalization algorithm, specifically designed for complex symmetric matrices. The key observation is that, after a suitable projection of the pseudo–Hermitian Hamiltonian onto a finite basis of Hilbert space vectors, the Householder reflections can be generalized to effectively tridiagonalize a complex symmetric matrix provided that the inner product is replace with the indefinite inner product. It then takes only $(n-2)$ iterations to transform a fully populated complex symmetric matrix into a symmetric tridiagonal matrix. Once the tridiagonal form is obtained, one utilizes the obvious generalization of QL decompositions. The resulting algorithm was used to great effect in determining the wave–functions used to fuel our earlier discussion, as well as obtaining high precision eigenvalues.

Part II

Dirac Hamiltonians and Foldy–Wouthuysen Transforms

7. INTRODUCTION

In this part we try to answer the question: How can the eigenvalues and corresponding wave–functions of relativistic Dirac Hamiltonians, including relevant degrees of freedom, be rotated onto a decoupled basis, in appropriate limits? To answer this question we investigate a number of generalized Dirac Hamiltonians, and employ both the classic Foldy–Wouthuysen transformation [22], as well as the "chiral" variant [24].

Generalized Dirac equations are used to describe quantum particles moving at relativistic speeds as they interact with different potentials [23, 60, 86] (also see appendix C). To answer the central question of this part, we must first and foremost understand how to derive the Hamiltonians for our example cases. In some cases the derivation of these Hamiltonians can be rather straightforward, as the correspondence principle can be applied. Other cases are not so simple, as the use of the correspondence principle neglects to take into account the curvature of space-time, which comes about when considering, for example, gravitational potentials, and particles in a non–inertial reference frame. These interactions must be covariantly coupled to the Dirac equation, resulting in more complicated Hamiltonians.

Once the generalized Dirac Hamiltonians are obtained, we find that the equations for the particles and antiparticles are entangled, making it difficult to interpret how the potentials affect the particles and antiparticles. Traditionally a Foldy–Wouthuysen transformation is used to disentangle the two equations, making it significantly easier to understand the Dirac Hamiltonians [22, 23, 87–89]. Unfortunately, the Foldy–Wouthuysen transformation can only be applied exactly to the simplest of such equations, the free particle. To perform the transformation on anything more complicated, one must approximate to the non–relativistic limit. This gives rise to a well defined iterative procedure, in which one approximates the Dirac Hamiltonian up to a desired order. The resulting transformed Hamiltonian then reveals itself in the familiar form of a Schrödinger equation, with relativistic corrections. Thus the general Foldy–Wouthuysen transformation both decouples the particles and antiparticles, as well as transforming the Hamiltonians into an easily understandable form. This is in contrast to the exact transformation, which when applied to the free particle, still leaves the need for a Taylor series expansion to enter the nonrelativistic limit.

While the Foldy–Wouthuysen transformation is well defined in terms of its procedure, it can become a rather complicated computation, especially when higher–order terms are desired. This, coupled with the inexact nature of the transform has lead to attempts aimed at finding either an easier method, or a method which yields an exact result. We will be looking closely at an example of the former. In what we call the "chiral Foldy–Wouthuysen" transformation, a rather good, if ultimately unsuccessful attempt is made to simplify the procedure. While it can be used to find the non-relativistic limit of the free Dirac Hamiltonian, it is shown to be unsuccessful when applied to a variety of generalized Hamiltonians. Still, it is instructive to examine the transform in detail, both to appreciate the algebraic properties, as well as to understand the possible pitfalls of the approach.

In the chapters that follow we will be examining a number of generalized Dirac equations, and performing both the standard, as well as the chiral Foldy–Wouthuysen transformations. While most of the Hamiltonians can be obtained using the correspondence principle, some of them cannot, thus in chapter 8 we shall derive these Hamiltonians. In chapter 9 we perform the standard transformation on these Hamiltonians, including the exact and general transformation of a free particle, as well as the textbook example of the Dirac–Coulomb Hamiltonian, which will serve to ground the discussion of the transformed Dirac equation coupled to a gravitational field. In chapter 10 we perform the chiral transform on the same Hamiltonians, and compare the results. Finally, concluding remarks are in chapter 11.

Throughout part II of this thesis, we will be using units such that $\hbar = c = \epsilon_0 = 1$.

8. DIRAC EQUATION IN CURVED–SPACETIME

8.1. SOME BASICS

It is helpful to first clarify our conventions for the indices used throughout this work, as they pertain to both Lorentz as well as spatial components of the vectors. Namely, we shall be using lowercase Greek characters for the curved–spacetime ($\mu, \nu, ... = 0, 1, 2, 3$), lower case Latin characters starting at i for curved–space ($i, j, k, ... = 1, 2, 3$), capital Latin characters for flat–spacetime, i.e., the anholonomic basis ($A, B, C, ... = 0, 1, 2, 3$), and capital Latin characters starting at I for anholonomic space ($I, J, K... = 1, 2, 3$). Additionally, the symbol η will be used for the Minkowski metric, $[\eta_{AB}] = \mathrm{diag}[1, -1, -1, -1]$, and g for the curved space metric, $g_{\mu\nu}(x)$. Finally, we shall use $\overline{\gamma}$ and $\widetilde{\gamma}$ for the curved– and flat–spacetime Dirac γ matrices, inspired by the conventions used in [90]. However, as we shall see below, sometimes, the contraction of indices with Kronecker symbols will induce the necessity to intertwine the conventions. Using the vierbein, we relate the curved and flat Dirac γ matrices as

$$\overline{\gamma}_\mu(x) = e^A_\mu(x)\, \widetilde{\gamma}_A\,, \quad \overline{\gamma}^\mu(x) = e^\mu_A(x)\, \widetilde{\gamma}^A\,. \tag{8.1}$$

Then

$$g_{\mu\nu}(x) = \frac{1}{2}\left\{\overline{\gamma}_\mu(x), \overline{\gamma}_\nu(x)\right\} = \frac{1}{2}\left\{e^A_\mu(x)\widetilde{\gamma}_\mu, e^B_\nu(x)\widetilde{\gamma}\right\} = e^A_\mu(x) e^B_\nu(x)\eta_{AB}\,, \tag{8.2}$$

$$g^{\mu\nu}(x) = \frac{1}{2}\left\{\overline{\gamma}_\mu(x), \overline{\gamma}^\nu(x)\right\} = \frac{1}{2}\left\{e^\mu_A(x)\widetilde{\gamma}^\mu, e^\nu_B(x)\widetilde{\gamma}\right\} = e^\mu_A(x) e^\nu_B(x)\eta^{AB}\,. \tag{8.3}$$

Note: from here on we will be suppressing the "(x)", i.e., $g_{\mu\nu} = g_{\mu\nu}(x)$, and have a similar convention (suppression of the argument) for the vierbein coefficients. We know that $g^{\mu\rho}\, g_{\rho\nu} = \delta^\mu_\nu$, thus

$$g^{\mu\rho} g_{\rho\nu} = e^\mu_A e^\rho_B \eta^{AB}\, e^B_\rho e^C_\nu \eta_{BC} = e^\mu_A e^\rho_B e^B_\rho e^C_\nu \delta^A_C = e^\mu_A e^A_\nu e^\rho_B e^B_\rho \,, \tag{8.4}$$

since we must find that this expansion is equal to δ^μ_ν we conclude that the matrices composed of the vierbein and inverse vierbeins are themselves the inverses of each other,

$$\left(e^\rho_B\right)^{-1} = e^B_\rho \,, \quad \left(e^B_\rho\right)^{-1} = e^\rho_B \,, \quad e^\mu_A e^A_\nu = \delta^\mu_\nu \,, \quad e^A_\mu e^\mu_B = \delta^A_B \,. \tag{8.5}$$

We also define

$$e_{\mu A} = e^\nu_A\, g_{\nu\mu} = e^B_\nu\, \eta_{BA} \,, \quad e^{\mu A} = e^A_\nu\, g^{\nu\mu} = e^\nu_B\, \eta^{BA} \,. \tag{8.6}$$

Form (8.2) and (8.3) we easily find

$$e^A_\mu\, e_{\nu A} = g_{\mu\nu} \,, \quad e^\mu_A\, e^{\nu A} = g^{\mu\nu} \,, \tag{8.7}$$

$$e^\mu_A\, e_{\mu B} = \eta_{AB} \,, \quad e^A_\mu\, e^{\mu B} = \eta^{AB} \,. \tag{8.8}$$

8.2. COVARIANT DERIVATIVE OF A SPINOR

We now want to construct the covariant derivative for a spinor in curved space. The key observation, made by Brill and Wheeler [91] is that in the Dirac equation going from flat space to curved space, the derivative transforms as

$$\partial_\mu \psi \to \nabla_\mu \psi = \left(\partial_\mu - \Gamma_\mu\right) \psi \,. \tag{8.9}$$

The Dirac equation then reads as

$$\overline{\psi}\left[\gamma^{\mu}\left(\partial_{\mu}-\Gamma_{\mu}\right)-m\right]\psi=0\,. \tag{8.10}$$

We will now use the remainder of this section, as well as the following two sections to determine precisely what the spin connection matrix Γ_{μ} is.

As with the flat–space Dirac equation (see appendix C), we require that the Dirac equation in curved space is Lorentz invariant. A Lorentz transformation in the "internal" space (flat–space), reads as

$$e'^{\mu A}=\Lambda^{A}_{B}\,e^{\mu B}\,, \tag{8.11}$$

we can then use (8.6), as well as $\eta_{CA}\;\Lambda^{A}{}_{B}\;\eta^{BD}=\Lambda_{C}{}^{D}$, and multiply the previous expression by η_{CA} to find

$$e'^{\mu}_{C}=\Lambda_{C}{}^{D}\,e^{\mu}_{D}=\left(\Lambda^{-1}\right)^{C}{}_{D}\,e^{\mu}_{D}\,. \tag{8.12}$$

We now turn our attention to the curved space Dirac γ functions, defined in (8.1). A Lorentz transform will alter the vierbein, however the structure will remain valid, i.e.,

$$\overline{\gamma}'^{\mu}=e'^{\mu}_{A}\,\widetilde{\gamma}^{A}=\Lambda_{B}{}^{A}\,e^{\mu}_{B}\,\widetilde{\gamma}^{A}\,, \tag{8.13}$$

where we did not use "$\widetilde{\gamma}'^{A}$", because we know $\widetilde{\gamma}'^{A}=\widetilde{\gamma}^{A}$ (see appendix C). By including the spinor representation of the Lorentz transform, we find

$$\overline{\gamma}'^{\mu}=S(\Lambda)\,\overline{\gamma}^{\nu}\,S(\Lambda)^{-1}=\Lambda_{B}{}^{A}\,e^{\mu}_{B}\,\widetilde{\gamma}^{A}\,. \tag{8.14}$$

This result is similar to the result in flat–space (see appendix C, though we note that unlike in flat–space, $\overline{\gamma}'^{\mu} \neq \overline{\gamma}^{\mu}$). By rewriting the relation (C.43) as

$$\Lambda_A{}^B\, \eta^{AC}\, \Lambda_C{}^D = \eta^{BD}\,, \tag{8.15}$$

we can then show that although the curved–space Dirac γ matrices do change under a Lorentz transformation, the metric g remains constant, i.e.,

$$\begin{aligned} g'^{\mu\nu} &= \frac{1}{2}\left\{\overline{\gamma}'^{\mu}, \overline{\gamma}'^{\nu}\right\} = \frac{1}{2}\left\{\Lambda_B{}^A\, e^{\mu}_B\, \widetilde{\gamma}^A, \Lambda_C{}^D\, e^{\nu}_D\, \widetilde{\gamma}^C\right\} = e^{\mu}_B\, e^{\nu}_D\, \Lambda_A{}^B\, \frac{1}{2}\left\{\widetilde{\gamma}^A, \widetilde{\gamma}^C\right\} \Lambda_C{}^D \\ &= e^{\mu}_B\, e^{\nu}_D\, \Lambda_A{}^B\, \eta^{AC}\, \Lambda_C{}^D = e^{\mu}_B e^{\nu}_D \eta^{BD} = g^{\mu\nu}\,. \end{aligned} \tag{8.16}$$

The same is true for a metric with lower indices, i.e., $g'_{\mu\nu} = g_{\mu\nu}$, following a virtually identical derivation. We can now look at the Dirac equation, which transforms under a Lorentz transform as

$$\overline{\psi}\, (\mathrm{i}\, \overline{\gamma}^{\mu}\, \nabla_{\mu} - m)\, \psi \to \overline{\psi}'\, (\mathrm{i}\, \overline{\gamma}'^{\nu}\, \nabla'_{\nu} - m)\, \psi'\,. \tag{8.17}$$

We now recall that

$$\psi' = S(\Lambda)\, \psi\,, \quad \overline{\psi}' = \overline{\psi}\, S(\Lambda)^{-1}\,, \tag{8.18}$$

in which case (8.17) becomes

$$\begin{aligned} &\overline{\psi}\, S(\Lambda)^{-1}\, \left[\mathrm{i}\left(S(\Lambda)\, \overline{\gamma}^{\mu}\, S(\Lambda)^{-1}\right) \nabla'_{\mu} - m\right] S(\Lambda)\, \psi \\ &= \overline{\psi}\left[\mathrm{i}\, \overline{\gamma}^{\mu}\, \left(S(\Lambda)^{-1}\, \nabla'_{\mu}\, S(\Lambda)\right) - m\right] \psi\,. \end{aligned} \tag{8.19}$$

Then for the curved–space Dirac equation to be Lorentz invariant, we require that

$$S(\Lambda)^{-1}\,\nabla'_\mu\,S(\Lambda)\psi = \nabla_\mu\,\psi\,, \tag{8.20}$$

i.e.,

$$\nabla'_\mu\,S(\Lambda)\,\psi = S(\Lambda)\,\nabla_\mu\,\psi\,. \tag{8.21}$$

Furthermore, we do not want the Lorentz transformation to alter the fundamental structure of the covariant derivative acting on a spinor, i.e.,

$$\nabla'_\mu = \partial_\mu - \Gamma'_\mu\,. \tag{8.22}$$

Plugging (8.22) into the l.h.s. of (8.21) we find,

$$\begin{aligned}\left(\partial_\mu - \Gamma'_\mu\right) S(\Lambda)\psi &= \left(\partial_\mu\,S(\Lambda)\right)\psi + S(\Lambda)\,\partial_\mu\,\psi - \Gamma'_\mu\,S(\Lambda)\,\psi \\ &= S(\Lambda)\,\left[\partial_\mu + S(\Lambda)^{-1}\,\partial_\mu\,S(\Lambda) - S(\Lambda)^{-1}\,\Gamma'_\mu\,S(\Lambda)\right]\psi\,. \end{aligned}\tag{8.23}$$

Comparing this result with the l.h.s. of (8.21), we are left to conclude that

$$\Gamma_\mu = S(\Lambda)^{-1}\,\Gamma'_\mu\,S(\Lambda) - S(\Lambda)^{-1}\,\partial_\mu\,S(\Lambda)\,, \tag{8.24}$$

which we can reformulate as

$$\Gamma'_\mu = S(\Lambda)\,\Gamma_\mu\,S(\Lambda)^{-1} + \left[\partial_\mu\,S(\Lambda)\right]\,S(\Lambda)^{-1}\,. \tag{8.25}$$

As discussed in appendix C,

$$S(\Lambda) = \exp\left(-\frac{\mathrm{i}}{4}\Omega^{AB}\sigma_{AB}\right)\,, \quad S(\Lambda)^{-1} = \exp\left(\frac{\mathrm{i}}{4}\Omega^{AB}\sigma_{AB}\right)\,, \tag{8.26}$$

where Ω^{AB} are the generators of the Lorentz transformation. Then (8.24) and (8.25) become

$$\Gamma_\mu = S(\Lambda)^{-1}\,\Gamma'_\mu\,S(\Lambda) - \frac{\mathrm{i}}{4}\left(\partial_\mu\,\Omega^{AB}\right)\sigma_{AB}\,, \quad \Gamma'_\mu = S(\Lambda)\,\Gamma_\mu\,S(\Lambda)^{-1} + \frac{\mathrm{i}}{4}\left(\partial_\mu\,\Omega^{AB}\right)\sigma_{AB}\,. \tag{8.27}$$

Despite the fact that the Γ_μs are changed by Lorentz transformations, their overall structure should be consistent. As such, based on (8.27) we deduce that

$$\Gamma_\mu = \frac{\mathrm{i}}{4} C_\mu^{AB} \sigma_{AB}\,, \tag{8.28}$$

where C_μ^{AB} is antisymmetric since Ω^{AB} is antisymmetric (see chapter 13.1 of [126]). In the following sections we will calculate the C_μ^{AB} coefficients.

8.3. COVARIANT DERIVATIVE OF THE DIRAC γ MATRICES

An obvious extension of the formalism outlined above pertains to the covariant derivative of Dirac γ matrices which is useful to clarify in a more general context. We begin with the Dirac equation in curved space,

$$\overline{\psi}\,[\mathrm{i}\overline{\gamma}^\mu\left(\partial_\mu - \Gamma_\mu\right) - m]\,\psi = 0\,. \tag{8.29}$$

We then take the adjoint,

$$\psi^+\left[-\mathrm{i}\left(\overleftarrow{\partial}_\mu - \Gamma_\mu^+\right)(\overline{\gamma}^\mu)^+ - m\right]\overline{\psi}^+ = 0\,. \tag{8.30}$$

Regardless of the space we are working in, we define $\overline{\psi}$ as

$$\overline{\psi} \equiv \psi^+\widetilde{\gamma}^0\,, \tag{8.31}$$

with the flat-space $\widetilde{\gamma}^0$, ensuring that $\overline{\psi}$ transforms with the inverse of the local Lorentz transform. With this in mind, and the fact that $(\widetilde{\gamma}^0)^2 = 1$, (8.30) becomes

$$\begin{aligned} &\psi^+\widetilde{\gamma}^0\widetilde{\gamma}^0 \left[-\mathrm{i}\left(\overleftarrow{\partial}_\mu - \Gamma^+_\mu\right)(\overline{\gamma}^\mu)^+ - m\right]\widetilde{\gamma}^0\psi \\ &= \overline{\psi}\left[-\mathrm{i}\left(\overleftarrow{\partial}_\mu - \widetilde{\gamma}^0\Gamma^+_\mu\widetilde{\gamma}^0\right)\widetilde{\gamma}^0(\overline{\gamma}^\mu)^+\widetilde{\gamma}^0 - m\right]\psi = 0\,. \end{aligned} \tag{8.32}$$

Since $\widetilde{\gamma}^0(\widetilde{\gamma}^\mu)^+\widetilde{\gamma}^0 = \widetilde{\gamma}^\mu$, it is trivial to show that $\widetilde{\gamma}^0(\overline{\gamma}^\mu)^+\widetilde{\gamma}^0 = \overline{\gamma}^\mu$. Recall the form of Γ_μ as given in (8.28), then

$$\widetilde{\gamma}^0(\Gamma_\mu)^+\widetilde{\gamma}^0 = \widetilde{\gamma}^0\left(\frac{\mathrm{i}}{4}C^{AB}_\mu\sigma_{AB}\right)^+\widetilde{\gamma}^0 = -\frac{\mathrm{i}}{4}C^{AB}_\mu\widetilde{\gamma}^0(\sigma_{AB})^+\widetilde{\gamma}^0\,, \tag{8.33}$$

where

$$\begin{aligned} \widetilde{\gamma}^0(\sigma_{AB})^+\widetilde{\gamma}^0 &= \widetilde{\gamma}^0\left(\frac{\mathrm{i}}{2}[\widetilde{\gamma}_A,\widetilde{\gamma}_B]\right)^+\widetilde{\gamma}^0 = -\frac{\mathrm{i}}{2}\widetilde{\gamma}^0\left[(\widetilde{\gamma}_B)^+,(\widetilde{\gamma}_A)^+\right]\widetilde{\gamma}^0 \\ &= -\frac{\mathrm{i}}{2}\left[\widetilde{\gamma}^0(\widetilde{\gamma}_B)^+\widetilde{\gamma}^0, \widetilde{\gamma}^0(\widetilde{\gamma}_A)^+\widetilde{\gamma}^0\right] \\ &= -\frac{\mathrm{i}}{2}[\widetilde{\gamma}_B,\widetilde{\gamma}_A] = \frac{\mathrm{i}}{2}[\widetilde{\gamma}_A,\widetilde{\gamma}_B] = \sigma_{AB}\,, \end{aligned} \tag{8.34}$$

where we used

$$\widetilde{\gamma}^0(\widetilde{\gamma}_A)^+\widetilde{\gamma}^0 = \widetilde{\gamma}^0\left(\eta_{AB}\widetilde{\gamma}^B\right)^+\widetilde{\gamma}^0 = \eta_{AB}\widetilde{\gamma}^0\left(\widetilde{\gamma}^B\right)^+\widetilde{\gamma}^0 = \eta_{AB}\widetilde{\gamma}^B = \widetilde{\gamma}_A\,. \tag{8.35}$$

Plugging this into (8.33) we find

$$\widetilde{\gamma}^0\Gamma^+_\mu\widetilde{\gamma}^0 = -\frac{\mathrm{i}}{4}C^{AB}_\mu\sigma_{AB} = -\Gamma_\mu\,. \tag{8.36}$$

Thus (8.32) becomes

$$\overline{\psi}\left[-\mathrm{i}\left(\overleftarrow{\partial}_\mu + \Gamma_\mu\right)\overline{\gamma}^\mu - m\right]\psi = 0\,. \tag{8.37}$$

We now equate (8.29) and (8.37) and find

$$\begin{aligned}
&\overline{\psi}\left[\mathrm{i}\,\overline{\gamma}^{\mu}\left(\partial_{\mu}-\Gamma_{\mu}\right)-m\right]\psi=\overline{\psi}\left[-\mathrm{i}\left(\overleftarrow{\partial}_{\mu}+\Gamma_{\mu}\right)\overline{\gamma}^{\mu}-m\right]\psi \\
&\Rightarrow \overline{\psi}\,\overline{\gamma}^{\mu}\left(\partial_{\mu}\,\psi\right)+\left(\partial_{\mu}\,\overline{\psi}\right)\overline{\gamma}^{\mu}\,\psi=\overline{\psi}\left(\overline{\gamma}^{\mu}\,\Gamma_{\mu}-\Gamma_{\mu}\,\overline{\gamma}^{\mu}\right)\psi \\
&\Rightarrow \partial_{\mu}\left(\overline{\psi}\,\overline{\gamma}^{\mu}\,\psi\right)=\overline{\psi}\left(\partial_{\mu}\,\overline{\gamma}^{\mu}\right)\psi-\overline{\psi}\left[\Gamma_{\mu},\overline{\gamma}^{\mu}\right]\psi\,.
\end{aligned} \tag{8.38}$$

Much like in flat–space (as discussed in appendix C), the curved space probability current is given as

$$j^{\mu}=\overline{\psi}\,\overline{\gamma}^{\mu}\,\psi\,, \tag{8.39}$$

and is conserved using the covariant derivative (D.29), i.e.,

$$\nabla_{\mu}\,j^{\mu}=0\,. \tag{8.40}$$

Thus,

$$\partial_{\mu}\,j^{\mu}+\Gamma^{\mu}_{\mu\rho}\,j^{\rho}=\partial_{\mu}\left(\overline{\psi}\,\overline{\gamma}^{\mu}\,\psi\right)+\Gamma^{\mu}_{\mu\rho}\left(\overline{\psi}\,\overline{\gamma}^{\rho}\,\psi\right)=0\,. \tag{8.41}$$

We now apply (8.38), yielding

$$\overline{\psi}\left(\partial_{\mu}\,\overline{\gamma}^{\mu}\right)\psi-\overline{\psi}\left[\Gamma_{\mu},\overline{\gamma}^{\mu}\right]\psi+\overline{\psi}\,\Gamma^{\mu}_{\mu\rho}\,\gamma^{\rho}\,\psi=\overline{\psi}\,\delta^{\mu}_{\nu}\left(\left(\partial_{\mu}\,\overline{\gamma}^{\nu}\right)+\Gamma^{\nu}_{\mu\rho}\,\overline{\gamma}^{\rho}-\left[\Gamma_{\mu},\overline{\gamma}^{\nu}\right]\right)\psi=0\,. \tag{8.42}$$

Because this equation has to be valid for any ψ, we immediately have

$$\delta^{\mu}_{\nu}\left(\partial_{\mu}\,\overline{\gamma}^{\nu}\right)+\Gamma^{\nu}_{\mu\rho}\,\overline{\gamma}^{\rho}-\left[\Gamma_{\mu},\overline{\gamma}^{\nu}\right]=\delta^{\mu}_{\nu}\,\nabla_{\mu}\,\overline{\gamma}^{\nu}=0\,, \tag{8.43}$$

where we define the covariant derivative of a Dirac γ matrix as

$$\nabla_\mu \overline{\gamma}^\nu = \partial_\mu \overline{\gamma}^\nu + \Gamma^\nu_{\mu\rho} \overline{\gamma}^\rho - [\Gamma_\mu, \gamma^\rho] \,. \tag{8.44}$$

8.4. FINALLY SOLVING FOR C_μ^{AB} AND Γ_μ

In order to find C_μ^{AB}, we impose the restriction that $\nabla_\mu \overline{\gamma}^\nu = 0$, which of course still satisfies (8.43), and apply the the vierbein ($\overline{\gamma}^\mu = e^\mu_A \widetilde{\gamma}^A$) to (8.44), and find

$$\left(\partial_\mu e^\nu_B + \Gamma^\nu_{\mu\rho} e^\rho_B\right) \widetilde{\gamma}^B - e^\nu_B \left[\Gamma_\mu, \widetilde{\gamma}^B\right] = \nabla_\mu e^\nu_B \widetilde{\gamma}^B - e^\nu_B \left[\Gamma_\mu, \widetilde{\gamma}^B\right] = 0 \,. \tag{8.45}$$

We now multiply by e^A_ν to obtain

$$e^A_\nu \nabla_\mu e^\nu_B \widetilde{\gamma}^B - e^A_\nu e^\nu_B \left[\Gamma_\mu, \widetilde{\gamma}^B\right] = \omega^A_{\mu B} \widetilde{\gamma}^B - \delta^A_B \left[\Gamma_\mu, \widetilde{\gamma}^B\right] = \omega^A_{\mu B} \widetilde{\gamma}^B - \left[\Gamma_\mu, \widetilde{\gamma}^A\right] = 0 \,. \tag{8.46}$$

In order to proceed, we use our ansatz of Γ_μ (8.28), and (C.67), to find

$$\begin{aligned}\left[\Gamma_\mu, \widetilde{\gamma}^A\right] &= \frac{\mathrm{i}}{4} C^{BC}_\mu \left[\sigma_{BC}, \widetilde{\gamma}^A\right] = \frac{\mathrm{i}}{4} C^{BC}_\mu \left[2\,\mathrm{i}\left(\delta^A_C \widetilde{\gamma}_B - \delta^A_B \widetilde{\gamma}_C\right)\right] = -\frac{1}{2}\left(C^{BA}_\mu \widetilde{\gamma}_B - C^{AC}_\mu \widetilde{\gamma}_C\right) \\ &= \frac{1}{2}\left(C^{AB}_\mu \widetilde{\gamma}_B + C^{AC}_\mu \widetilde{\gamma}_C\right) = C^{AB}_\mu \widetilde{\gamma}_B = \eta_{BC} C^{AB}_\mu \widetilde{\gamma}^C = \eta_{BC} C^{AC}_\mu \widetilde{\gamma}^B \,,\end{aligned} \tag{8.47}$$

where we used the assumption that C^{AB}_μ is antisymmetric ($C^{AB}_\mu = -C^{BA}_\mu$). Combining this result with (8.46) we obtain

$$\omega^A_{\mu B} \widetilde{\gamma}^B - \eta_{BC} C^{AC}_\mu \widetilde{\gamma}^B = \eta_{BC}\left(\omega^{AC}_\mu - C^{AC}_\mu\right) \widetilde{\gamma}^B = \left(\omega^{AC}_\mu - C^{AC}_\mu\right) \widetilde{\gamma}_B = 0 \,. \tag{8.48}$$

We then conclude that

$$C^{AB}_\mu = \omega^{AB}_\mu \,, \tag{8.49}$$

and

$$\Gamma_\mu = \frac{\mathrm{i}}{4}\omega_\mu^{AB}\sigma_{AB}\,. \tag{8.50}$$

We also note that our assumption that C_μ^{AB} is antisymmetric has been validated as we know ω_μ^{AB} is antisymmetric (see appendix D). Finally we can write the covariant derivative operating on a spinor as

$$\nabla_\mu\psi = \left(\partial_\mu - \Gamma_\mu\right)\psi = \left(\partial_\mu - \frac{\mathrm{i}}{4}\omega_\mu^{AB}\,\sigma_{ab}\right)\psi\,. \tag{8.51}$$

In principle, this result is well known [87–98], including a wrong prefactor in the article of Brill and Wheeler [91], as pointed out by Jentschura in [92], however the derivations in the literature are not detailed.

8.5. DERIVATION OF CURVED–SPACE DIRAC HAMILTONIANS

Here we will use the tools developed in the preceding sections to derive both the Dirac Hamiltonians for a particle coupled to a gravitational field, and a particle in a non-inertial, rotating reference frame. The derivation presented here is a slight variation on the derivations presented in [87, 92, 94]. While these are not the only types of interactions we will be examining, they share a unifying concept in that we cannot simply apply the correspondence principle, in which the potential $V(r)$ is simply added to the Dirac Hamiltonian for a free particle ($H = H_F + V(r)$). The curvature of spacetime due to the presence of a gravitational field is not taken into consideration by the correspondence principle if we simply use the Newtonian potential $-GmM/r$ for the potential $V(r)$, and neither can the non-inertial nature of the particle be taken into account by such a simple formula. Fortunately, both the isotropic Schwarzschild metric, and the uniformly accelerated non-inertial metric, can be represented in the form

$$[g_{\mu\nu}] = \text{diag}\left(w^2, -v^2, -v^2, -v^2\right) \,, \quad [g^{\mu\nu}] = \text{diag}\left(w^{-2}, -v^{-2}, -v^{-2}, -v^{-2}\right) \,, \tag{8.52}$$

where $w = w(\vec{r})$ and $v = v(\vec{r})$ are functions of the spatial coordinates $\vec{r}$ (and are independent of time, t). Recall that according to (8.2) and (8.3),

$$g_{\mu\nu} = e^A_\mu \, e^B_\nu \, \eta_{AB} \,, \quad g^{\mu\nu} = e^\mu_A \, e^\nu_B \, \eta^{AB} \,.$$

Each of these equations generates a system of 16 equations, solved by

$$e^A_0 = w \, \delta^A_0 \,, \quad e^0_\mu = w \, \delta^0_\mu \,, \quad e^A_i = v \, \delta^A_i \,, \quad e^I_\mu = v \, \delta^I_\mu \,, \tag{8.53}$$

$$e^0_A = \frac{\delta^0_A}{w} \,, \quad e^\mu_0 = \frac{\delta^\mu_a}{w} \,, \quad e^i_A = \frac{\delta^i_A}{v} \,, \quad e^\mu_I = \frac{\delta^\mu_I}{v} \,, \tag{8.54}$$

where δ is the Kronecker symbol. We know that the Dirac equation in curved–spacetime is given as (8.10)

$$(\mathrm{i}\overline{\gamma}^\mu \nabla_\mu - m) \, \psi = (\mathrm{i}\overline{\gamma}^\mu \partial_\mu - \mathrm{i}\overline{\gamma}^\mu \Gamma_\mu - m) \, \psi = 0 \,. \tag{8.55}$$

By multiplying by $\overline{\gamma}^0$ on the left, we can rewrite the curved–space Dirac equation as

$$\mathrm{i} \left(\overline{\gamma}^0\right)^2 \partial_o \psi = \left(-\mathrm{i}\overline{\gamma}^0 \, \overline{\gamma}^i \partial_i + \mathrm{i}\overline{\gamma}^0 \, \overline{\gamma}^\mu \, \Gamma_\mu + m\right) \psi \,. \tag{8.56}$$

To proceed we need to calculate the connection matrices $\Gamma_\mu = \frac{\mathrm{i}}{4} \omega^{AB}_\mu \widetilde{\sigma}_{AB}$ (equation (8.50)) explicitly. As we continue, the flat–space and curved–space indices will start to mix, so we have written the flat–space spin connection matrix with a tilde, "$\widetilde{\sigma}$," to avoid any possible confusion. From equation (D.35) we know

$$\omega_\mu^{AB} = e_\nu^A \nabla_\mu e^{\nu B} = e_\nu^A \partial_\mu e_C^\nu \eta^{CB} + e_\nu^A \Gamma_{\mu\rho}^\nu e_C^\rho \eta^{CB} . \tag{8.57}$$

We begin by solving for the first term on the r.h.s., i.e.,

$$\begin{aligned} e_\nu^A \partial_\mu e_C^\nu \eta^{CB} &= e_0^A \partial_\mu e_C^0 \eta^{CB} + e_i^A \partial_\mu e_C^i \eta^{CB} = w\, \delta_0^A \partial_\mu \frac{1}{w} \delta_C^0 \eta^{CB} + v\, \delta_i^A \partial_\mu \frac{1}{v} \delta_C^i \eta^{CB} \\ &= -\frac{\partial_\mu w}{w} \eta^{0B} \delta_0^A - \frac{\partial_\mu v}{v} \eta^{IB} \delta_i^A . \end{aligned} \tag{8.58}$$

The connection matrix Γ_μ becomes

$$\begin{aligned} \Gamma_\mu &= \frac{\mathrm{i}}{4} \left(-\frac{\partial_\mu w}{w} \eta^{0B} \delta_0^A \widetilde{\sigma}_{AB} - \frac{\partial_\mu v}{v} \eta^{IB} \delta_i^A \widetilde{\sigma}_{AB} + e_\nu^A \Gamma_{\mu\rho}^\nu e_C^\rho \eta^{CB} \widetilde{\sigma}_{AB} \right) \\ &= \frac{\mathrm{i}}{4} \left(-\frac{\partial_\mu w}{w} \eta^{0B} \widetilde{\sigma}_{0B} - \frac{\partial_\mu v}{v} \eta^{IB} \widetilde{\sigma}_{IB} + e_\nu^A \Gamma_{\mu\rho}^\nu e_C^\rho \eta^{CB} \widetilde{\sigma}_{AB} \right) . \end{aligned} \tag{8.59}$$

But by definition, $\widetilde{\sigma}_{AB} = \frac{\mathrm{i}}{2} [\widetilde{\gamma}_A, \widetilde{\gamma}_B]$, thus if $A = B$, $\widetilde{\sigma}_{AB} = 0$. Additionally, our metric η is diagonal, so if $A \neq B$ then $\eta^{AB} = 0$. With this in mind, it is clear that the first two terms in Γ_μ must be 0, and (8.59) becomes

$$\Gamma_\mu = \frac{\mathrm{i}}{4} e_\nu^A \Gamma_{\mu\rho}^\nu e_C^\rho \eta^{CB} \widetilde{\sigma}_{AB} . \tag{8.60}$$

We now need to calculate the Christoffel symbols, and the spin matrices. Let us begin with the Christoffel symbols, bearing in mind that $\Gamma_{\mu\nu}^\rho = \Gamma_{\nu\mu}^\rho$, and that w, and v are time independent, and thus $\partial_0 g_{\mu\nu} = 0$. We find

$$\Gamma_{00}^0 = 0 , \quad \Gamma_{0i}^0 = \Gamma_{i0}^0 = \frac{\partial_i w^2}{2w^2} , \quad \Gamma_{ij}^0 = 0 , \quad \Gamma_{00}^i = \frac{\partial_i w^2}{2v^2} , \quad \Gamma_{0j}^i = \Gamma_{j0}^i = 0 , \tag{8.61a}$$

$$\Gamma_{jk}^i = \frac{1}{2v^2} \left(\delta_k^i \partial_j v^2 + \delta_j^i \partial_k v^2 - \delta_{jk} \partial^i v^2 \right) . \tag{8.61b}$$

We now need to consider Γ_0 and Γ_i separately. Setting $\mu = 0$, the expression in equation (8.60) becomes

$$\begin{aligned}\Gamma_0 &= \frac{\mathrm{i}}{4} e_\nu^A \, \Gamma^\nu_{0\rho} \, e_C^\rho \, \eta^{CB} \, \widetilde{\sigma}_{AB} \\ &= \frac{\mathrm{i}}{4}\left(e_0^A \, \Gamma^0_{00} \, e_C^0 + e_0^A \, \Gamma^0_{0i} \, e_C^i + e_i^A \, \Gamma^i_{00} \, e_C^0 + e_i^A \, \Gamma^i_{0j} \, e_C^j\right) \eta^{CB} \, \widetilde{\sigma}_{AB} \,. \end{aligned} \tag{8.62}$$

We now plug our results from (8.61) into this equation to find

$$\begin{aligned}\Gamma_0 &= \frac{\mathrm{i}}{4}\left(e_0^A \frac{\partial_i w^2}{2w^2} \, e_C^i + e_i^A \frac{\partial_i w^2}{2v^2} \, e_C^0\right) \eta^{CB} \, \widetilde{\sigma}_{AB} \\ &= \frac{\mathrm{i}}{4}\left(w \, \delta_0^A \frac{\partial_i w^2}{2w^2} \frac{\delta_C^i}{v} + v \, \delta_i^A \frac{\partial_i w^2}{2v^2} \frac{\delta_C^0}{w}\right) \eta^{CB} \, \widetilde{\sigma}_{AB} = \frac{\mathrm{i}}{4} \frac{\partial_i w^2}{2wv}\left(\eta^{iB} \, \widetilde{\sigma}_{0B} + \eta^{0B} \, \widetilde{\sigma}_{iB}\right) \,. \end{aligned} \tag{8.63}$$

We now simplify the bracketed terms,

$$\begin{aligned}\eta^{iB} \, \widetilde{\sigma}_{0B} + \eta^{0B} \, \widetilde{\sigma}_{iB} &= \frac{\mathrm{i}}{2}\left(\eta^{iB} \left[\widetilde{\gamma}_0, \widetilde{\gamma}_B\right] + \eta^{0B} \left[\widetilde{\gamma}_i, \widetilde{\gamma}_B\right]\right) = \frac{\mathrm{i}}{2}\left(\left[\widetilde{\gamma}_0, \widetilde{\gamma}^i\right] + \left[\widetilde{\gamma}_i, \widetilde{\gamma}^0\right]\right) \\ &= \frac{\mathrm{i}}{2}\left(\left[\widetilde{\gamma}^0, \widetilde{\gamma}^i\right] - \left[\widetilde{\gamma}^i, \widetilde{\gamma}^0\right]\right) = \frac{\mathrm{i}}{2}\left(\left[\widetilde{\gamma}^0, \widetilde{\gamma}^i\right] + \left[\widetilde{\gamma}^0, \widetilde{\gamma}^i\right]\right) \\ &= \mathrm{i}\left[\widetilde{\gamma}^0, \widetilde{\gamma}^i\right] = 2\mathrm{i}\widetilde{\gamma}^0\widetilde{\gamma}^i \,, \end{aligned} \tag{8.64}$$

where we used the (somewhat trivial) identities

$$\widetilde{\gamma}_0 = \eta_{0A}\widetilde{\gamma}^A = (1)\widetilde{\gamma}^0 + (0)\widetilde{\gamma}^1 + (0)\widetilde{\gamma}^2 + (0)\widetilde{\gamma}^3 = \widetilde{\gamma}^0 \,, \tag{8.65a}$$

$$\widetilde{\gamma}_1 = \eta_{1A}\widetilde{\gamma}^A = (0)\widetilde{\gamma}^0 + (-1)\widetilde{\gamma}^1 + (0)\widetilde{\gamma}^2 + (0)\widetilde{\gamma}^3 = -\widetilde{\gamma}^1 \,. \tag{8.65b}$$

Hence we can show that $\widetilde{\gamma}_I = -\widetilde{\gamma}^I$ by generalizing the final equation for $\widetilde{\gamma}_2$ and $\widetilde{\gamma}_3$. Thus, equation (8.63) becomes

$$\Gamma_0 = -\widetilde{\gamma}^0\widetilde{\gamma}^i \frac{\partial_i \, w^2}{4wv} = -\widetilde{\gamma}^0\widetilde{\gamma}^i \frac{\partial_i \, w}{2v} = -\frac{\vec{\alpha} \cdot \vec{\nabla} w}{2v} \,. \tag{8.66}$$

We now caculate Γ_i,

$$\begin{aligned}\Gamma_i &= \frac{\mathrm{i}}{4} e^A_\nu \, \Gamma^\nu_{i\rho} \, e^\rho_C \, \eta^{CB} \, \widetilde{\sigma}_{AB} \\ &= \frac{\mathrm{i}}{4} \left(e^A_0 \, \Gamma^0_{i0} \, e^0_C + e^A_0 \, \Gamma^0_{ij} \, e^j_C + e^A_j \, \Gamma^j_{i0} \, e^0_C + e^A_j \, \Gamma^j_{ik} \, e^k_C \right) \eta^{CB} \, \widetilde{\sigma}_{AB} \, . \end{aligned} \tag{8.67}$$

As before, we plug in our results form (8.61) to find

$$\Gamma_i = \frac{\mathrm{i}}{4} \left(e^A_0 \frac{\partial_i w^2}{2w^2} e^0_C + e^A_j \frac{1}{2v^2} \left(\delta^j_k \, \partial_i \, v^2 + \delta^j_i \, \partial_k v^2 - \delta_{ik} \, \partial^j v^2 \right) e^k_C \right) \eta^{CB} \, \widetilde{\sigma}_{AB} \, . \tag{8.68}$$

Note that $e^A_0 e^0_C \eta^{CB} = \eta^{AB}$, which means the first term will be proportional to $\eta^{AB}\sigma_{AB} = 0$ (if $A = B$ then $\sigma_{AB} = 0$, and if $A \neq B$ then $\eta^{AB} = 0$). Thus, equation (8.68) becomes

$$\begin{aligned}\Gamma_i &= \frac{\mathrm{i}}{4} \left(v \delta^A_j \right) \frac{1}{2v^2} \left(\delta^j_k \, \partial_i v^2 + \delta^j_i \, \partial_k v^2 - \delta_{ik} \, \partial_j v^2 \right) \left(\frac{1}{v} \delta^k_C \right) \eta^{CB} \, \widetilde{\sigma}_{AB} \\ &= \frac{\mathrm{i}}{8v^2} \left(\delta^j_k \, \partial_i v^2 + \delta^j_i \, \partial_k v^2 - \delta_{ik} \, \partial^j v^2 \right) \eta^{kB} \, \widetilde{\sigma}_{jB} \\ &= \frac{\mathrm{i}}{8v^2} \left(\eta^{jB} \, \widetilde{\sigma}_{jB} \, \partial_j v^2 + \eta^{kB} \, \widetilde{\sigma}_{iB} \partial_k v^2 - \eta^B_i \, \widetilde{\sigma}_{jB} \partial^j v^2 \right) \, . \end{aligned} \tag{8.69}$$

We again use the fact that $\eta^{AB}\sigma_{AB} = 0$, and switch the implicit sum over k in the second term, to an implicit sum over j, yielding

$$\Gamma_i = \mathrm{i} \frac{\partial_j v^2}{8v^2} \left(\eta^{jB} \, \widetilde{\sigma}_{iB} - \eta^B_i \, \eta^{jA} \, \widetilde{\sigma}_{AB} \right) \, , \tag{8.70}$$

where we used the fact that

$$\eta^B_i \, \widetilde{\sigma}_{jB} \, \partial^j = \eta^B_i \, \widetilde{\sigma}^j_B \partial_j = \eta^B_i \, \eta^{jA} \widetilde{\sigma}_{AB} \partial_j \, . \tag{8.71}$$

We now look at the bracketed term in (8.70), in which it is clear that if $i = j$ the term will vanish (as will Γ_i), so we must assume $i \neq j$

$$\eta^{jB}\,\widetilde{\sigma}_{iB} - \eta_i^B\,\eta^{jA}\,\widetilde{\sigma}_{AB} = \frac{\mathrm{i}}{2}\left(\left[\widetilde{\gamma}_i, \widetilde{\gamma}^j\right] - \left[\widetilde{\gamma}^j, \widetilde{\gamma}_i\right]\right) = \mathrm{i}\left[\widetilde{\gamma}_i, \widetilde{\gamma}^j\right] = -\mathrm{i}\left[\widetilde{\gamma}^i, \widetilde{\gamma}^j\right] . \tag{8.72}$$

Plugging this result into (8.70) we find

$$\Gamma_i = \frac{\partial_j v^2}{8v^2}\left[\widetilde{\gamma}^i, \widetilde{\gamma}^j\right] = \frac{\partial_j v}{4v}\left[\widetilde{\gamma}^i, \widetilde{\gamma}^j\right] . \tag{8.73}$$

We now note that the term involving Γ_μ in (8.56) is $\overline{\gamma}^0\overline{\gamma}^\mu\Gamma_\mu$, which we must calculate. Using our results from (8.66) we begin with

$$\overline{\gamma}^0\overline{\gamma}^0\Gamma_0 = \frac{1}{w^2}\widetilde{\gamma}^0\widetilde{\gamma}^0\left(-\frac{\vec{\alpha}\cdot\vec{\nabla}w}{2v}\right) = -\frac{\vec{\alpha}\cdot\nabla w}{2vw^2} . \tag{8.74}$$

We now use the identity $\widetilde{\gamma}^i[\widetilde{\gamma}^i, \widetilde{\gamma}^j] = -2\widetilde{\gamma}^j$ *where we do not have an implicit sum over* i. With a sum over the index i we find $\sum_i \widetilde{\gamma}^i[\widetilde{\gamma}^i, \widetilde{\gamma}^j] = -4\widetilde{\gamma}^j$ (i can take on 3 values, but when $i = j$ the commutator vanishes, leaving us with two terms). With this in mind, as well as (8.73) we find

$$\overline{\gamma}^0\overline{\gamma}^i\Gamma_i = \frac{1}{vw}\widetilde{\gamma}^0\widetilde{\gamma}^i\left(\frac{\partial_j v}{4v}\left[\widetilde{\gamma}^i, \widetilde{\gamma}^j\right]\right) = \frac{\partial_j v}{4v^2 w}\widetilde{\gamma}^0\left(-4\widetilde{\gamma}^j\right) = -\frac{\vec{\alpha}\cdot\vec{\nabla}v}{v^2 w} . \tag{8.75}$$

Thus

$$\overline{\gamma}^0\overline{\gamma}^\mu\Gamma_\mu = \overline{\gamma}^0\overline{\gamma}^0\Gamma_0 + \overline{\gamma}^0\overline{\gamma}^i\Gamma_i = -\frac{\vec{\alpha}\cdot\vec{\nabla}w}{2vw^2} - \frac{\vec{\alpha}\cdot\vec{\nabla}v}{v^2 w} . \tag{8.76}$$

Let us quickly simplify the l.h.s. of (8.56), i.e.,

$$\mathrm{i}(\overline{\gamma}^0)^2\partial_0\psi = \mathrm{i}\frac{1}{w^2}(\widetilde{\gamma}^0)^2\partial_0\psi = \frac{\mathrm{i}}{w^2}\partial_0\psi . \tag{8.77}$$

We now apply (8.76) to the r.h.s. of (8.56), multiply both sides by w^2, and utilize the vierbein connecting the curved– and flat–space Dirac γ matrices to find,

$$\begin{aligned} \mathrm{i}\partial_0\psi &= w^2\left(-\mathrm{i}\frac{1}{vw}\widetilde{\gamma}^0\widetilde{\gamma}^j\partial_j + \mathrm{i}\left(-\frac{\vec{\alpha}\cdot(\vec{\nabla}w)}{2vw^2} - \frac{\vec{\alpha}\cdot(\vec{\nabla}v)}{v^2w}\right) + \frac{1}{w}\widetilde{\gamma}^0 m\right)\psi \\ &= \left(\frac{w}{v}\vec{\alpha}\cdot\vec{p} + \frac{\vec{\alpha}\cdot(\vec{p}\,w)}{2v} + \frac{w}{v}\frac{\vec{\alpha}\cdot(\vec{p}\,v)}{v} + w\beta m\right)\psi\,, \end{aligned} \tag{8.78}$$

notice that we have parenthesis around "$\vec{p}\,w$" and "$\vec{p}\,v$". This indicates that the momentum operator $\vec{p}$ is only acting on the function w or v (respectively), and *not* on the wave–function ψ. The resulting equation has the familiar form of the time dependent Schödinger equation, $\mathrm{i}\partial_t\psi = H\,\psi$, since $\partial_0 = \partial_t$. It is then clear that the Hamiltonian should be identified as

$$H = \frac{w}{v}\vec{\alpha}\cdot\vec{p} + \frac{\vec{\alpha}\cdot(\vec{p}\,w)}{2v} + \frac{w}{v}\frac{\vec{\alpha}\cdot(\vec{p}\,v)}{v} + \beta m\,w\,. \tag{8.79}$$

However, this form is not Hermitian, and therefore cannot act on a well-defined Hilbert space of functions. It is possible to massage the Hamiltonian into a more compact form. To do so we we rescale the wave–function according to

$$\psi' = v^{3/2}\,\psi\,, \quad H' = v^{3/2}\,H\,v^{-3/2}\,, \tag{8.80}$$

and it is immediately clear that the only term in our Hamiltonian that will be affected is the $\vec{\alpha}\cdot\vec{p}$ term that operates on ψ, i.e.,

$$v^{3/2}\vec{\alpha}\cdot\vec{p}\,v^{-3/2} = \vec{\alpha}\cdot\vec{p} + v^{3/2}\vec{\alpha}\cdot\left(\vec{p}\,v^{-3/2}\right) = \vec{\alpha}\cdot\vec{p} - v^{3/2}\frac{3}{2}\frac{\vec{\alpha}\cdot(\vec{p}\,v)}{v^{5/2}} = \vec{\alpha}\cdot\vec{p} - \frac{3}{2}\frac{\vec{\alpha}\cdot(\vec{p}\,v)}{v}\,. \tag{8.81}$$

It then follows that the scaled Hamiltonian is

$$\begin{aligned} H' &= \frac{w}{v}\vec{\alpha}\cdot\vec{p} - \frac{3}{2}\frac{w}{v}\frac{\vec{\alpha}\cdot(\vec{p}\,v)}{v} + \frac{w}{v}\frac{\vec{\alpha}\cdot(\vec{p}\,v)}{v} + \frac{\vec{\alpha}\cdot(\vec{p}\,w)}{2v} + \beta m\,w \\ &= \frac{w}{v}\vec{\alpha}\cdot\vec{p} - \frac{1}{2}w\frac{\vec{\alpha}\cdot(\vec{p}\,v)}{v^2} + \frac{1}{2}\frac{1}{v}\vec{\alpha}\cdot(\vec{p}\,w) + \beta m\,w \\ &= \frac{w}{v}\vec{\alpha}\cdot\vec{p} + \frac{1}{2}w\,\vec{\alpha}\cdot\left(\vec{p}\frac{1}{v}\right) + \vec{\alpha}\cdot(\vec{p}\,w) + \beta m\,w \\ &= \frac{1}{2}\frac{w}{v}\vec{\alpha}\cdot\vec{p} + \left(\frac{1}{2}\frac{w}{v}\vec{\alpha}\cdot\vec{p} + \frac{1}{2}\vec{\alpha}\cdot\left(\vec{p}\frac{w}{v}\right)\right) + \beta m\,w \\ &= \frac{1}{2}\mathcal{F}\,\vec{\alpha}\cdot\vec{p} + \frac{1}{2}\vec{\alpha}\cdot\vec{p}\,\mathcal{F} + \beta m\,w = \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p},\mathcal{F}\right\} + \beta m\,w\,, \end{aligned} \tag{8.82}$$

where we defined $\mathcal{F} = w/v$. We now have a general Dirac Hamiltonian in a useful form, and we need only substitute for v and w. In the case of the Dirac–Einstein–Schwarzschild we use the Eddington parameterization (see references [92,94,99–101]) of the Schwarzschild metric, which is the isotropic form. Approximating for a small Schwarzschild radius ($r_{\rm s} = 2\,M\,G/c$, where $c = 1$ in our coordinate system), we have

$$w \approx 1 - \frac{r_{\rm s}}{2r}\,, \quad v \approx 1 + \frac{r_{\rm s}}{2r}\,, \quad \mathcal{F} = \frac{w}{v} \approx 1 - \frac{r_{\rm s}}{r}\,. \tag{8.83}$$

Then to the first order in the Schwarzschild radius ($r_{\rm s}$), the Dirac–Einstein–Schwarzschild Hamiltonian is thus found to be

$$H_{\rm DS} \approx \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p},\left(1-\frac{r_{\rm s}}{r}\right)\right\} + \beta\,m\,\left(1-\frac{r_{\rm s}}{2r}\right)\,. \tag{8.84}$$

Similarly the functions for a non-rotating, accelerating frame are given as (see [94, 99,101,102])

$$w = 1\,, \qquad v = 1 + \vec{a}\cdot\vec{r}\,, \tag{8.85}$$

where $\vec{a}$ is the acceleration of the frame. Using these equations with (8.82), the Dirac Hamiltonian for a non-rotating non-inertial reference frame is found to be

$$H_{\mathrm{NR}} = \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, (1-\vec{a}\cdot\vec{r})\right\} + \beta\, m\ (1-\vec{a}\cdot\vec{r})\ . \tag{8.86}$$

In reference [103], Mashhoon showed that for a Hamiltonian H in a non-rotating frame, the same system as viewed by a rotating observer is given by

$$H' = U^{-1}\, H\, U - \vec{\omega}\cdot\vec{J}, \tag{8.87}$$

where $\vec{J} = \vec{L} + \frac{1}{2}\vec{\Sigma}$ is the total angular momentum, and

$$U = \exp\left(\mathrm{i}\,\omega\cdot\vec{J}\right)\ . \tag{8.88}$$

We can show that the operator $\vec{\alpha}\cdot\vec{p}$ commutes with the operator $\vec{\omega}\cdot\vec{J}$, i.e.,

$$\begin{aligned}
[\vec{\alpha}\cdot\vec{p},\vec{\omega}\cdot\vec{J}] = \left[\vec{\alpha}\cdot\vec{p},\vec{\omega}\cdot\left(\vec{L}+\frac{1}{2}\vec{\Sigma}\right)\right] &= [\vec{\alpha}\cdot\vec{p},\vec{\omega}\cdot\vec{L}] + \frac{1}{2}[\vec{\alpha}\cdot\vec{p},\vec{\omega}\cdot\vec{\Sigma}] \\
&= \alpha^L\epsilon^{IJK}\omega^I(p^L r^J)p^K + \frac{1}{2}\gamma^5\alpha^I\alpha^J(p^I\omega^J - p^J\omega^I) \\
&= \alpha^L\epsilon^{IJK}\omega^I(-\mathrm{i}\delta^{LJ})p^K + \frac{1}{2}\gamma^5(\cancel{\delta^{IJ}} + \mathrm{i}\sigma^K\epsilon^{IJK})(p^I\omega^J - p^J\omega^I) \\
&= -\,\mathrm{i}\epsilon^{IJK}\omega^I\alpha^J p^K + \frac{\mathrm{i}}{2}\epsilon^{IJK}(\gamma^5\sigma^K)(2p^I\omega^J) \\
&= -\,\mathrm{i}\epsilon^{IJK}\alpha^K p^I\omega^J + \mathrm{i}\epsilon^{IJK}\alpha^K p^I\omega^J = 0\,.
\end{aligned} \tag{8.89}$$

Thus $U^{-1}\, H_{\mathrm{NR}}\, U = H_{\mathrm{NR}}$, and the Dirac Hamiltonian for a rotating non–inertial reference frame is found to be

$$H_{\mathrm{NF}} = \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, (1-\vec{a}\cdot\vec{r})\right\} + \beta\, m\ (1-\vec{a}\cdot\vec{r}) - \vec{\omega}\cdot\left(\vec{L}+\frac{1}{2}\vec{\Sigma}\right)\ , \tag{8.90}$$

which we recognize as H_{NR} plus the Mashhoon term. The Hamiltonian in equation (8.90) is explicitly Hermitian, and no further scaling factor for the wave–function is required. With these results in hand, we are ready to investigate a number of Foldy–Wouthuysen transformations.

9. DIRAC EQUATIONS AND FOLDY–WOUTHUYSEN TRANSFORMS

9.1. ORIENTATION

In 1950, Foldy and Wouthuysen discovered a transformation which decouples relativistic quantum Hamiltonians into particle and anti-particle components [22]. This transformation can be exact for the free particle, while for other Hamiltonians we must use a well defined iterative process. While this process is useful in determining the effects of a given Dirac Hamiltonian in the non-relativistic limit, it does have drawbacks. Chief among these is that the process of the Foldy-Wouthuysen transformation can be rather complicated, especially when higher orders of precision are desired. Despite this drawback, we will use standard Foldy-Wouthuysen transformation on a number of Hamiltonians, ranging from the simplest, well known Hamiltonians, such as the free particle, to less well known Hamiltonians, such as the transformation for the Dirac Hamiltonians in a non-inertial reference frame. Some of the results below have been discussed in references [87,88], while detailed calculations are described in the following. For an additional application of the Foldy–Wouthuysen transformation, in which a linear superposition of two confining potentials are added to the Dirac–Coulomb Hamiltonian, see references [89].

9.2. FREE PARTICLE

For the free particle, we will discuss two alternative ways of performing the Foldy-Wouthuysen transformation. The first way, the exact transformation, has been described in [23], chapter 3, which will lend some context as to how the transformation works. This will be followed by the iterative procedure of the process, which will demonstrate how the Foldy–Wouthuysen transformation works in general. We shall see that the results given by the general transformation are in fact an approximation given by the exact transformation.

For the exact Fold-Wouthuysen transformation, we are looking to completely eliminate the odd operators from the Dirac Hamiltonian of the free particle, which in its unrotated form is given as

$$H_{\mathrm{F}} = \vec{\alpha} \cdot \vec{p} + \beta\, m\,. \tag{9.1}$$

To do this we rotate into H' using a unitary transform $U = \exp[\mathrm{i}\, S]$, i.e.,

$$H' = U\, H_{\mathrm{F}}\, U^{+} = \mathrm{e}^{\mathrm{i}\, S}\, H_{\mathrm{F}}\, \mathrm{e}^{-\mathrm{i}\, S}\,, \tag{9.2}$$

where S is Hermitian. For the exact transformation we try the Hermitian operator $S = -\mathrm{i}\,\beta\,\vec{\alpha}\cdot\vec{p}\,\theta$, thus

$$\begin{aligned} U = \mathrm{e}^{\beta\,\vec{\alpha}\cdot\vec{p}\,\theta} &= \sum_{n=0}^{\infty} \frac{(\beta\,\vec{\alpha}\cdot\vec{p}\,\theta)^n}{n!} = \sum_{m=0}^{\infty} \frac{(\beta\,\vec{\alpha}\cdot\vec{p}\,\theta)^{2m}}{(2m)!} + \sum_{m=0}^{\infty} \frac{(\beta\,\vec{\alpha}\cdot\vec{p}\,\theta)^{2m+1}}{(2m+1)!} \\ &= \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!}(p\,\theta)^{2m} + \beta\frac{\vec{\alpha}\cdot\vec{p}}{p}\sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!}(p\,\theta)^{2m+1} = \cos p\,\theta + \beta\frac{\vec{\alpha}\cdot\vec{p}}{p}\sin p\,\theta\,, \end{aligned} \tag{9.3}$$

where $p = |\vec{p}|$. To do this calculation, we used $(\beta\,\vec{\alpha}\cdot\vec{p})^2 = -\vec{p}^{\,2}$. We can now apply this transformation to H,

$$\begin{aligned} H' &= U\, H\, U^{+} = \mathrm{e}^{\beta\,\vec{\alpha}\cdot\vec{p}\,\theta}(\vec{\alpha}\cdot\vec{p} + \beta m)\mathrm{e}^{-\beta\,\vec{\alpha}\cdot\vec{p}\,\theta} \\ &= \left(\cos p\,\theta + \beta\frac{\vec{\alpha}\cdot\vec{p}}{p}\sin p\,\theta\right)(\vec{\alpha}\cdot\vec{p} + \beta m)\left(\cos p\,\theta - \beta\frac{\vec{\alpha}\cdot\vec{p}}{p}\sin p\,\theta\right) \\ &= (\vec{\alpha}\cdot\vec{p} + \beta m)\left(\cos p\,\theta - \beta\frac{\vec{\alpha}\cdot\vec{p}}{p}\sin p\,\theta\right)^2 = (\vec{\alpha}\cdot\vec{p} + \beta m)\mathrm{e}^{-2\,\beta\,\vec{\alpha}\cdot\vec{p}\,\theta} \end{aligned}$$

$$
\begin{aligned}
&= (\vec{\alpha}\cdot\vec{p} + \beta m)\left(\cos 2p\,\theta - \beta\frac{\vec{\alpha}\cdot\vec{p}}{p}\sin 2p\,\theta\right) \\
&= \vec{\alpha}\cdot\vec{p}\left(\cos 2p\,\theta - \frac{m}{p}\sin 2p\,\theta\right) + \beta\left(m\,\cos 2p\,\theta + p\,\sin 2p\,\theta\right) \\
&= \vec{\alpha}\cdot\vec{p}\,\cos 2p\,\theta\left(1 - \frac{m}{p}\tan 2p\,\theta\right) + \beta\,\cos 2p\,\theta\left(m + p\,\tan 2p\,\theta\right)\,. \qquad (9.4)
\end{aligned}
$$

Then to eliminate the odd part, it becomes necessary to choose θ such that $\tan 2p\,\theta = p/m$. As illustrated by figure 9.1, $\cos 2p\,\theta = m/\sqrt{\vec{p}^2 + m^2}$. Using this information we can then finish solving for H',

$$
H' = \beta\,\frac{m}{\sqrt{\vec{p}^2+m^2}}\left(m + p\,\frac{p}{m}\right) = \beta\frac{\vec{p}^2+m^2}{\sqrt{\vec{p}^2+m^2}} = \beta\sqrt{\vec{p}^2+m^2}\,. \qquad (9.5)
$$

Now that we have exactly solved the Foldy-Wouthuysen transformation for the free particle, we will move on to the general transformation. For the general Foldy-Wouthuysen transformation we use the approximation

$$
H' = \mathrm{e}^{\mathrm{i}S}\,H\,\mathrm{e}^{-\mathrm{i}S} \approx H + \mathrm{i}\,[S,H] + \frac{(\mathrm{i})^2}{2!}[S,[S,H]] + \dots\,, \qquad (9.6)
$$

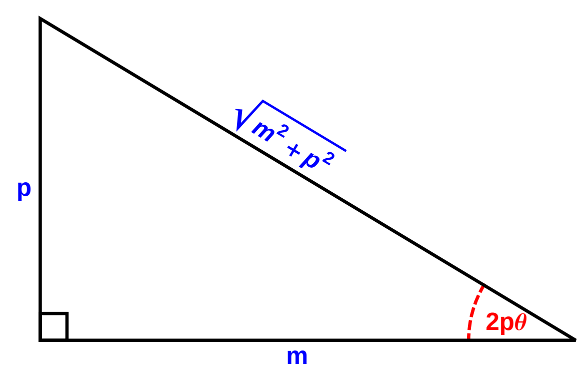


Figure 9.1: This triangle represents our choice of setting $\tan 2p\,\theta = p/m$, in that the side opposite the angle is of length p, while the adjacent side is of length m. Thus the length of the hypotenuse is $\sqrt{m^2+\vec{p}^2}$. This construction is used to aid us in determining the value of the trigonometric functions in the exact Foldy-Wouthuysen transformation of the free particle.

where

$$S = -\frac{\mathrm{i}\,\beta\,\mathcal{O}}{2m}\,, \tag{9.7}$$

and $\mathcal{O}$ is the odd term in the Hamiltonian. As we know the Hamiltonian for the free particle is

$$H_{\mathrm{F}} = \vec{\alpha}\cdot\vec{p} + \beta\,m = \mathcal{O} + \beta\,m\,. \tag{9.8}$$

We are going to keep everything to the order $(\vec{\alpha}\cdot\vec{p})^3$, and ignore terms of higher order. To perform the transformation we use the fact that $\{\beta, \mathcal{O}\} = 0$ (i.e., $\beta\mathcal{O} = -\mathcal{O}\beta$), and $\mathcal{O}^2 = \vec{p}^{\,2}$. We begin by solving for the single commutator

$$\begin{aligned}
[S, H_{\mathrm{F}}] &= -\frac{\mathrm{i}}{2m}[\beta\mathcal{O}, \mathcal{O} + \beta\,m] = -\frac{\mathrm{i}}{2m}\left(\beta\mathcal{O}\mathcal{O} - \mathcal{O}\beta\mathcal{O} + \beta\mathcal{O}\beta m - \beta m\beta\mathcal{O}\right)\\
&= -\frac{\mathrm{i}}{2m}(2\beta\mathcal{O}^2 - 2\beta^2\mathcal{O}m) = -\mathrm{i}\left(\beta\frac{\vec{p}^{\,2}}{m} - \mathcal{O}\right)\,,
\end{aligned} \tag{9.9}$$

followed by the double commutator,

$$\begin{aligned}
[S, [S, H_{\mathrm{F}}]] &= -\frac{\mathrm{i}}{2m}(-\mathrm{i})\left[\beta\mathcal{O}, \beta\frac{\mathcal{O}^2}{m} - \mathcal{O}\right]\\
&= -\frac{1}{2m}\left(\frac{1}{m}\left(\beta\mathcal{O}\beta\mathcal{O}^2 - \beta\mathcal{O}^2\beta\mathcal{O}\right) - \beta\mathcal{O}\mathcal{O} + \mathcal{O}\beta\mathcal{O}\right)\\
&= -\frac{1}{2m}\left(-\frac{2\mathcal{O}^3}{m} - 2\beta\mathcal{O}^2\right) = \frac{\mathcal{O}^3}{m^2} + \beta\frac{\vec{p}^{\,2}}{m}\,,
\end{aligned} \tag{9.10}$$

and finally we calculate the triple commutator, which is the final commutator that will yield new terms to our desired precision,

$$\begin{aligned}
[S, [S, [S, H_{\mathrm{F}}]]] &= -\frac{\mathrm{i}}{2m}\left[\beta\mathcal{O}, \frac{\mathcal{O}^3}{m^2} + \beta\frac{\mathcal{O}^2}{m}\right]\\
&= -\frac{\mathrm{i}}{2m}\left(\frac{1}{m^2}\left(\beta\mathcal{O}^4 - \mathcal{O}^3\beta\mathcal{O}\right) + \frac{1}{m}\left(\beta\mathcal{O}\beta\mathcal{O}^2 - \beta\mathcal{O}^2\beta\mathcal{O}\right)\right)\\
&= -\frac{\mathrm{i}}{2m}\left(\frac{2\beta\mathcal{O}^4}{m^2} - \frac{2\mathcal{O}^3}{m}\right) = \mathrm{i}\left(-\cancel{\frac{\beta\mathcal{O}^4}{m^3}} + \frac{\mathcal{O}^3}{m^2}\right) = \mathrm{i}\frac{\mathcal{O}^3}{m^2}\,,
\end{aligned} \tag{9.11}$$

where we crossed out the term that is of higher order in $(\vec{\alpha}\cdot\vec{p})$ than we are interested in. We then find that after the first transformation the Hamiltonian for a free particle is

$$H' = \beta\, m + \beta\frac{\vec{p}^{\,2}}{2m} - \frac{\mathcal{O}^3}{2m^2} + \frac{1}{6}\frac{\mathcal{O}^3}{m^2} = \beta\left(m + \frac{\vec{p}^{\,2}}{2m}\right) - \frac{\mathcal{O}^3}{3m^2}\,, \tag{9.12}$$

to our desired precision. Notice that there is still an odd term in the rotated Hamiltonian, meaning the particles and antiparticle degrees of freedom are not actually decoupled. To correct this we perform the rotation again, this time setting

$$\mathcal{O}' = -\frac{\mathcal{O}^3}{3m^2} \quad \text{and} \quad S' = -\frac{\mathrm{i}\beta\mathcal{O}'}{2m}\,. \tag{9.13}$$

We begin by calculating $[S', H']$,

$$\begin{aligned}
[S', H'] = -&\frac{\mathrm{i}}{2m}\left[\beta\mathcal{O}', \beta\left(m + \frac{\vec{p}^{\,2}}{2m}\right) - \mathcal{O}'\right] \\
= -&\frac{\mathrm{i}}{2m}\left((\beta\mathcal{O}'\beta - \beta\beta\mathcal{O}')\left(m + \frac{\vec{p}^{\,2}}{2m}\right) - (\beta\mathcal{O}'\mathcal{O}' - \mathcal{O}'\beta\mathcal{O}')\right) \\
= -&\frac{\mathrm{i}}{2m}\left(-2\mathcal{O}'\left(m + \frac{\vec{p}^{\,2}}{2m}\right) - 2\beta(\mathcal{O}')^2\right) = -\mathrm{i}\left(-\mathcal{O}' - \mathcal{O}'\frac{\vec{p}^{\,2}}{2m^2} - \beta\frac{(\mathcal{O}')^2}{m}\right) \\
= -&\mathrm{i}\left(\frac{\mathcal{O}^3}{3m^2} + \cancel{\frac{\mathcal{O}^3\vec{p}^{\,2}}{6m^4}} - \beta\cancel{\frac{\mathcal{O}^6}{9m^5}}\right)\,,
\end{aligned} \tag{9.14}$$

again crossing out terms that are outside our scope of accuracy. We now note that all the higher order commutators will be of order $(\vec{\alpha}\cdot\vec{p})^4$ or higher, so we find

$$H_{\mathrm{F}}^{(\mathrm{FW})} = H'' = \beta\left(m + \frac{\vec{p}^{\,2}}{2m}\right) - \frac{\mathcal{O}^3}{3m^2} + \frac{\mathcal{O}^3}{3m^2} = \beta\left(m + \frac{\vec{p}^{\,2}}{2m}\right)\,, \tag{9.15}$$

where we use the superscript "(FW)" to indicate that the Hamiltonian has been full transformed (to our desired order) using the Foldy–Wouthuysen transform. If we then perform a Taylor series expansion on the exact transformation from equation (9.5) to the same order, we find

$$\beta\sqrt{m^2+\vec{p}^{\,2}} = \beta\, m\sqrt{1+\frac{\vec{p}^{\,2}}{m^2}} \approx \beta\left(m+\frac{\vec{p}^{\,2}}{2m}\right). \tag{9.16}$$

The two solutions agree. This tests the methodology of the Foldy-Wouthuysen transformation, and certainly helps to validate the power of the generalized transformation.

9.3. DIRAC-COULOMB HAMILTONIAN

The Dirac-Coulomb Hamiltonian is a well known operator, and is treated in a number of standard works, (including equation (2.91) of reference [60]), and is given as

$$H_{\rm DC} = \vec{\alpha}\cdot\vec{p} + \beta m - \frac{Z\alpha}{r}\,, \tag{9.17}$$

where Z us the nuclear charge number and α is the fine structure constant. As with the free particle, the odd part of this operator is just $\vec{\alpha}\cdot\vec{p}$, thus

$$\mathcal{O} = \vec{\alpha}\cdot\vec{p}\,, \quad \text{and} \quad S = -\frac{{\rm i}\beta\mathcal{O}}{2m}\,, \tag{9.18}$$

and of course,

$$H_{\rm DC} = \mathcal{O} + \beta m - \frac{Z\alpha}{r}\,. \tag{9.19}$$

We can now calculate the series of nested commutators, keeping terms to the order of $(Z\alpha)^4 m$, recalling that for atomic systems $p \sim Z\alpha m$ and $r \sim 1/(Z\alpha m)$ (see [105, 106]). We begin with the single commutator,

$$\begin{aligned}[S, H_{\rm DC}] &= -\frac{\rm i}{2m}\left[\beta\mathcal{O}, \mathcal{O}+\beta m - \frac{Z\alpha}{r}\right] = -\frac{\rm i}{2m}\left(2\beta\mathcal{O}\mathcal{O} - 2m\mathcal{O} - \beta\left[\mathcal{O}, \frac{Z\alpha}{r}\right]\right) \\ &= -{\rm i}\left(\beta\frac{\mathcal{O}^2}{m} - \mathcal{O} - \frac{\beta}{2m}\left[\mathcal{O}, \frac{Z\alpha}{r}\right]\right)\,, \end{aligned} \tag{9.20}$$

followed by the double commutator,

$$
\begin{aligned}
[S,[S,H_{\mathrm{DC}}]] &= -\frac{\mathrm{i}}{2m}(-\mathrm{i})\left[\beta\mathcal{O},\beta\frac{\mathcal{O}^2}{m}-\mathcal{O}-\frac{\beta}{2m}\left[\mathcal{O},\frac{Z\alpha}{r}\right]\right] \\
&= -\frac{1}{2m}\left(-\frac{2\mathcal{O}^3}{m}-2\beta\mathcal{O}^2+\frac{1}{2m}\left[\mathcal{O},\left[\mathcal{O},\frac{Z\alpha}{r}\right]\right]\right) \\
&= \frac{\mathcal{O}^3}{m^2}+\beta\frac{\mathcal{O}^2}{m}-\frac{1}{4m^2}\left[\mathcal{O},\left[\mathcal{O},\frac{Z\alpha}{r}\right]\right] ,
\end{aligned}
\tag{9.21}
$$

then the triple commutator,

$$
\begin{aligned}
[S,[S,[S,H_{\mathrm{DC}}]]] &= -\frac{\mathrm{i}}{2m}\left[\beta\mathcal{O},\frac{\mathcal{O}^3}{m^2}+\beta\frac{\mathcal{O}^2}{m}-\cancel{\frac{1}{4m^2}\left[\mathcal{O},\left[\mathcal{O},\frac{Z\alpha}{r}\right]\right]}\right] \\
&= -\frac{\mathrm{i}}{2m}\left(2\beta\frac{\mathcal{O}^4}{m^2}-\frac{2\mathcal{O}^3}{m}\right) = -\mathrm{i}\left(\beta\frac{\mathcal{O}^4}{m^3}-\frac{\mathcal{O}^3}{m^2}\right) .
\end{aligned}
\tag{9.22}
$$

where we identified the higher order term early, and crossed it out. Finally we calculate the quadruple commutator,

$$
[S,[S,[S,[S,H_{\mathrm{DC}}]]]] = -\frac{\mathrm{i}}{2m}(-\mathrm{i})\left[\beta\mathcal{O},\cancel{\beta\frac{\mathcal{O}^4}{m^3}}-\frac{\mathcal{O}^3}{m^2}\right] = -\frac{1}{2m}\left(-\beta\frac{2\mathcal{O}^4}{m^2}\right) = \beta\frac{\mathcal{O}^4}{m^3} ,
\tag{9.23}
$$

where the canceled term is again outside our scope. From which we find, to our desired order of precision,

$$
\begin{aligned}
H' &= \beta m - \frac{Z\alpha}{r} + \beta\frac{\mathcal{O}^2}{2m} - \frac{\mathcal{O}^3}{3m^2} - \beta\frac{\mathcal{O}^4}{8m^3} - \frac{\beta}{2m}\left[\mathcal{O},\frac{Z\alpha}{r}\right] + \frac{1}{8m^2}\left[\mathcal{O},\left[\mathcal{O},\frac{Z\alpha}{r}\right]\right] \\
&= \beta\left(m+\frac{\mathcal{O}^2}{2m}-\frac{\mathcal{O}^4}{8m^3}\right) - \frac{Z\alpha}{r} + \frac{1}{8m^2}\left[\mathcal{O},\left[\mathcal{O},\frac{Z\alpha}{r}\right]\right] + \mathcal{O}' ,
\end{aligned}
\tag{9.24}
$$

where we have already anticipated the next step, and defined $\mathcal{O}'$. We need to repeat the entire process so we can eliminate all odd terms, to the desired order, i.e.,

$$
\mathcal{O}' = -\frac{\mathcal{O}^3}{3m^2} - \frac{\beta}{2m}\left[\mathcal{O},\frac{Z\alpha}{r}\right] , \quad \text{and} \quad S' = -\frac{\mathrm{i}\mathcal{O}'}{2m} .
\tag{9.25}
$$

By inspection, it is clear that the only term we will get from this iteration is

$$[S', H'] = \mathrm{i}\mathcal{O}' \,. \tag{9.26}$$

Thus the Foldy-Wouthuysen transformation gives us,

$$H_{\mathrm{DC}}^{(\mathrm{FW})} = H'' = \beta\left(m + \frac{\mathcal{O}^2}{2m} - \frac{\mathcal{O}^4}{8m^3}\right) - \frac{Z\alpha}{r} + \frac{Z\alpha}{8m^2}\left[\mathcal{O}, \left[\mathcal{O}, \frac{1}{r}\right]\right] . \tag{9.27}$$

We know that $\mathcal{O}^2 = (\vec{\alpha}\cdot\vec{p})^2 = \vec{p}^{\,2}$, and $\mathcal{O}^4 = \vec{p}^{\,4}$, leaving only the final term in need of simplification, i.e.,

$$\begin{aligned}
\left[\vec{\alpha}\cdot\vec{p}, \left[\vec{\alpha}\cdot\vec{p}, \frac{1}{r}\right]\right] = \left[\vec{\alpha}\cdot\vec{p}, \vec{\alpha}\cdot\vec{p}\frac{1}{r} - \frac{1}{r}\vec{\alpha}\cdot\vec{p}\right] &= \left[\vec{\alpha}\cdot\vec{p}, \vec{\alpha}\cdot\left(\vec{p}\frac{1}{r}\right) + \cancel{\frac{1}{r}\vec{\alpha}\cdot\vec{p}} - \cancel{\frac{1}{r}\vec{\alpha}\cdot\vec{p}}\right] \\
&= \vec{\alpha}\cdot\vec{p}\,\vec{\alpha}\cdot\left(\vec{p}\frac{1}{r}\right) - \vec{\alpha}\cdot\left(\vec{p}\frac{1}{r}\right)\vec{\alpha}\cdot\vec{p} \\
&= \sigma^I p^I \sigma^J\left(p^J\frac{1}{r}\right) - \sigma^I\left(p^I\frac{1}{r}\right)\sigma^J p^J \\
&= \sigma^I\sigma^J\left(p^I p^J\frac{1}{r}\right) + \sigma^I\sigma^J\left(p^J\frac{1}{r}\right)p^I - \sigma^I\sigma^J\left(p^I\frac{1}{r}\right)p^J \\
&= \sigma^I\sigma^J\left(p^I p^J\frac{1}{r}\right) + \frac{\mathrm{i}}{r^3}\sigma^I\sigma^J(r^J p^I - r^I p^J) \,,
\end{aligned} \tag{9.28}$$

where we used

$$\left(\vec{p}\frac{1}{r}\right) = \mathrm{i}\frac{\vec{r}}{r^3} \,, \quad p^I\left(p^J\frac{1}{r}\right) = \left(p^I p^J\frac{1}{r}\right) + \left(p^J\frac{1}{r}\right)p^I \,. \tag{9.29}$$

We will finish by calculating each of the remaining terms individually,

$$\sigma^I\sigma^J\left(p^I p^J\frac{1}{r}\right) = \delta^{IJ}\left(p^I p^J\frac{1}{r}\right) + \cancel{\mathrm{i}\epsilon^{IJK}\sigma^K\left(p^I p^J\frac{1}{r}\right)} = \vec{p}^{\,2}\frac{1}{r} = 4\pi\delta^{(3)}(\vec{r}) \,, \tag{9.30}$$

and

$$
\begin{aligned}
\sigma^I\sigma^J(r^Jp^I - r^Ip^J) &= \cancel{\delta^{IJ}(r^Jp^I - r^Ip^J)} + \mathrm{i}\epsilon^{IJK}\sigma^K(r^Jp^I - r^Ip^J) \\
&= \mathrm{i}\epsilon^{IJK}\sigma^K(-r^Ip^J - r^Ip^J) = -2\mathrm{i}\sigma^k\epsilon^{IJK}r^Ip^J \\
&= -2\mathrm{i}\vec{\Sigma}\cdot(\vec{r}\times\vec{p}) = -2\mathrm{i}\vec{\Sigma}\cdot\vec{L}\,.
\end{aligned}
\tag{9.31}
$$

Putting this all together, we find

$$
\left[\vec{\alpha}\cdot\vec{p},\left[\vec{\alpha}\cdot\vec{p},\frac{1}{r}\right]\right] = 4\pi\delta^{(3)}(\vec{r}) + 2\frac{\vec{\Sigma}\cdot\vec{L}}{r^3}\,. \tag{9.32}
$$

Finally we find that the Foldy-Wouthuysen Dirac-Coulomb Hamiltonian is given as

$$
H_{\mathrm{DC}}^{(\mathrm{FW})} = \beta\left(m + \frac{\vec{p}^{\,2}}{2m} - \frac{\vec{p}^{\,4}}{8m^3}\right) - \frac{Z\alpha}{r} + \frac{Z\alpha\pi}{2m^2}\delta^{(3)}(\vec{r}) + \frac{Z\alpha}{4m^2r^3}\vec{\Sigma}\cdot\vec{L}\,. \tag{9.33}
$$

Here the first term is the familiar corrected term for the free Dirac Hamiltonian. The second term is the Coulomb potential. The next two terms are higher order corrections, the zitterbewegung term and the Thomas precession, respectively.

9.4. DIRAC HAMILTONIAN WITH A SCALAR POTENTIAL

The Dirac Hamiltonian with a scalar potential [88] is given as

$$
H_{\mathrm{SP}} = \vec{\alpha}\cdot\vec{p} + \beta\left(m - \frac{\lambda}{r}\right)\,, \tag{9.34}
$$

where λ is a coupling parameter. To perform the transformation, we set

$$
\mathcal{O} = \vec{\alpha}\cdot\vec{p}, \quad \text{and} \quad S = -\frac{\mathrm{i}\beta\mathcal{O}}{2m}\,, \tag{9.35}
$$

in which case

$$H_{\rm SP} = \mathcal{O} + \beta\left(m - \frac{\lambda}{r}\right) . \tag{9.36}$$

We can now calculate the nested commutators, keeping terms up to the forth order in momenta ($\vec{p}^{\,4}$), the first order in λ, and up to the second order in momenta when multiplied by λ ($\vec{p}^{\,2}\,\lambda$). We begin by calculating the single commutator,

$$\begin{aligned}
[S, H_{\rm SP}] &= -\frac{\rm i}{2m}\left[\beta\mathcal{O}, \mathcal{O} + \beta\left(m - \frac{\lambda}{r}\right)\right] = -\frac{\rm i}{2m}\left(2\beta\mathcal{O}^2 - \left\{\mathcal{O}, m - \frac{\lambda}{r}\right\}\right)\\
&= -\frac{\rm i}{2m}\left(2\beta\mathcal{O}^2 - 2m\mathcal{O} + m\left\{\mathcal{O}, \frac{\lambda}{r}\right\}\right) = -{\rm i}\left(\beta\frac{\mathcal{O}^2}{m} - \mathcal{O} + \frac{1}{2m}\left\{\mathcal{O}, \frac{\lambda}{r}\right\}\right) ,
\end{aligned} \tag{9.37}$$

then the double commutator,

$$\begin{aligned}
[S, [S, H_{\rm SP}]] &= -\frac{\rm i}{2m}(-{\rm i})\left[\beta\mathcal{O}, \beta\frac{\mathcal{O}^2}{m} - \mathcal{O} + \frac{1}{2m}\left\{\mathcal{O}, \frac{\lambda}{r}\right\}\right]\\
&= -\frac{1}{2m}\left(\frac{1}{m}\left(\beta\mathcal{O}\beta\mathcal{O}^2 - \beta\mathcal{O}^2\beta\mathcal{O}\right) - \beta\mathcal{O}\mathcal{O} + \mathcal{O}\beta\mathcal{O}\right.\\
&\quad\left.+\frac{1}{2m}\left(\beta\mathcal{O}\left\{\mathcal{O}, \frac{\lambda}{r}\right\} - \left\{\mathcal{O}, \frac{\lambda}{r}\right\}\beta\mathcal{O}\right)\right)\\
&= -\frac{1}{2m}\left(-\frac{2\mathcal{O}^3}{m} - 2\beta\mathcal{O}^2 + \frac{1}{2m}\beta\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{\lambda}{r}\right\}\right\}\right)\\
&= \frac{\mathcal{O}^3}{m^2} + \beta\frac{\mathcal{O}^2}{m} - \beta\frac{1}{4m^2}\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{\lambda}{r}\right\}\right\} ,
\end{aligned} \tag{9.38}$$

next the triple commutator,

$$\begin{aligned}
[S, [S, [S, H_{\rm SP}]]] &= -\frac{\rm i}{2m}\left[\beta\mathcal{O}, \frac{\mathcal{O}^3}{m^2} + \beta\frac{\mathcal{O}^2}{m} - \beta\frac{1}{4m^2}\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{\lambda}{r}\right\}\right\}\right]\\
&= -\frac{\rm i}{2m}\left(\frac{1}{m^2}\left(\beta\mathcal{O}\mathcal{O}^3 - \mathcal{O}^3\beta\mathcal{O}\right) + \frac{1}{m}\left(\beta\mathcal{O}\beta\mathcal{O}^2 - \beta\mathcal{O}^2\beta\mathcal{O}\right)\right.\\
&\quad\left.-\frac{1}{4m^2}\left(\beta\mathcal{O}\beta\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{\lambda}{r}\right\}\right\} - \beta\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{\lambda}{r}\right\}\right\}\beta\mathcal{O}\right)\right)\\
&= -\frac{\rm i}{2m}\left(\beta\frac{2\mathcal{O}^4}{m^2} - \frac{2\mathcal{O}}{m} + \frac{1}{4m^2}\left\{\mathcal{O}, \left\{\mathcal{O}, \left\{\mathcal{O}, \frac{\lambda}{r}\right\}\right\}\right\}\right)
\end{aligned}$$

$$
\begin{aligned}
&= - \mathrm{i} \left(\beta \frac{\mathcal{O}^4}{m^3} - \frac{\mathcal{O}}{m^2} + \cancel{\frac{1}{8m^3} \left\{ \mathcal{O}, \left\{ \mathcal{O}, \left\{ \mathcal{O}, \frac{\lambda}{r} \right\} \right\} \right\}} \right) \\
&= - \mathrm{i} \left(\beta \frac{\mathcal{O}^4}{m^3} - \frac{\mathcal{O}^3}{m^2} \right) , \qquad (9.39)
\end{aligned}
$$

where we cancel the higher order term. Finally the quadruple commutator,

$$
\begin{aligned}
[S,[S,[S,[S,H_{\mathrm{SP}}]]]] &= - \frac{\mathrm{i}}{2m} (-\mathrm{i}) \left[\beta \mathcal{O}, \beta \frac{\mathcal{O}^4}{m^3} - \frac{\mathcal{O}^3}{m^2} \right] \\
&= - \frac{1}{2m} \left(\frac{1}{m^2} \left(\beta \mathcal{O} \beta \mathcal{O}^4 - \beta \mathcal{O}^4 \beta \mathcal{O} \right) - \frac{1}{m^2} \left(\beta \mathcal{O} \mathcal{O}^3 - \mathcal{O}^3 \beta \mathcal{O} \right) \right) \\
&= - \frac{1}{2m} \left(- \frac{2\mathcal{O}^5}{m^3} - \beta \frac{2\mathcal{O}^4}{m^2} \right) = \cancel{\frac{\mathcal{O}^6}{m^4}} + \beta \frac{\mathcal{O}^4}{m^3} , \qquad (9.40)
\end{aligned}
$$

where the canceled term is of high enough order that we can approximate it to zero. From which we find

$$
\begin{aligned}
H' &= \beta \left(m - \frac{\lambda}{r} \right) + \beta \frac{\mathcal{O}^2}{2m} + \frac{1}{2m} \left\{ \mathcal{O}, \frac{\lambda}{r} \right\} - \frac{\mathcal{O}^3}{3m^2} + \beta \frac{1}{8m^2} \left\{ \mathcal{O}, \left\{ \mathcal{O}, \frac{\lambda}{r} \right\} \right\} - \beta \frac{\mathcal{O}^4}{8m^3} \\
&= \beta \left(m + \frac{\mathcal{O}^2}{2m} - \frac{\mathcal{O}^4}{8m^3} - \frac{\lambda}{r} \right) - \frac{\mathcal{O}^3}{3m^2} + \frac{1}{2m} \left\{ \mathcal{O}, \frac{\lambda}{r} \right\} + \beta \frac{1}{8m^2} \left\{ \mathcal{O}, \left\{ \mathcal{O}, \frac{\lambda}{r} \right\} \right\} .
\end{aligned}
\qquad (9.41)
$$

Once again there are odd terms in the Hamiltonian, so we must perform the procedure again, this time

$$
\mathcal{O}' = - \frac{\mathcal{O}^3}{3m^2} + \frac{1}{2m} \left\{ \mathcal{O}, \frac{\lambda}{r} \right\} , \quad \text{and} \quad S' = - \frac{\mathrm{i} \beta \mathcal{O}}{2m} . \qquad (9.42)
$$

By inspection it is clear that the only term of $[S', H']$ that is of low enough order that it will not go to zero is

$$
- \frac{\mathrm{i}}{2m} \left[\beta \mathcal{O}', \beta m \right] = - \frac{\mathrm{i}}{2m} \left(-2\, m\, \mathcal{O}' \right) = \mathrm{i} \mathcal{O}' . \qquad (9.43)
$$

This will clearly cancel out the odd term from H', leaving us with

$$H_{\rm SP}^{(\rm FW)} = H'' = \beta\left(m + \frac{\mathcal{O}^2}{2m} - \frac{\mathcal{O}^4}{8m^3} - \frac{\lambda}{r}\right) + \beta\frac{1}{8m^2}\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{\lambda}{r}\right\}\right\}. \tag{9.44}$$

We now need to calculate

$$\left\{\vec{\alpha}\cdot\vec{p}, \left\{\vec{\alpha}\cdot\vec{p}, \frac{\lambda}{r}\right\}\right\}.$$

To do so we will use the following relations

$$\{A,\{A,B\}\} = \{A^2,B\} + 2ABA\,, \qquad 2ABA = \{A^2,B\} - [A,[A,B]]\,. \tag{9.45}$$

Thus

$$\left\{\vec{\alpha}\cdot\vec{p}, \left\{\vec{\alpha}\cdot\vec{p}, \frac{\lambda}{r}\right\}\right\} = 2\left\{\vec{p}^2, \frac{\lambda}{r}\right\} - \left[\vec{\alpha}\cdot\vec{p}, \left[\vec{\alpha}\cdot\vec{p}, \frac{\lambda}{r}\right]\right], \tag{9.46}$$

We recall the results found in (9.32), and allowing for the constant λ we find,

$$\left\{\vec{\alpha}\cdot\vec{p}, \left\{\vec{\alpha}\cdot\vec{p}, \frac{\lambda}{r}\right\}\right\} = 2\left\{\vec{p}^2, \frac{\lambda}{r}\right\} - 4\pi\lambda\delta^{(3)}(\vec{r}) - 2\frac{\lambda}{r^3}\vec{\Sigma}\cdot\vec{L}\,. \tag{9.47}$$

Applying this result to our Hamiltonian we obtain

$$\begin{aligned} H_{\rm SP}^{(\rm FW)} = \beta\Bigg(m &+ \frac{p^2}{2m} - \frac{p^4}{8m^3} - \frac{\lambda}{r} + \frac{1}{4m^2}\left\{\vec{p}^2, \frac{\lambda}{r}\right\} \\ &- \frac{\pi\lambda}{2m^2}\delta^{(3)}(\vec{r}) - \frac{\lambda}{4m^2 r^3}\vec{\Sigma}\cdot\vec{L}\Bigg)\,. \end{aligned} \tag{9.48}$$

Here the first three terms represent the the transformed free particle. The fourth term is the potential. The fifth term is a kinetic correction. The final two terms are again the zitterbewegung term and a spin orbit coupling term respectively. Notice that there is a β prefactor for all the terms, giving us particle–antiparticle symmetry (i.e., both particles and antiparticles will be affected by the scalar potential in the same way).

9.5. DIRAC HAMILTONIAN WITH A SCALAR CONFINING POTENTIAL

The Dirac Hamiltonian with a scalar confining potential (for slightly more complicated variations on such a Hamiltonian see [89, 107, 108]) is given as

$$H_{\rm LC} = \vec{\alpha}\cdot\vec{p} + \beta\,\left(m + \alpha^2 m^2 r\right)\,. \tag{9.49}$$

where α must be small in order for there to be a physically meaningful non–relativistic limit. Thus to perform a Foldy–Wouthuysen transformation, we must assume that $\alpha^2\, r \sim |\vec{p}|$. To perform the transformation, we set

$$\mathcal{O} = \vec{\alpha}\cdot\vec{p}\,, \quad \text{and} \quad S = -\frac{{\rm i}\beta\mathcal{O}}{2m}\,, \tag{9.50}$$

in which case

$$H_{\rm LC} = \mathcal{O} + \beta\,\left(m + \alpha^2 m^2 r\right)\,. \tag{9.51}$$

We can now calculate the series of nested commutators, keeping terms up to third order in momenta ($(\vec{\alpha}\cdot\vec{p})^3$), bearing in mind that we have assumed $\alpha^2\, r \sim |\vec{p}|$ (note that any canceled out terms are being approximated to zero, unless specified otherwise),

$$\begin{aligned} [S, H_{\rm LC}] =& -\frac{\rm i}{2m}\left[\beta\mathcal{O}, \mathcal{O} + \beta\,\left(m + \alpha^2 m^2 r\right)\right] = -\frac{\rm i}{2m}\left(2\beta\mathcal{O}^2 - \left\{\mathcal{O}, m + \alpha^2 m^2 r\right\}\right) \\ =& -\frac{\rm i}{2m}\left(2\beta\mathcal{O}^2 - 2m\mathcal{O} - \alpha^2 m^2 \left\{\mathcal{O}, r\right\}\right) = -{\rm i}\left(\beta\frac{\mathcal{O}^2}{m} - \mathcal{O} - \frac{\alpha^2 m}{2}\left\{\mathcal{O}, r\right\}\right)\,, \end{aligned} \tag{9.52}$$

then the double commutator,

$$\begin{aligned} [S, [S, H_{\rm LC}]] =& -\frac{\rm i}{2m}(-{\rm i})\left[\beta\mathcal{O}, \beta\frac{\mathcal{O}^2}{m} - \mathcal{O} - \frac{\alpha^2 m}{2}\left\{\mathcal{O}, r\right\}\right] \\ =& -\frac{1}{2m}\left(-\frac{2\mathcal{O}^3}{m} - 2\beta\mathcal{O}^2 - \beta\frac{\alpha^2 m}{2}\left\{\mathcal{O}, \left\{\mathcal{O}, r\right\}\right\}\right) \end{aligned}$$

$$= \frac{\mathcal{O}^3}{m^2} + \beta\frac{\mathcal{O}^2}{m} + \beta\frac{\alpha^2}{4}\left\{\mathcal{O},\{\mathcal{O},r\}\right\}\,, \tag{9.53}$$

and finally the triple commutator,

$$\begin{aligned}
[S,[S,[S,H_{\mathrm{LC}}]]] =& -\frac{\mathrm{i}}{2m}\left[\beta\mathcal{O},\frac{\mathcal{O}^3}{m^2} + \beta\frac{\mathcal{O}^2}{m} + \beta\frac{\alpha^2}{4}\left\{\mathcal{O},\{\mathcal{O},r\}\right\}\right] \\
=& -\frac{\mathrm{i}}{2m}\left(\cancel{2\beta\frac{\mathcal{O}^4}{m^2}} - 2\frac{\mathcal{O}^3}{m} - \cancel{\frac{\alpha^2}{4}\left\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},r\}\}\right\}}\right) = \mathrm{i}\frac{\mathcal{O}^3}{m^2}\,.
\end{aligned} \tag{9.54}$$

Then

$$\begin{aligned}
H' =& \,\mathcal{O} + \beta\,\left(m + \alpha^2 m^2 r\right) + \beta\frac{\mathcal{O}^2}{m} - \mathcal{O} - \frac{\alpha^2 m}{2}\left\{\mathcal{O},r\right\} \\
& - \frac{1}{2}\left(\frac{\mathcal{O}^3}{m^2} + \beta\frac{\mathcal{O}^2}{m} + \beta\frac{\alpha^2}{4}\left\{\mathcal{O},\{\mathcal{O},r\}\right\}\right) + \frac{\mathcal{O}^3}{6m^2} \\
=& \,\beta\,\left(m + \alpha^2 m^2 r\right) + \beta\frac{\vec{p}^{\,2}}{2m} - \beta\frac{\alpha^2}{8}\left\{\mathcal{O},\{\mathcal{O},r\}\right\} + \mathcal{O}'\,,
\end{aligned} \tag{9.55}$$

where $\mathcal{O}'$ is a new odd term. When we repeat the process, we will eliminate the odd term, without introducing any new terms of order $(\vec{\alpha}\cdot\vec{p})^3$ or lower, i.e.,

$$H_{\mathrm{LC}}^{(\mathrm{FW})} = \beta\,\left(m + \alpha^2 m^2 r\right) + \beta\frac{\vec{p}^{\,2}}{2m} - \beta\frac{\alpha^2}{8}\left\{\mathcal{O},\{\mathcal{O},r\}\right\}\,, \tag{9.56}$$

We can now calculate the unknown term:

$$\begin{aligned}
\left\{\mathcal{O},\{\mathcal{O},r\}\right\} =& \left\{\vec{\alpha}\cdot\vec{p},\{\vec{\alpha}\cdot\vec{p},r\}\right\} = 2\{\vec{p}^{\,2},r\} - [\vec{\alpha}\cdot\vec{p},[\vec{\alpha}\cdot\vec{p},r]] \\
=& \,2\{\vec{p}^{\,2},r\} - (\delta^{IJ} + \mathrm{i}\epsilon^{IJK}\sigma^K)\left((p^I p^J r) + \frac{\mathrm{i}}{r}(r^I p^J - r^J p^I)\right) \\
=& \,2\{\vec{p}^{\,2},r\} - \left(\delta^{IJ}(p^I p^J r) - \frac{1}{r}\epsilon^{IJK}\sigma^K(r^I p^J - r^J p^I)\right) \\
=& \,2\{\vec{p}^{\,2},r\} - \left(\delta^{IJ}(p^I p^J r) - 2\frac{1}{r}\epsilon^{IJK}\sigma^K r^I p^J\right) \\
=& \,2\{\vec{p}^{\,2},r\} - \vec{p}^{\,2} r + \frac{2}{r}\vec{\Sigma}\cdot(\vec{r}\times\vec{p}) = 2\{\vec{p}^{\,2},r\} + \frac{2}{r} + \frac{2}{r}\vec{\Sigma}\cdot\vec{L}\,.
\end{aligned} \tag{9.57}$$

Thus

$$H_{\rm LC}^{\rm (FW)} = \beta\,\left(m + \alpha^2 m^2 r\right) + \beta\frac{\vec{p}^{\,2}}{2m} - \beta\frac{\alpha^2}{4}\{\vec{p}^{\,2}, r\} - \beta\frac{\alpha^2}{4\,r} - \beta\frac{\alpha^2}{4\,r}\vec{\Sigma}\cdot\vec{L}\,. \qquad (9.58)$$

While the resulting Hamiltonian is not written such that they appear together, the Hamiltonian does contain the transformed free particle equation. This is contained in the first two terms, along with the added linear potential. We again find a kinetic correction term as well as a spin orbit coupling term. Again we note that all the terms have a β prefactor, again preserving the particle–antiparticle symmetry of the system.

9.6. DIRAC–EINSTEIN–SCHWARZSCHILD HAMILTONIAN

As we found in chapter 8, the Dirac–Einstein–Schwarzschild Hamiltonian is given as (see (8.84))

$$H_{\rm DS} = \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \left(1 - \frac{r_{\rm s}}{r}\right)\right\} + \beta\, m\left(1 - \frac{r_{\rm s}}{2r}\right)\,. \qquad (9.59)$$

To perform the transformation, we set

$$\mathcal{O} = \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \left(1 - \frac{r_{\rm s}}{r}\right)\right\}\,, \quad \text{and} \quad S = -\frac{{\rm i}\beta\mathcal{O}}{2m}\,, \qquad (9.60)$$

in which case

$$H_{\rm DS} = \mathcal{O} + \beta\, m\left(1 - \frac{r_{\rm s}}{2r}\right)\,. \qquad (9.61)$$

We can now calculate the nested commutators, keeping terms up to the forth order in momentum, the first order in $r_{\rm s}$, and the second order in momentum when multiplied by $r_{\rm s}$. Note that all canceled out terms are of sufficiently high order that we can approximate them to be zero. We begin with the single commutator,

$$
\begin{aligned}
[S, H_{\mathrm{DS}}] &= -\frac{\mathrm{i}}{2m}\left[\beta\mathcal{O}, \mathcal{O} + \beta\, m\left(1 - \frac{r_{\mathrm{s}}}{2r}\right)\right] \\
&= -\frac{\mathrm{i}}{2m}\left(\beta\mathcal{O}\mathcal{O} - \mathcal{O}\beta\mathcal{O} + \beta\mathcal{O}\beta\, m\left(1 - \frac{r_{\mathrm{s}}}{2r}\right) - \beta\, m\left(1 - \frac{r_{\mathrm{s}}}{2r}\right)\beta\mathcal{O}\right) \\
&= -\frac{\mathrm{i}}{2m}\left(2\beta\mathcal{O}^2 - m\left\{\mathcal{O}, 1 - \frac{r_{\mathrm{s}}}{2r}\right\}\right) \\
&= -\frac{\mathrm{i}}{2m}\left(2\beta\mathcal{O}^2 - m\left\{\mathcal{O}, 1\right\} + m\left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\}\right) \\
&= -\frac{\mathrm{i}}{2m}\left(2\beta\mathcal{O}^2 - 2m\mathcal{O} + m\left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\}\right) = -\mathrm{i}\left(\beta\frac{\mathcal{O}^2}{m} - \mathcal{O} + \frac{1}{2}\left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\}\right) ,
\end{aligned}
\tag{9.62}
$$

followed by the double commutator,

$$
\begin{aligned}
[S, [S, H_{\mathrm{DS}}]] &= -\frac{\mathrm{i}}{2m}(-\mathrm{i})\left[\beta\mathcal{O}, \beta\frac{\mathcal{O}^2}{m} - \mathcal{O} + \frac{1}{2}\left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\}\right] \\
&= -\frac{1}{2m}\left(\frac{1}{m}\left(\beta\mathcal{O}\beta\mathcal{O}^2 - \beta\mathcal{O}^2\beta\mathcal{O}\right) - \beta\mathcal{O}\mathcal{O} + \mathcal{O}\beta\mathcal{O}\right. \\
&\qquad \left. + \frac{1}{2}\left(\beta\mathcal{O}\left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\} - \left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\}\beta\mathcal{O}\right)\right) \\
&= -\frac{1}{2m}\left(-\frac{2\mathcal{O}^3}{m} - 2\beta\mathcal{O}^2 + \frac{1}{2}\beta\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\}\right\}\right) \\
&= \frac{\mathcal{O}^3}{m^2} + \beta\frac{\mathcal{O}^2}{m} - \beta\frac{1}{4m}\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\}\right\} ,
\end{aligned}
\tag{9.63}
$$

then the triple commutator,

$$
\begin{aligned}
[S, [S, [S, H_{\mathrm{DS}}]]] &= -\frac{\mathrm{i}}{2m}\left[\beta\mathcal{O}, \frac{\mathcal{O}^3}{m^2} + \beta\frac{\mathcal{O}^2}{m} - \beta\frac{1}{4m}\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\}\right\}\right] \\
&= -\frac{\mathrm{i}}{2m}\left(\frac{1}{m^2}\left(\beta\mathcal{O}\mathcal{O}^3 - \mathcal{O}^3\beta\mathcal{O}\right) + \frac{1}{m}\left(\beta\mathcal{O}\beta\mathcal{O}^2 - \beta\mathcal{O}^2\beta\mathcal{O}\right)\right. \\
&\qquad \left. - \frac{1}{4m}\left(\beta\mathcal{O}\beta\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\}\right\} - \beta\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\}\right\}\beta\mathcal{O}\right)\right) \\
&= -\frac{\mathrm{i}}{2m}\left(\beta\frac{2\mathcal{O}^4}{m^2} - \frac{2\mathcal{O}}{m} + \cancel{\frac{1}{2m}\left\{\mathcal{O}, \left\{\mathcal{O}, \left\{\mathcal{O}, \frac{r_{\mathrm{s}}}{2r}\right\}\right\}\right\}}\right) \\
&= -\,\mathrm{i}\left(\beta\frac{\mathcal{O}^4}{m^3} - \frac{\mathcal{O}^3}{m^2}\right) ,
\end{aligned}
\tag{9.64}
$$

and finally the quadruple commutator,

$$\begin{aligned}[S,[S,[S,[S,H_{\mathrm{DS}}]]]] =& -\frac{\mathrm{i}}{2m}(-\mathrm{i})\left[\beta\mathcal{O},\beta\frac{\mathcal{O}^4}{m^3}-\frac{\mathcal{O}^3}{m^2}\right] \\ =& -\frac{1}{2m}\left(\frac{1}{m^2}\left(\beta\mathcal{O}\beta\mathcal{O}^4-\beta\mathcal{O}^4\beta\mathcal{O}\right)-\frac{1}{m^2}\left(\beta\mathcal{O}\mathcal{O}^3-\mathcal{O}^3\beta\mathcal{O}\right)\right) \\ =& -\frac{1}{2m}\left(-\frac{2\mathcal{O}^5}{m^3}-\beta\frac{2\mathcal{O}^4}{m^2}\right)=\cancel{\frac{\mathcal{O}^5}{m^4}}+\beta\frac{\mathcal{O}^4}{m^3}\,.\end{aligned} \tag{9.65}$$

Putting it together we find

$$\begin{aligned}H' =&\beta\, m\left(1-\frac{r_{\mathrm{s}}}{2r}\right)+\beta\frac{\mathcal{O}^2}{2m}+\frac{1}{2}\left\{\mathcal{O},\frac{r_{\mathrm{s}}}{2r}\right\}-\frac{\mathcal{O}^3}{3m^2}+\beta\frac{1}{8m}\left\{\mathcal{O},\left\{\mathcal{O},\frac{r_{\mathrm{s}}}{2r}\right\}\right\}-\beta\frac{\mathcal{O}^4}{8m^3} \\ =&\beta\left(m+\frac{\mathcal{O}^2}{2m}-\frac{\mathcal{O}^4}{8m^3}\right)-\beta\frac{m\,r_{\mathrm{s}}}{2r}-\frac{\mathcal{O}^3}{3m^2}+\frac{1}{2}\left\{\mathcal{O},\frac{r_{\mathrm{s}}}{2r}\right\}+\beta\frac{1}{8m}\left\{\mathcal{O},\left\{\mathcal{O},\frac{r_{\mathrm{s}}}{2r}\right\}\right\}\,.\end{aligned} \tag{9.66}$$

Once again there are odd terms in the Hamiltonian, so we must perform the procedure again, this time

$$\mathcal{O}'=-\frac{\mathcal{O}^3}{3m^2}+\frac{1}{2}\left\{\mathcal{O},\frac{r_{\mathrm{s}}}{2r}\right\}\,,\quad\text{and}\quad S'=-\frac{\mathrm{i}\beta\mathcal{O}'}{2m}\,. \tag{9.67}$$

By inspection it is clear that the only term of $[S',H']$ that is of low enough order that it will not go to zero is

$$-\frac{\mathrm{i}}{2m}\left[\beta\mathcal{O}',\beta m\right]=-\frac{\mathrm{i}}{2m}\left(-2\,m\,\mathcal{O}'\right)=\mathrm{i}\mathcal{O}'\,. \tag{9.68}$$

This will clearly cancel out the odd term from H', leaving us with

$$H_{\mathrm{DS}}^{(\mathrm{FW})}=H''=\beta\left(m+\frac{\mathcal{O}^2}{2m}-\frac{\mathcal{O}^4}{8m^3}\right)-\beta\frac{m\,r_{\mathrm{s}}}{2r}+\beta\frac{1}{8m}\left\{\mathcal{O},\left\{\mathcal{O},\frac{r_{\mathrm{s}}}{2r}\right\}\right\}\,. \tag{9.69}$$

We now calculate all terms involving the odd part. We begin by noting

$$\mathcal{O} = \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, 1 - \frac{r_\mathrm{s}}{r}\right\} = \vec{\alpha}\cdot\vec{p} - \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\} . \tag{9.70}$$

Then to the desired order we find the square of the odd part,

$$\begin{aligned}
\mathcal{O}^2 = &\left(\vec{\alpha}\cdot\vec{p} - \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}\right)^2 \\
=&(\vec{\alpha}\cdot\vec{p})^2 - \frac{1}{2}\vec{\alpha}\cdot\vec{p}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\} - \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}\vec{\alpha}\cdot\vec{p} + \cancel{\frac{1}{4}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}^2} \\
=&\vec{p}^{\,2} - \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}\right\} ,
\end{aligned} \tag{9.71}$$

as well as the odd part to the fourth power,

$$\mathcal{O}^4 = \left(\vec{p}^{\,2} - \cancel{\frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}\right\}}\right)^2 = \vec{p}^{\,4} . \tag{9.72}$$

We then calculate the nested anticommutator,

$$\begin{aligned}
\left\{\mathcal{O}, \left\{\mathcal{O}, \frac{r_\mathrm{s}}{2r}\right\}\right\} &= \left\{\vec{\alpha}\cdot\vec{p} - \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \left\{\vec{\alpha}\cdot\vec{p} - \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \frac{r_\mathrm{s}}{2r}\right\}\right\} \\
&= \left\{\vec{\alpha}\cdot\vec{p}, \left\{\vec{\alpha}\cdot\vec{p} - \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \frac{r_\mathrm{s}}{2r}\right\}\right\} \\
&\quad \cancel{- \frac{1}{2}\left\{\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{2r}\right\}, \left\{\vec{\alpha}\cdot\vec{p} - \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \frac{r_\mathrm{s}}{2r}\right\}\right\}} \\
&= \left\{\vec{\alpha}\cdot\vec{p}, \left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{2r}\right\}\right\} - \cancel{\left\{\vec{\alpha}\cdot\vec{p}, \left\{\frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \frac{r_\mathrm{s}}{2r}\right\}\right\}} \\
&= \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}\right\} .
\end{aligned} \tag{9.73}$$

Plugging these results into our equation for the Hamiltonian we find

$$H_\mathrm{DS}^{(\mathrm{FW})} = \beta\left(m + \frac{\vec{p}^{\,2}}{2m} - \frac{\vec{p}^{\,4}}{8m^3}\right) - \beta\frac{m\, r_\mathrm{s}}{2r} - \frac{3\,\beta}{16m}\left\{\vec{\alpha}\cdot\vec{p}, \left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}\right\} . \tag{9.74}$$

The first two terms of this Hamiltonian are immediately recognizable, leaving us needing to simplify only the double anticommutator in the final term. To do so we recall the result found in (9.47), and let $\lambda \to r_{\mathrm{s}}$, giving us the result

$$\left\{\vec{\alpha}\cdot\vec{p}, \left\{\vec{\alpha}\cdot\vec{p}, \frac{r_{\mathrm{s}}}{r}\right\}\right\} = 2\left\{\vec{p}^{\,2}, \frac{r_{\mathrm{s}}}{r}\right\} - 4\pi r_{\mathrm{s}}\delta^{(3)}(\vec{r}) - 2\frac{r_{\mathrm{s}}}{r^3}\vec{\Sigma}\cdot\vec{L}\,. \tag{9.75}$$

Plugging this result into our equation for the rotated Hamiltonian we find

$$\begin{aligned} H_{\mathrm{DS}}^{(\mathrm{FW})} = & \beta\left(m + \frac{\vec{p}^{\,2}}{2m} - \frac{\vec{p}^{\,4}}{8m^3}\right) - \beta\frac{m\,r_{\mathrm{s}}}{2r} \\ & - \beta\frac{3}{8m}\left\{\vec{p}^{\,2}, \frac{r_{\mathrm{s}}}{r}\right\} + \beta\frac{3\pi\,r_{\mathrm{s}}}{4m}\delta^{(3)}(\vec{r}) + \beta\frac{3\,r_{\mathrm{s}}}{8m}\frac{\vec{\Sigma}\cdot\vec{L}}{r^3}\,. \end{aligned} \tag{9.76}$$

Once again, and especially in view of our previous calculations, we find that all of the terms are recognizable. The first term is the transformed equation for a free particle, with its corrections up to the fourth order in momenta. The second term is the gravitational potential, which can be more clearly seen when one considers that $r_{\mathrm{s}} = 2\,G\,M$ $(c = 1)$, where G is the universal gravitational constant, and M is the mass of the gravitational center. With these considerations, the second term becomes $\beta\,G\,M\,m/r$, where the prefactor β ensures that both the particles and antiparticles will be attracted by gravity. The forth term is a kinetic correction to the gravitational coupling. The final two terms are the gravitational zitterbewegung term, and the spin–orbit coupling term, otherwise known as Fokker precession. Furthermore, the Fokker precession term is in full agreement with the classical result found in [115].

9.7. DIRAC HAMILTONIAN IN A ROTATING NON–INERTIAL REFERENCE FRAME

So far all of the Foldy-Wouthuysen transformations we performed consisted of exactly two iterations of the prescribed procedure. The first revealed the correction terms to our desired precision, and the second iteration simply eliminated the remaining odd terms. Based on this, one might be tempted to conclude that working under the precision that we have, all Foldy–Wouthuysen transformations are done in two iterations, provided one wants to calculate the correction terms up to the forth order in momenta. Granted that based solely on the previous examples, this is a quite an assumption, and as should be expected, an incorrect one. Here we will look at the Foldy–Wouthuysen transformation for the Dirac Hamiltonian in a rotating non–inertial frame, and as we will see it takes three iterations. Recall that this Hamiltonian is given as (see (8.90)),

$$H_{\rm NF} = (1 + \vec{a} \cdot \vec{r})\beta m + \frac{1}{2}\{1 + \vec{a} \cdot \vec{r}, \vec{\alpha} \cdot \vec{p}\} - \vec{\omega} \cdot \left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right) , \qquad (9.77)$$

and is valid for a reference frame with a uniform acceleration $\vec{a}$. In addition to the usual constraints, we will keep both $\vec{a}$ and $\vec{\omega}$ to the first order, and keep terms up to the forth order in momenta ($\vec{p}^4$). The Hamiltonian can be rewritten as

$$H_{\rm NF} = (1+\vec{a}\cdot\vec{r})\beta m - \vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right) + \mathcal{O} = \beta m + \beta m\vec{a}\cdot\vec{r} - \vec{\omega}\cdot\vec{L} - \frac{1}{2}\vec{\omega}\cdot\vec{\Sigma} + \mathcal{O}\,, \quad (9.78)$$

where as usual, $\mathcal{O}$ is the odd part of the Hamiltonian. It is clear that

$$\mathcal{O} = \frac{1}{2}\{1 + \vec{a} \cdot \vec{r}, \vec{\alpha} \cdot \vec{p}\} = \vec{\alpha} \cdot \vec{p} + \frac{1}{2}\{\vec{a} \cdot \vec{r}, \vec{\alpha} \cdot \vec{p}\}\,, \quad S = -\frac{\mathrm{i}\,\beta\,\mathcal{O}}{2\,m}\,. \qquad (9.79)$$

We are then ready to perform the transformation, again, we are keeping terms to the fourth order in momenta, and the first order in $\vec{a}$. We will additionally use the fact that $\vec{\alpha}\cdot\vec{p}$ and $\vec{\omega}\cdot\vec{J}$ commute, as shown in (8.89). As in the previous sections, canceled out terms are of high enough order that they can be approximated to zero. We begin with the single commutator

$$\begin{aligned}[S,H_{\rm NF}] &= -\frac{\rm i}{2\,m}\left[\beta\,\mathcal{O}\,,\;\beta m+\beta m\vec{a}\cdot\vec{r}-\vec{\omega}\cdot\vec{L}-\frac{1}{2}\vec{\omega}\cdot\vec{\Sigma}+\mathcal{O}\right]\\ &= -\frac{\rm i}{2\,m}\left(-2m\mathcal{O}-m\{\mathcal{O},\vec{a}\cdot\vec{r}\}-\cancel{\left[\beta\,\mathcal{O}\,,\vec{\omega}\cdot\left(\vec{L}+\frac{1}{2}\vec{\Sigma}\right)\right]}+2\beta\mathcal{O}^2\right)\\ &={\rm i}\left(\mathcal{O}+\frac{1}{2}\{\mathcal{O},\vec{a}\cdot\vec{r}\}-\frac{\beta\,\mathcal{O}^2}{m}\right)\,,\end{aligned}\tag{9.80}$$

then the double commutator,

$$\begin{aligned}[S,[S,H_{\rm NF}]] &= -\frac{{\rm i}^2}{2\,m}\left[\beta\,\mathcal{O}\,,\;\mathcal{O}+\frac{1}{2}\{\mathcal{O},\vec{a}\cdot\vec{r}\}-\frac{\beta\,\mathcal{O}^2}{m}\right]\\ &=\frac{1}{2\,m}\left(2\beta\,\mathcal{O}^2+\frac{1}{2}\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}+\frac{2\,\mathcal{O}^3}{m}\right)\\ &=\frac{\beta}{m}\mathcal{O}^2+\frac{\beta}{4\,m}\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}+\frac{\mathcal{O}^3}{m^2}\,,\end{aligned}\tag{9.81}$$

followed by the triple commutator,

$$\begin{aligned}[S,[S,[S,H_{\rm NF}]]] &= -\frac{\rm i}{2\,m}\left[\beta\,\mathcal{O}\,,\;\frac{\beta\,\mathcal{O}^2}{m}+\frac{\beta}{4\,m}\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}+\frac{\mathcal{O}^3}{m^2}\right]\\ &= -\frac{\rm i}{2\,m}\left(-\frac{2\,\mathcal{O}^3}{m}-\frac{1}{4\,m}\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}\}+\frac{2\,\beta\,\mathcal{O}^4}{m^2}\right)\\ &={\rm i}\left(\frac{\mathcal{O}^3}{m^2}+\frac{1}{8\,m^2}\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}\}-\frac{\beta\,\mathcal{O}^4}{m^3}\right)\,,\end{aligned}\tag{9.82}$$

and finally we calculate the quadruple commutator,

$$\begin{aligned}[S,[S,[S,[S,H_{\rm NF}]]]] &= -\frac{{\rm i}^2}{2\,m}\left[\beta\,\mathcal{O}\,,\;\frac{\mathcal{O}^3}{m^2}+\frac{1}{8\,m^2}\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}\}-\cancel{\frac{\beta\,\mathcal{O}^4}{m^3}}\right]\\ &=\frac{1}{2m}\left(\frac{2\beta\mathcal{O}^4}{m^2}+\frac{\beta}{8\,m^2}\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}\}\}\right)\\ &=\frac{\beta\,\mathcal{O}^4}{m^3}+\frac{\beta}{16\,m^3}\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}\}\}\,.\end{aligned}\tag{9.83}$$

The first iteration of the procedure then leads to the Hamiltonian,

$$
\begin{aligned}
H' =& \beta\, m + \beta\, m\, \vec{a}\cdot\vec{r} - \vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right) + \frac{\beta\,\mathcal{O}^2}{2\,m} - \frac{\beta}{8\,m}\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\} + \frac{\beta\,\mathcal{O}^4}{8\,m^3} \\
&+ \frac{\beta}{384\,m^3}\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}\}\} + \mathcal{O}'\,,
\end{aligned}
\tag{9.84}
$$

where

$$
\mathcal{O}' = -\frac{1}{2}\{\mathcal{O},\vec{a}\cdot\vec{r}\} - \frac{\mathcal{O}^3}{3\,m^2} + \frac{1}{48\,m^2}\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}\}\,. \tag{9.85}
$$

For this iteration we use the operator,

$$
S' = -\frac{\mathrm{i}\,\beta\,\mathcal{O}'}{2\,m}\,. \tag{9.86}
$$

Before we begin calculating all the relevant terms, we note that the only term in $\mathcal{O}'$ which does not contain an acceleration term, $\vec{a}$, (outside of the odd part, $\mathcal{O}$) is of the third order in momenta. As such most of the terms of $[S', H']$ can be ignored. Accordingly, the only terms we explicitly write out will be the (possibly) relevant terms. We begin with the single commutator,

$$
\begin{aligned}
[S', H'] =& -\frac{\mathrm{i}}{2\,m}\left[\beta\,\mathcal{O}'\,,\ \beta\, m + \beta\, m\,\vec{a}\cdot\vec{r} - \vec{\omega}\cdot\left(\vec{L}+\frac{1}{2}\vec{\Sigma}\right) + \frac{\beta\,\mathcal{O}^2}{2\,m} + \mathcal{O}'\right] \\
=& -\frac{\mathrm{i}}{2\,m}\left(-2\,m\,\mathcal{O}' - m\{\mathcal{O}',\vec{a}\cdot\vec{r}\} - \cancel{\beta\left[\mathcal{O}',\vec{\omega}\cdot\left(\vec{L}+\frac{1}{2}\vec{\Sigma}\right)\right]}\right. \\
&\left. -\frac{1}{2\,m}\{\mathcal{O}',\mathcal{O}^2\} + 2\beta\,\mathcal{O}'^2\right) \\
=& \,\mathrm{i}\left(\mathcal{O}' + \frac{1}{2}\{\mathcal{O}',\vec{a}\cdot\vec{r}\} + \frac{1}{4\,m^2}\{\mathcal{O}',\mathcal{O}^2\} - \frac{\beta\,\mathcal{O}'^2}{m}\right)\,,
\end{aligned}
\tag{9.87}
$$

and we follow up with the double commutator,

$$[S',[S',H']] = -\frac{\mathrm{i}^2}{2\,m}\left[\beta\,\mathcal{O}',\mathcal{O}' + \frac{1}{2}\{\mathcal{O}',\vec{a}\cdot\vec{r}\} + \frac{1}{4\,m^2}\{\mathcal{O}',\mathcal{O}^2\} - \frac{\beta\,\mathcal{O}'^2}{m}\right]$$
$$= \frac{1}{2\,m}\left(2\,\beta\,\mathcal{O}'^2 + \cancel{\frac{\beta}{2}\{\mathcal{O}',\{\mathcal{O}',\vec{a}\cdot\vec{r}\}\}} + \cancel{\frac{\beta}{4\,m^2}\{\mathcal{O}',\{\mathcal{O}',\mathcal{O}^2\}\}} + \cancel{\frac{2\,\beta\,\mathcal{O}'^3}{m}}\right)$$
$$= \frac{\beta\,\mathcal{O}'^2}{m}\,. \tag{9.88}$$

Any additional terms will be of a higher order then we are interested in. Thus after two iterations of the Foldy–Wouthuysen transformation, the resulting Hamiltonian reads as

$$H'' = \beta\,m + \beta\,m\,\vec{a}\cdot\vec{r} - \vec{\omega}\cdot\left(\vec{L}+\frac{1}{2}\vec{\Sigma}\right) + \frac{\beta\,\mathcal{O}^2}{2\,m} + \frac{\beta}{8\,m}\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\} - \frac{\beta\,\mathcal{O}^4}{8\,m^3}$$
$$+ \frac{\beta}{384\,m^3}\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}\}\} - \frac{1}{4\,m^2}\{\mathcal{O}',\mathcal{O}^2\} + \frac{\beta\,\mathcal{O}'^2}{2\,m}\,. \tag{9.89}$$

In order to simplify this expression we will express all the terms involving the odd operator $\mathcal{O}'$ in terms of $\mathcal{O}$, ignoring the higher order terms, i.e.,

$$\frac{1}{2}\{\mathcal{O}',\vec{a}\cdot\vec{r}\} = -\frac{1}{2}\left\{\frac{\mathcal{O}^3}{3\,m^2},\vec{a}\cdot\vec{r}\right\} = -\frac{1}{6\,m^3}\left\{\mathcal{O}^3,\vec{a}\cdot\vec{r}\right\}\,, \tag{9.90}$$
$$\frac{1}{4\,m^2}\{\mathcal{O}',\mathcal{O}^2\} = -\frac{1}{4\,m^2}\left\{\frac{1}{2}\{\mathcal{O},\vec{a}\cdot\vec{r}\},\mathcal{O}^2\right\} = -\frac{1}{8\,m^2}\{\{\mathcal{O},\vec{a}\cdot\vec{r}\},\mathcal{O}^2\}\,, \tag{9.91}$$
$$\frac{\beta\,\mathcal{O}'^2}{2m} = \frac{\beta}{2m}\left(\frac{1}{2}\{\mathcal{O},\vec{a}\cdot\vec{r}\}\frac{\mathcal{O}^3}{3\,m^2} + \frac{\mathcal{O}^3}{3\,m^2}\frac{1}{2}\{\mathcal{O},\vec{a}\cdot\vec{r}\}\right) = \frac{\beta}{12\,m^3}\{\mathcal{O}^3,\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}\,. \tag{9.92}$$

As such the twice rotated Hamiltonian becomes,

$$H'' = \beta\,m + \beta\,m\,\vec{a}\cdot\vec{r} - \vec{\omega}\cdot\left(\vec{L}+\frac{1}{2}\vec{\Sigma}\right) + \frac{\beta\,\mathcal{O}^2}{2\,m} - \frac{\beta}{8\,m}\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\} - \frac{\beta\,\mathcal{O}^4}{8\,m^3}$$
$$+ \frac{\beta}{384\,m^3}\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}\}\} + \frac{\beta}{12\,m^3}\{\mathcal{O}^3,\{\mathcal{O},\vec{a}\cdot\vec{r}\}\} + \mathcal{O}''\,, \tag{9.93}$$

where

$$\mathcal{O}'' = \frac{1}{6\,m^3}\{\mathcal{O}^3, \vec{a}\cdot\vec{r}\} + \frac{1}{8\,m^2}\{\{\mathcal{O}, \vec{a}\cdot\vec{r}\}, \mathcal{O}^2\}\,. \tag{9.94}$$

Notice that every term in $\mathcal{O}''$ is at least a first-order term in acceleration, and they are all third-order terms in momenta. As such, it is clear by inspection that the third iteration of the Foldy–Wouthuysen transformation will serve only to eliminate the remaining odd part of the Hamiltonian, leading us to conclude that

$$\begin{aligned} H_{\rm NF}^{\rm (FW)} = {} & \beta\,m + \beta\,m\,\vec{a}\cdot\vec{r} - \vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right) + \frac{\beta\,\mathcal{O}^2}{2\,m} - \frac{\beta}{8\,m}\{\mathcal{O}, \{\mathcal{O}, \vec{a}\cdot\vec{r}\}\} - \frac{\beta\,\mathcal{O}^4}{8\,m^3} \\ & + \frac{\beta}{384\,m^3}\{\mathcal{O}, \{\mathcal{O}, \{\mathcal{O}, \{\mathcal{O}, \vec{a}\cdot\vec{r}\}\}\}\} + \frac{\beta}{12\,m^3}\{\mathcal{O}^3, \{\mathcal{O}, \vec{a}\cdot\vec{r}\}\}\,. \end{aligned} \tag{9.95}$$

We now need to simplify the Hamiltonian by substituting for the odd terms according to (9.79). We begin with $\mathcal{O}^2$,

$$\begin{aligned} \mathcal{O}^2 &= \left(\vec{\alpha}\cdot\vec{p} + \frac{1}{2}\{\vec{a}\cdot\vec{r}, \vec{\alpha}\cdot\vec{p}\}\right)^2 \\ &= (\vec{\alpha}\cdot\vec{p})^2 + \vec{\alpha}\cdot\vec{p}\,\frac{1}{2}\{\vec{a}\cdot\vec{r}, \vec{\alpha}\cdot\vec{p}\} + \frac{1}{2}\{\vec{a}\cdot\vec{r}, \vec{\alpha}\cdot\vec{p}\}\,\vec{\alpha}\cdot\vec{p} + \cancel{\frac{1}{4}\{\vec{a}\cdot\vec{r}, \vec{\alpha}\cdot\vec{p}\}^2} \\ &= \vec{p}^{\,2} + \frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}\}\,, \end{aligned} \tag{9.96}$$

which we can use to calculate $\mathcal{O}^4$,

$$\begin{aligned} \mathcal{O}^4 &= \left(\vec{p}^{\,2} + \frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}\}\right)^2 \\ &= \vec{p}^{\,4} + \frac{1}{2}\left\{\vec{p}^{\,2}, \{\vec{\alpha}\cdot\vec{p}, \{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}\}\right\} + \cancel{\frac{1}{4}\{\vec{\alpha}\cdot\vec{p}, \{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}\}^2} \\ &= \vec{p}^{\,4} + \frac{1}{2}\{\vec{p}^{\,2}, \{\vec{\alpha}\cdot\vec{p}, \{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}\}\}\,. \end{aligned} \tag{9.97}$$

We now need to calculate the double anticommutator in the final term of $\mathcal{O}^2$. To do so we again turn to the relation given in equation (9.45), from which we find

$$\{\vec{\alpha}\cdot\vec{p},\{\vec{\alpha}\cdot\vec{p},\vec{a}\cdot\vec{r}\}\} = 2\,\{\vec{p}^{\,2},\vec{a}\cdot\vec{r}\} - [\vec{\alpha}\cdot\vec{p},[\vec{\alpha}\cdot\vec{p},\vec{a}\cdot\vec{r}]]\,, \tag{9.98}$$

leaving us with the double commutator, which we calculate using using the fact that $\vec{\alpha}\cdot(\vec{p}\,\vec{a}\cdot\vec{r}) = -\mathrm{i}\,\vec{\alpha}\cdot\vec{a}$,

$$\begin{aligned}
[\vec{\alpha}\cdot\vec{p},[\vec{\alpha}\cdot\vec{p},\vec{a}\cdot\vec{r}]] &= [\vec{\alpha}\cdot\vec{p},(\vec{\alpha}\cdot\vec{p}\,\vec{a}\cdot\vec{r} - \vec{a}\cdot\vec{r}\,\vec{a}\cdot\vec{r})] = [\vec{\alpha}\cdot\vec{p},-\mathrm{i}\,\vec{\alpha}\cdot\vec{a}] \\
&= -\mathrm{i}\left(\alpha^I p^I \alpha^J a^J - \alpha^I a^I \alpha^J p^J\right) = -\mathrm{i}\alpha^I\alpha^J\left(a^J p^I - a^I p^J\right) \\
&= -\mathrm{i}(\cancel{\delta^{IJ}} + \mathrm{i}\sigma^K\epsilon^{IJK})\left(a^J p^I - a^I p^J\right) = \sigma^k\epsilon^{IJK}\left(-a^I p^J - a^I p^J\right) \\
&= -2\epsilon^{KIJ}\sigma^K a^I p^J = -2\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\,.
\end{aligned} \tag{9.99}$$

The Kronecker symbols δ^{IJ} vanish due to the arithmetic involved. Thus,

$$\{\vec{\alpha}\cdot\vec{p},\{\vec{\alpha}\cdot\vec{p},\vec{a}\cdot\vec{r}\}\} = 2\{\vec{p}^{\,2},\vec{a}\cdot\vec{r}\} + 2\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\,, \tag{9.100}$$

and we plug this result into our equation for $\mathcal{O}^2$, leaving

$$\mathcal{O}^2 = \vec{p}^{\,2} + \{\vec{p}^{\,2},\vec{a}\cdot\vec{r}\} + \vec{\Sigma}\cdot(\vec{a}\times\vec{p})\,. \tag{9.101}$$

Furthermore, these results can be used to simplify $\mathcal{O}^4$, i.e.,

$$\begin{aligned}
\{\vec{p}^{\,2},\{\vec{\alpha}\cdot\vec{p},\{\vec{\alpha}\cdot\vec{p},\vec{a}\cdot\vec{r}\}\}\} &= 2\{\vec{p}^{\,2},\{\vec{p}^{\,2},\vec{a}\cdot\vec{r}\}\} + 2\{\vec{p}^{\,2},\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\} \\
&= 4\{\vec{p}^{\,4},\vec{a}\cdot\vec{r}\} - \cancelto{0}{2[\vec{p}^{\,2},[\vec{p}^{\,2},\vec{a}\cdot\vec{r}]]} + 4\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\vec{p}^{\,2}\,,
\end{aligned} \tag{9.102}$$

and thus we find,

$$\mathcal{O}^4 = \vec{p}^{\,4} + 2\{\vec{p}^{\,4}, \vec{a}\cdot\vec{r}\} + 2\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\,\vec{p}^{\,2}\,. \tag{9.103}$$

We are now left with only the anticommutators to simplify. Since we are still keeping the acceleration to the first order, and the odd term $\mathcal{O}$ which appears in the anticommutators can be taken as $\mathcal{O} = \vec{\alpha}\cdot\vec{p}$ to our desired order, we find that the first double anticommutator was already solved in (9.100). This result, along with the identities,

$$\begin{aligned}
[A,[A,\{A,\{A,B\}\}]] &= [A^2,[A^2,B]]\,,\\
\{A^3,\{A,B\}\} &= 2\{A^4,B\} - \frac{1}{2}[A^2,[A^2,B]] - \frac{1}{2}\{A^2,[A,[A,B]]\}\,,
\end{aligned} \tag{9.104}$$

can then be used to solve the remaining terms, i.e.,

$$\begin{aligned}
\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\{\mathcal{O},\vec{a}\cdot\vec{r}\}\}\}\} &= \{\vec{\alpha}\cdot\vec{p},\{\vec{\alpha}\cdot\vec{p},\{\vec{\alpha}\cdot\vec{p},\{\vec{\alpha}\cdot\vec{p},\vec{a}\cdot\vec{r}\}\}\}\}\\
&= 2\{\vec{p}^{\,2},\{\vec{\alpha}\cdot\vec{p},\{\vec{\alpha}\cdot\vec{p},\vec{a}\cdot\vec{r}\}\}\} - [\vec{\alpha}\cdot\vec{p},[\vec{\alpha}\cdot\vec{p},\{\vec{\alpha}\cdot\vec{p},\{\vec{\alpha}\cdot\vec{p},\vec{a}\cdot\vec{r}\}\}]]\\
&= 2(4\{\vec{p}^{\,4},\vec{a}\cdot\vec{r}\} + 4\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\vec{p}^{\,2}) - \cancelto{0}{[\vec{p}^{\,2},[\vec{p}^{\,2},\vec{a}\cdot\vec{r}]]}\\
&= 8\{\vec{p}^{\,4},\vec{a}\cdot\vec{r}\} + 8\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\,\vec{p}^{\,2}\,,
\end{aligned} \tag{9.105}$$

and

$$\begin{aligned}
\{\mathcal{O}^3,\{\mathcal{O},\vec{a}\cdot\vec{r}\}\} &= 2\{\vec{p}^{\,4},\vec{a}\cdot\vec{r}\} - \frac{1}{2}\cancelto{0}{[\vec{p}^{\,2},[\vec{p}^{\,2},\vec{a}\cdot\vec{r}]]} - \frac{1}{2}\{\vec{p}^{\,2},[\vec{\alpha}\cdot\vec{p},[\vec{\alpha}\cdot\vec{p},\vec{a}\cdot\vec{r}]]\}\\
&= 2\{\vec{p}^{\,4},\vec{a}\cdot\vec{r}\} - \frac{1}{2}\{\vec{p}^{\,2},-2\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\}\\
&= 2\{\vec{p}^{\,4},\vec{a}\cdot\vec{r}\} + 2\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\;\vec{p}^{\,2}\,.
\end{aligned} \tag{9.106}$$

Notice that we have used the fact that $[\vec{p}^{\,2},[\vec{p}^{\,2},\vec{a}\cdot\vec{r}]] = 0$ quite a few times when solving for the terms. This can be easily verified by solving the commentator nested within, i.e.,

$$\begin{aligned}\left[\vec{p}^{\,2}, \vec{a}\cdot\vec{r}\right] &= \vec{p}^{\,2}\,\vec{a}\cdot\vec{r} - \vec{a}\cdot\vec{r}\,\vec{p}^{\,2} = p^I p^I\, a^J r^J - a^I r^I\, p^J p^J \\ &= -\,\mathrm{i}\, p^I\, a^J \delta^{IJ} + p^I\, a^J r^J\, p^I - a^I r^I\, p^J p^J \\ &= -\,\mathrm{i}\, a^I p^I - \mathrm{i}\, a^J \delta^{IJ} p^I + a^J r^J\, p^I p^I - a^I r^I\, p^J p^J = -2\mathrm{i}\,\vec{a}\cdot\vec{p}\,, \end{aligned} \tag{9.107}$$

which of course commutes with $\vec{p}^{\,2}$. To reiterate, $[\vec{p}^{\,2}, \vec{a}\cdot\vec{r}]$ commutes with $\vec{p}^{\,2}$, thus the double commutator is zero. Finally we can plug all of our results into our Hamiltonian, and then simplify, yielding

$$\begin{aligned} H^{(\mathrm{FW})}_{\mathrm{NF}} &= \beta\left(m + \frac{\vec{p}^{\,2}}{2\,m} - \frac{\vec{p}^{\,4}}{8\,m^3} + m\,\vec{a}\cdot\vec{r} + \frac{1}{4\,m}\{\vec{p}^{\,2}, \vec{a}\cdot\vec{r}\} - \frac{1}{16\,m^3}\{\vec{p}^{\,4}, \vec{a}\cdot\vec{r}\}\right) \\ &\quad + \beta\,\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\left(\frac{1}{4\,m} - \frac{\vec{p}^{\,2}}{16\,m^3}\right) - \vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right). \end{aligned} \tag{9.108}$$

It is immediately obvious that while there are familiar terms, such as the free particle terms, there are terms here that are unlike the terms found in our other transformed Hamiltonians. This is due to the fact we are still dealing with a free particle, but it is as observed from a rotating and accelerating frame. We also notice that the Mashhoon term is unaffected by the transformation up to the fourth order in momenta. We also note that all the terms, save for the final Mashhoon term, are decorated with the prefactor β. We then have particle–antiparticle symmetry.

9.8. GRAVITATIONALLY COUPLED TRANSITION CURRENT

Here we derive the non–relativistic corrections to the gravitationally coupled transition current J^I. The nonrelativistic corrections to the transition current for emitted photons can be used to calculate the corrections to the interactions when coupling a system to photons, both real and virtual. The coupling to real photons gives rise to relativistic corrections in atomic physics, while the virtual photons result in

QED corrections (for more in depth discussions see [23,109–111]). One calculates the transition current by coupling the Dirac–Einstein–Schwarzschild Hamiltonian (9.76) to an external electromagnetic field, utilizing the usual replacement $\vec{p} \to \vec{p} - e\,\vec{A}$, where $\vec{A}$ is the vector potential. The interaction is known to be $H_{int} = -\vec{J}\cdot\vec{A}$. Thus the gravitationally coupled transition current is found to be

$$\begin{aligned} J^I &= \frac{1}{2}\left\{1 - \frac{r_{\mathrm{s}}}{r}, \alpha^I\, \mathrm{e}^{\mathrm{i}\vec{k}\cdot\vec{r}}\right\} \\ &\approx \left(1 - \frac{r_{\mathrm{s}}}{r}\right)\alpha^I + \left(1 - \frac{r_{\mathrm{s}}}{r}\right)\alpha^I\left(\mathrm{i}\vec{k}\cdot\vec{r}\right) - \frac{1}{2}\left(1 - \frac{r_{\mathrm{s}}}{r}\right)\alpha^I\left(\vec{k}\cdot\vec{r}\right)^2 . \end{aligned} \tag{9.109}$$

By defining

$$J_0^I \equiv \left(1 - \frac{r_{\mathrm{s}}}{r}\right)\alpha^I , \quad J_1^I \equiv \left(1 - \frac{r_{\mathrm{s}}}{r}\right)\alpha^I\left(\mathrm{i}\vec{k}\cdot\vec{r}\right) , \tag{9.110a}$$

$$J_2^I \equiv -\frac{1}{2}\left(1 - \frac{r_{\mathrm{s}}}{r}\right)\alpha^I\left(\vec{k}\cdot\vec{r}\right)^2 , \tag{9.110b}$$

we can calculate the approximate transformation of J^I in three parts ($J^I \approx J_0^I + J_1^I + J_2^I$). For convenience we recall that the first rotation of the transform is

$$U = \mathrm{e}^{\mathrm{i}\,S} , \quad S = -\mathrm{i}\frac{\beta\,\mathcal{O}}{2m} , \quad \mathcal{O} = \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \left(1 - \frac{r_{\mathrm{s}}}{r}\right)\right\} = \vec{\alpha}\cdot\vec{p} - \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_{\mathrm{s}}}{r}\right\} . \tag{9.111}$$

Before continuing on we take note of a few relations

$$\vec{\alpha}\cdot\vec{v}\,\alpha^I = v^I - \mathrm{i}\,(\vec{v}\times\vec{\sigma})^I = v^I + \mathrm{i}\,(\vec{\sigma}\times\vec{v})^I , \tag{9.112a}$$

$$\alpha^I\,\vec{\alpha}\cdot\vec{v} = v^I + \mathrm{i}\,(\vec{v}\times\vec{\sigma})^I = v^I - \mathrm{i}\,(\vec{\sigma}\times\vec{v})^I , \tag{9.112b}$$

$$\{\vec{\alpha}\cdot\vec{v}, \alpha^I\} = \vec{\alpha}\cdot\vec{v}\,\alpha^I + \alpha^I\,\vec{\alpha}\cdot\vec{v} = 2\,v^I , \tag{9.112c}$$

where $\vec{v}$ is an arbitrary vector. Armed with these equations, we are prepared to begin the first transformation. As with the Dirac–Einstein–Schwarzschild Hamiltonian, we are keeping terms up to the first order in gravity, to the fourth order in momenta, and to the second order in momenta when combined with gravity. Additionally, we are assuming that the exchanged photons are soft, i.e., $|\vec{k}| \sim \vec{p}^{\,2}/m$. As before, the canceled out terms are approximated to zero. We begin by calculating the single commutators. First J_0^I

$$
\begin{aligned}
[S, J_0^I] =& -\frac{\mathrm{i}}{2m}\left[\beta\,\vec{\alpha}\cdot\vec{p} - \frac{\beta}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \alpha^I - \alpha^I\,\frac{r_\mathrm{s}}{r}\right] \\
=& -\frac{\mathrm{i}\,\beta}{2m}\left(\left\{\vec{\alpha}\cdot\vec{p}, \alpha^I\right\} - \left\{\vec{\alpha}\cdot\vec{p}, \alpha^I\frac{r_\mathrm{s}}{r}\right\} - \frac{1}{2}\left\{\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \alpha^I\right\}\right. \\
& \left. + \cancel{\frac{1}{2}\left\{\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \alpha^I\frac{r_\mathrm{s}}{r}\right\}}\right) \\
=& -\frac{\mathrm{i}\,\beta}{2m}\left(2p^I - \vec{\alpha}\cdot\vec{p}\alpha^I\frac{r_\mathrm{s}}{r} - \alpha^I\frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p} - \frac{1}{2}\left(\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}\alpha^I + \alpha^I\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}\right)\right) \\
=& -\frac{\mathrm{i}\,\beta}{2m}\left(2p^I - \left(\vec{\alpha}\cdot\left(\vec{p}\frac{r_\mathrm{s}}{r}\right)\alpha^I + \frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p}\alpha^I + \frac{r_\mathrm{s}}{r}\alpha^I\vec{\alpha}\cdot\vec{p}\right)\right. \\
& \left. -\frac{1}{2}\left(\vec{\alpha}\cdot\vec{p}\frac{r_\mathrm{s}}{r}\alpha^I + \frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p}\alpha^I + \alpha^I\vec{\alpha}\cdot\vec{p}\frac{r_\mathrm{s}}{r} + \alpha^I\frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p}\right)\right) \\
=& -\frac{\mathrm{i}\,\beta}{2m}\left(2p^I - \left(\left(p\frac{r_\mathrm{s}}{r}\right)^I + \mathrm{i}\left(\vec{\sigma}\times\left(\vec{p}\frac{r_\mathrm{s}}{r}\right)\right)^I + \frac{r_\mathrm{s}}{r}\{\vec{\alpha}\cdot\vec{p}, \alpha^I\}\right)\right. \\
& \left. -\frac{1}{2}\left(\{\vec{\alpha}\cdot\vec{p}, \alpha^I\}\frac{r_\mathrm{s}}{r} + \frac{r_\mathrm{s}}{r}\{\vec{\alpha}\cdot\vec{p}, \alpha^I\}\right)\right) \\
=& -\frac{\mathrm{i}\,\beta}{2m}\left(2p^I - \left(\left(p^I\frac{r_\mathrm{s}}{r}\right) + 2\frac{r_\mathrm{s}}{r}p^I + \mathrm{i}\left(\vec{\sigma}\times\left(\mathrm{i}\frac{\vec{r}r_\mathrm{s}}{r^3}\right)\right)^I\right)\right. \\
& \left. -\frac{1}{2}\left(2p^I\frac{r_\mathrm{s}}{r} + 2\frac{r_\mathrm{s}}{r}p^I\right)\right) \\
=& -\frac{\mathrm{i}\,\beta}{2m}\left(2p^I - \left\{p^I, \frac{r_\mathrm{s}}{r}\right\} + \frac{r_\mathrm{s}}{r^3}(\vec{\sigma}\times\vec{r})^I - \left\{p^I, \frac{r_\mathrm{s}}{r}\right\}\right) \\
=& -\mathrm{i}\,\beta\left(\frac{p^I}{m} - \left\{\frac{p^I}{m}, \frac{r_\mathrm{s}}{r}\right\} + \frac{r_\mathrm{s}}{2mr^3}(\vec{\sigma}\times\vec{r})^I\right)\,,
\end{aligned}
\tag{9.113}
$$

then J_1^I,

$$
\begin{aligned}
[S, J_1^I] =& -\frac{\mathrm{i}}{2m}\left[\beta\vec{\alpha}\cdot\vec{p} - \frac{\beta}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \alpha^I(\mathrm{i}\vec{k}\cdot\vec{r}) - \alpha^I\frac{r_\mathrm{s}}{r}(\mathrm{i}\vec{k}\cdot\vec{r})\right] \\
=& -\frac{\mathrm{i}\,\beta}{2m}\left(\left\{\vec{\alpha}\cdot\vec{p}, \alpha^I(\mathrm{i}\vec{k}\cdot\vec{r})\right\} - \left\{\vec{\alpha}\cdot\vec{p}, \alpha^I\frac{r_\mathrm{s}}{r}(\mathrm{i}\vec{k}\cdot\vec{r})\right\}\right. \\
&\left. -\frac{1}{2}\left\{\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \alpha^I(\mathrm{i}\vec{k}\cdot\vec{r})\right\} + \cancel{\frac{1}{2}\left\{\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \alpha^I\frac{r_\mathrm{s}}{r}(\mathrm{i}\vec{k}\cdot\vec{r})\right\}}\right) \\
=& -\frac{\mathrm{i}\,\beta}{2m}\left(\left(\mathrm{i}\vec{\alpha}\cdot\vec{p}\alpha^I\vec{k}\cdot\vec{r} + \mathrm{i}\alpha^I\vec{k}\cdot\vec{r}\vec{\alpha}\cdot\vec{p}\right) - \left(\mathrm{i}\vec{\alpha}\cdot\vec{p}\alpha^I\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r} + \mathrm{i}\alpha^I\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\vec{\alpha}\cdot\vec{p}\right)\right. \\
&\left. -\frac{\mathrm{i}}{2}\left(\vec{\alpha}\cdot\vec{p}\frac{r_\mathrm{s}}{r}\alpha^I\vec{k}\cdot\vec{r} + \frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p}\alpha^I\vec{k}\cdot\vec{r} + \alpha^I\vec{k}\cdot\vec{r}\vec{\alpha}\cdot\vec{p}\frac{r_\mathrm{s}}{r} + \alpha^I\vec{k}\cdot\vec{r}\frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p}\right)\right) \\
=& \frac{\beta}{2m}\left(\left(\vec{\alpha}\cdot(\vec{p}\vec{k}\cdot\vec{r})\alpha^I + \vec{k}\cdot\vec{r}\vec{\alpha}\cdot\vec{p}\alpha^I + \alpha^I\vec{k}\cdot\vec{r}\vec{\alpha}\cdot\vec{p}\right)\right. \\
& - \left(\vec{\alpha}\cdot\left(\vec{p}\vec{k}\cdot\vec{r}\frac{r_\mathrm{s}}{r}\right)\alpha^I + \vec{k}\cdot\vec{r}\frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p}\alpha^I + \alpha^I\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\vec{\alpha}\cdot\vec{p}\right) \\
& -\frac{1}{2}\left(\vec{\alpha}\cdot\left(\vec{p}\vec{k}\cdot\vec{r}\right)\frac{r_\mathrm{s}}{r}\alpha^I + \vec{k}\cdot\vec{r}\vec{\alpha}\cdot\vec{p}\frac{r_\mathrm{s}}{r}\alpha^I + \frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\left(\vec{p}\vec{k}\cdot\vec{r}\right)\alpha^I\right. \\
& \left.\left. +\vec{k}\cdot\vec{r}\frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p}\alpha^I + \alpha^I\vec{k}\cdot\vec{r}\vec{\alpha}\cdot\vec{p}\frac{r_\mathrm{s}}{r} + \alpha^I\vec{k}\cdot\vec{r}\frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p}\right)\right) \\
=& \frac{\beta}{2m}\left(\left(-\mathrm{i}\vec{\alpha}\cdot\vec{k}\alpha^I + 2\vec{k}\cdot\vec{r}p^I\right)\right. \\
& - \left(\vec{\alpha}\cdot\left(\vec{p}\vec{k}\cdot\vec{r}\frac{r_\mathrm{s}}{r}\right)\alpha^I + \vec{k}\cdot\vec{r}\frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p}\alpha^I + \vec{k}\cdot\vec{r}\frac{r_\mathrm{s}}{r}\alpha^I\vec{\alpha}\cdot\vec{p}^I\right) \\
& \left. -\frac{1}{2}\left(-\mathrm{i}\vec{\alpha}\cdot\vec{k}\frac{r_\mathrm{s}}{r}\alpha^I - \mathrm{i}\frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{k}\alpha^I + \vec{k}\cdot\vec{r}\left(\frac{r_\mathrm{s}}{r}\{\vec{\alpha}\cdot\vec{p}, \alpha^I\} + \{\vec{\alpha}\cdot\vec{p}, \alpha^I\}\frac{r_\mathrm{s}}{r}\right)\right)\right) \\
=& \frac{\beta}{2m}\left(\left((p^I\vec{k}\cdot\vec{r}) - (\vec{k}\times\vec{\sigma})^I + 2\vec{k}\cdot\vec{r}p^I\right) - \left(p^I\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\right) - \mathrm{i}\left(\vec{\sigma}\times\left(\vec{p}\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\right)\right)^I\right. \\
& \left. -2\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}p^I - \frac{1}{2}\left(-\mathrm{i}\frac{r_\mathrm{s}}{r}\{\vec{\alpha}\cdot\vec{k}, \alpha^I\} + 2\vec{k}\cdot\vec{r}\left\{p^I, \frac{r_\mathrm{s}}{r}\right\}\right)\right) \\
=& \frac{\beta}{2m}\left(\left(\{p^I, \vec{k}\cdot\vec{r}\} - (\vec{k}\times\vec{\sigma})^I\right) - \left\{\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}, p^I\right\} - \cancel{\frac{\mathrm{i}\vec{k}\cdot\vec{r}r_\mathrm{s}}{r^3}(\vec{\sigma}\times\vec{r})^I}\right. \\
& \left. -\frac{r_\mathrm{s}}{r}(\vec{\sigma}\times\vec{k})^I - \frac{1}{2}\left(-\mathrm{i}2k^I + 2\vec{k}\cdot\vec{r}\left\{p^I, \frac{r_\mathrm{s}}{r}\right\}\right)\right) \\
=& \frac{\beta}{2m}\left(\{p^I, \vec{k}\cdot\vec{r}\} - (\vec{k}\times\vec{\sigma})^I - \frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}p^I - p^I\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r} - \frac{r_\mathrm{s}}{r}(\vec{\sigma}\times\vec{k})^I\right. \\
& \left. -\frac{r_\mathrm{s}}{r}p^I\vec{k}\cdot\vec{r} - \vec{k}\cdot\vec{r}p^I\frac{r_\mathrm{s}}{r}\right) \\
=& -\mathrm{i}\,\beta\left(\frac{1}{2}\left\{\frac{p^I}{m}, \mathrm{i}\vec{k}\cdot\vec{r}\right\} - \frac{\mathrm{i}}{2m}(\vec{k}\times\vec{\sigma})^I - \frac{1}{2}\left\{\mathrm{i}\vec{k}\cdot\vec{r}, \left\{\frac{p^I}{m}, \frac{r_\mathrm{s}}{r}\right\}\right\}\right. \\
& \left. +\frac{\mathrm{i}}{2m}\frac{r_\mathrm{s}}{r}(\vec{k}\times\vec{\sigma})^I\right)\,,
\end{aligned}
\tag{9.114}
$$

and finally J_2^I,

$$
\begin{aligned}
[S, J_2^I] =& -\frac{\mathrm{i}}{2m}\left[\beta\vec{\alpha}\cdot\vec{p} - \cancel{\frac{\beta}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}}, -\frac{1}{2}\alpha^I(\vec{k}\cdot\vec{r})^2 + \frac{1}{2}\alpha^I\frac{r_\mathrm{s}}{r}(\vec{k}\cdot\vec{r})^2\right] \\
=& \frac{\mathrm{i}\,\beta}{4m}\left(\left\{\vec{\alpha}\cdot\vec{p}, \alpha^I(\vec{k}\cdot\vec{r})^2\right\} - \cancel{\left\{\vec{\alpha}\cdot\vec{p}, \alpha^I\frac{r_\mathrm{s}}{r}(\vec{k}\cdot\vec{r})^2\right\}}\right) \\
=& \frac{\mathrm{i}\,\beta}{4m}\left(\vec{\alpha}\cdot(\vec{p}(\vec{k}\cdot\vec{r})^2)\alpha^I + (\vec{k}\cdot\vec{r})^2\vec{\alpha}\cdot\vec{p}\alpha^I + (\vec{k}\cdot\vec{r})^2\alpha^I\vec{\alpha}\cdot\vec{p}\right) \\
=& -\frac{\mathrm{i}\,\beta}{4m}\left((p^I(\vec{k}\cdot\vec{r})^2) - \mathrm{i}((\vec{p}(\vec{k}\cdot\vec{r})^2)\times\vec{\sigma})^I + 2(\vec{k}\cdot\vec{r})^2 p^I\right) \\
=& \frac{\mathrm{i}\,\beta}{4m}\left(\{p^I, (\vec{k}\cdot\vec{r})^2\} - 2\vec{k}\cdot\vec{r}(\vec{k}\times\vec{\sigma})^I\right) \\
=& -\mathrm{i}\,\beta\left(-\frac{1}{4}\left\{(\vec{k}\cdot\vec{r})^2, \frac{p^I}{m}\right\} + \frac{1}{2m}(\vec{k}\cdot\vec{r})(\vec{k}\times\vec{\sigma})^I\right) .
\end{aligned}
\tag{9.115}
$$

We can now proceed to the double commutators, again starting with J_0^I,

$$
\begin{aligned}
[S, [S, J_0^I]] =& \frac{\mathrm{i}}{2m}\left[\beta\vec{\alpha}\cdot\vec{p} - \frac{\beta}{2}\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \mathrm{i}\beta\left(\frac{p^I}{m} - \left\{\frac{p^I}{m}, \frac{r_\mathrm{s}}{r}\right\} + \cancel{\frac{r_\mathrm{s}}{2mr^3}(\vec{\sigma}\times\vec{r})^I}\right)\right] \\
=& \frac{1}{2m}\left(\left\{\vec{\alpha}\cdot\vec{p}, \frac{p^I}{m}\right\} - \left\{\vec{\alpha}\cdot\vec{p}, \left\{\frac{p^I}{m}, \frac{r_\mathrm{s}}{r}\right\}\right\} - \frac{1}{2}\left\{\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \frac{p^I}{m}\right\}\right. \\
&\left. + \cancel{\frac{1}{2}\left\{\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \left\{\frac{p^I}{m}, \frac{r_\mathrm{s}}{r}\right\}\right\}}\right) \\
=& \vec{\alpha}\cdot\vec{p}\frac{p^I}{m^2} - \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \left\{\frac{p^I}{m^2}, \frac{r_\mathrm{s}}{r}\right\}\right\} - \frac{1}{4}\left\{\left\{\vec{\alpha}\cdot\vec{p}, \frac{r_\mathrm{s}}{r}\right\}, \frac{p^I}{m^2}\right\} ,
\end{aligned}
\tag{9.116}
$$

followed by J_1^I,

$$\begin{aligned}[S,[S,J_1^I]] = &-\frac{\mathrm{i}}{2m}\left[\beta\vec{\alpha}\cdot\vec{p}-\frac{\beta}{2}\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\},-\mathrm{i}\beta\left(\frac{1}{2}\left\{\frac{p^I}{m},\mathrm{i}\vec{k}\cdot\vec{r}\right\}-\frac{\mathrm{i}(\vec{k}\times\vec{\sigma})^I}{2m}\right.\right.\\ &\left.\left.\cancel{-\frac{1}{2}\left\{\mathrm{i}\vec{k}\cdot\vec{r},\left\{\frac{p^I}{m},\frac{r_\mathrm{s}}{r}\right\}\right\}}+\cancel{\frac{\mathrm{i}r_\mathrm{s}(\vec{k}\times\vec{\sigma})^I}{2mr}}\right)\right]\\ =&\frac{1}{2m}\left(\frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p},\left\{\frac{p^I}{m},\mathrm{i}\vec{k}\cdot\vec{r}\right\}\right\}-\left\{\vec{\alpha}\cdot\vec{p},\frac{\mathrm{i}(\vec{k}\times\vec{\sigma})^I}{2m}\right\}\right.\\ &\left.\cancel{-\frac{1}{4}\left\{\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\},\left\{\frac{p^I}{m},\mathrm{i}\vec{k}\cdot\vec{r}\right\}\right\}}+\frac{1}{2}\left\{\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\},\frac{\mathrm{i}(\vec{k}\times\vec{\sigma})^I}{2m}\right\}\right)\\ =&\frac{1}{4}\left\{\vec{\alpha}\cdot\vec{p},\left\{\frac{p^I}{m^2},\mathrm{i}\vec{k}\cdot\vec{r}\right\}\right\}-\left\{\vec{\alpha}\cdot\vec{p},\frac{\mathrm{i}(\vec{k}\times\vec{\sigma})^I}{4m^2}\right\}\\ &+\left\{\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\},\frac{\mathrm{i}(\vec{k}\times\vec{\sigma})^I}{8m^2}\right\}, \end{aligned} \tag{9.117}$$

and finally J_2^I,

$$\begin{aligned}[S,[S,J_2^I]] = &-\frac{\mathrm{i}}{2m}\left[\beta\vec{\alpha}\cdot\vec{p}-\cancel{\frac{\beta}{2}\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\}},\right.\\ &\left.-\mathrm{i}\,\beta\left(-\frac{1}{4}\left\{(\vec{k}\cdot\vec{r})^2,\frac{p^I}{m}\right\}+\frac{1}{2m}(\vec{k}\cdot\vec{r})(\vec{k}\times\vec{\sigma})^I\right)\right]\\ =&\frac{1}{2m}\left(-\frac{1}{4}\left\{\vec{\alpha}\cdot\vec{p},\left\{(\vec{k}\cdot\vec{r})^2,\frac{p^I}{m}\right\}\right\}+\left\{\vec{\alpha}\cdot\vec{p},\frac{(\vec{k}\cdot\vec{r})(\vec{k}\times\vec{\sigma})^I}{2m}\right\}\right). \end{aligned} \tag{9.118}$$

These results enable us to press forward to the triple commutators, as before, we begin with J_0^I,

$$\begin{aligned}[S,[S,[S,J_0^I]]] = &-\frac{\mathrm{i}}{2m}\left[\beta\vec{\alpha}\cdot\vec{p}-\cancel{\frac{\beta}{2}\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\}},\vec{\alpha}\cdot\vec{p}\frac{p^I}{m^2}\right.\\ &\left.\cancel{-\frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p},\left\{\frac{p^I}{m^2},\frac{r_\mathrm{s}}{r}\right\}\right\}}-\cancel{\frac{1}{4}\left\{\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\},\frac{p^I}{m^2}\right\}}\right]\\ =&-\frac{\mathrm{i}\,\beta}{2m}\left\{\vec{\alpha}\cdot\vec{p},\vec{\alpha}\cdot\vec{p}\frac{p^I}{m^3}\right\}=-\mathrm{i}\,\beta\frac{\vec{p}^{\,2}p^I}{m^3}. \end{aligned} \tag{9.119}$$

From here it is clear that the remainder of the triple commutators will yield results of higher order then we are interested in. As such we can then move on to the next rotation. Recall that for the Dirac–Einstein–Schwarzschild Hamiltonian (9.67)

$$\begin{aligned}\mathcal{O}' =& -\frac{1}{3m^2}\left(\vec{\alpha}\cdot\vec{p}-\frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\}\right)^3+\frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}-\cancel{\frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\}},\frac{r_\mathrm{s}}{2r}\right\}\\ \approx& -\frac{\vec{p}^2\vec{\alpha}\cdot\vec{p}}{3m^2}+\frac{1}{4}\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\}\,. \end{aligned}\tag{9.120}$$

Furthermore, we know that

$$(J^I)' = J_0^I + J_1^I + J_2^I + \mathrm{i}[S,J^I] - \frac{1}{2}[S,[S,J^I]] - \frac{\mathrm{i}}{6}[S,[S,[S,J^I]]] + \ldots\,, \tag{9.121}$$

Thus by looking at the results that have already been obtained as well as the impending rotation, it becomes clear that the only terms that will be of low enough order when rotated are J_0^I and J_1^I. We thus begin the second iteration of rotations,

$$\begin{aligned}[S',J_0^I] =& -\frac{\mathrm{i}}{2m}\left[-\beta\frac{\vec{p}^2}{3m^2}\vec{\alpha}\cdot\vec{p}+\frac{\beta}{4}\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\},\alpha^I-\alpha^I\frac{r_\mathrm{s}}{r}\right]\\ =& -\frac{\mathrm{i}\,\beta}{2m}\left(-\frac{\vec{p}^2}{3m^2}\{\vec{\alpha}\cdot\vec{p},\alpha^I\}+\cancel{\frac{\vec{p}^2}{3m^2}\left\{\vec{\alpha}\cdot\vec{p},\alpha^I\frac{r_\mathrm{s}}{r}\right\}}+\frac{1}{4}\left\{\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\},\alpha^I\right\}\right.\\ &\left.\cancel{-\frac{1}{4}\left\{\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\},\alpha^I\frac{r_\mathrm{s}}{r}\right\}}\right)\\ =& -\frac{\mathrm{i}\,\beta}{2m}\left(-\frac{2\vec{p}^2p^I}{3m^2}+\frac{1}{2}\left\{p^I,\frac{r_\mathrm{s}}{r}\right\}\right) = -\mathrm{i}\,\beta\left(-\frac{\vec{p}^2p^I}{3m^3}+\frac{1}{4}\left\{\frac{p^I}{m},\frac{r_\mathrm{s}}{r}\right\}\right)\,, \end{aligned}\tag{9.122}$$

and the final rotation needed for this order,

$$\begin{aligned}[S',J_1^I] =& -\frac{\mathrm{i}}{2m}\left[\cancel{-\beta\frac{\vec{p}^2}{3m^2}\vec{\alpha}\cdot\vec{p}}+\frac{\beta}{4}\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\},\alpha^I(\mathrm{i}\vec{k}\cdot\vec{r})-\cancel{\alpha^I\frac{r_\mathrm{s}}{r}(\mathrm{i}\vec{k}\cdot\vec{r})}\right]\\ =& \frac{\beta}{2m}\left(\frac{1}{4}\left\{\left\{\vec{\alpha}\cdot\vec{p},\frac{r_\mathrm{s}}{r}\right\},\alpha^I\vec{k}\cdot\vec{r}\right\}\right)\\ =& \frac{\beta}{8m}\left(\vec{\alpha}\cdot\vec{p}\frac{r_\mathrm{s}}{r}\alpha^I\vec{k}\cdot\vec{r}+\frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p}\alpha^I\vec{k}\cdot\vec{r}+\alpha^I\vec{k}\cdot\vec{r}\vec{\alpha}\cdot\vec{p}\frac{r_\mathrm{s}}{r}+\alpha^I\vec{k}\cdot\vec{r}\frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\vec{p}\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{\beta}{8m}\left(\vec{\alpha}\cdot\left(\vec{p}\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\right)\alpha^I + \frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\vec{\alpha}\cdot\vec{p}\alpha^I + \frac{r_\mathrm{s}}{r}\vec{\alpha}\cdot\left(\vec{p}\vec{k}\cdot\vec{r}\right)\alpha^I + \frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\vec{\alpha}\cdot\vec{p}\alpha^I\right.\\
&\quad\left.+\vec{k}\cdot\vec{r}\alpha^I\vec{\alpha}\cdot\left(\vec{p}\frac{r_\mathrm{s}}{r}\right) + \frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\alpha^I\vec{\alpha}\cdot\vec{p} + \frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\alpha^I\vec{\alpha}\cdot\vec{p}\right)\\
&= \frac{\beta}{8m}\left(2\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\left\{\vec{\alpha}\cdot\vec{p},\alpha^I\right\} + \left(p^I\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\right) + \mathrm{i}\left(\vec{\sigma}\times\left(\vec{p}\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\right)\right)^I\right.\\
&\quad\left.+\frac{r_\mathrm{s}}{r}\left(p^I\vec{k}\cdot\vec{r}\right) + \mathrm{i}\frac{r_\mathrm{s}}{r}\left(\vec{\sigma}\times\left(\vec{p}\vec{k}\cdot\vec{r}\right)\right)^I + \vec{k}\cdot\vec{r}\left(p^I\frac{r_\mathrm{s}}{r}\right) - \mathrm{i}\vec{k}\cdot\vec{r}\left(\vec{\sigma}\times\left(\vec{p}\frac{r_\mathrm{s}}{r}\right)\right)^I\right)\\
&= \frac{\beta}{8m}\left(4\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}p^I + 2\left(p^I\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r}\right) + \mathrm{i}\frac{r_\mathrm{s}}{r}(\vec{\sigma}\times(\vec{p}\vec{k}\cdot\vec{r}))^I + \cancel{\mathrm{i}\vec{k}\cdot\vec{r}\left(\vec{\sigma}\times\left(\vec{p}\frac{r_\mathrm{s}}{r}\right)\right)^I}\right.\\
&\quad\left.+\mathrm{i}\frac{r_\mathrm{s}}{r}(\vec{\sigma}\times(\vec{p}\vec{k}\cdot\vec{r}))^I - \cancel{\mathrm{i}\vec{k}\cdot\vec{r}\left(\vec{\sigma}\times\left(\vec{p}\frac{r_\mathrm{s}}{r}\right)\right)^I}\right)\\
&= \frac{\beta}{8m}\left(2\left\{\frac{r_\mathrm{s}}{r}\vec{k}\cdot\vec{r},p^I\right\} + 2\frac{r_\mathrm{s}}{r}(\vec{\sigma}\times\vec{k})^I\right)\\
&= -\,\mathrm{i}\,\beta\left(\frac{1}{4}\left\{\frac{r_\mathrm{s}}{r}(\mathrm{i}\vec{k}\cdot\vec{r}),\frac{p^I}{m}\right\} - \frac{\mathrm{i}r_\mathrm{s}}{4r}\frac{(\vec{k}\times\vec{\sigma})^I}{m}\right)\,. \qquad (9.123)
\end{aligned}
$$

To finalize the calculation of the current, we throw out the odd terms, set $\beta = 1$ and add all the remaining terms with the appropriate prefactors. For convenience these are listed below;

$$\mathrm{i}[S,J_0^I] = \frac{p^I}{m} - \left\{\frac{p^I}{m},\frac{r_\mathrm{s}}{r}\right\} + \frac{r_\mathrm{s}}{2r}\frac{(\vec{\sigma}\times\vec{r})^I}{mr^2}\,, \qquad (9.124a)$$

$$
\begin{aligned}
\mathrm{i}[S,J_1^I] &= \frac{1}{2}\left\{\frac{p^I}{m},\mathrm{i}\vec{k}\cdot\vec{r}\right\} - \frac{\mathrm{i}}{2m}(\vec{k}\times\vec{\sigma})^I\\
&\quad - \frac{1}{2}\left\{\mathrm{i}\vec{k}\cdot\vec{r},\left\{\frac{p^I}{m},\frac{r_\mathrm{s}}{r}\right\}\right\} + \frac{\mathrm{i}r_\mathrm{s}}{2r}\frac{(\vec{k}\times\vec{\sigma})^I}{m}\,, \qquad (9.124b)
\end{aligned}
$$

$$\mathrm{i}[S,J_2^I] = -\frac{1}{4}\left\{(\vec{k}\cdot\vec{r})^2,\frac{p^I}{m}\right\} + \frac{(\vec{k}\cdot\vec{r})}{2m}(\vec{k}\times\vec{\sigma})^I\,, \qquad (9.124c)$$

$$-\frac{\mathrm{i}}{6}[S,[S,[S,J_0^I]]] = -\frac{p^I\vec{p}^2}{6m^3}\,, \qquad (9.124d)$$

$$\mathrm{i}[S',J_0^I] = -\frac{p^I\vec{p}^2}{3m^3} + \frac{1}{4}\left\{\frac{p^I}{m},\frac{r_\mathrm{s}}{r}\right\}\,, \qquad (9.124e)$$

$$\mathrm{i}[S',J_1^I] = \frac{1}{4}\left\{\frac{r_\mathrm{s}}{r}(\mathrm{i}\vec{k}\cdot\vec{r}),\frac{p^I}{m}\right\} - \frac{\mathrm{i}r_\mathrm{s}}{4r}\frac{(\vec{k}\times\vec{\sigma})^I}{m}\,. \qquad (9.124f)$$

Putting this all together we find

$$\begin{aligned} J^I_{\mathrm{FW}} = &\frac{p^I}{m} - \frac{p^I \vec{p}^2}{2m^3} - \frac{\mathrm{i}}{2m}(\vec{k}\times\vec{\sigma})^I + \frac{1}{2}\left\{\frac{p^I}{m}, \mathrm{i}\vec{k}\cdot\vec{r}\right\} \\ &- \frac{1}{4}\left\{(\vec{k}\cdot\vec{r})^2, \frac{p^I}{m}\right\} + \frac{(\vec{k}\cdot\vec{r})}{2m}(\vec{k}\times\vec{\sigma})^I \\ &- \frac{3}{4}\left\{\frac{p^I}{m}, \frac{r_{\mathrm{s}}}{r}\right\} + \frac{r_{\mathrm{s}}}{2r}\frac{(\vec{\sigma}\times\vec{r})^I}{mr^2} - \frac{1}{2}\left\{\mathrm{i}\vec{k}\cdot\vec{r}, \left\{\frac{p^I}{m}, \frac{r_{\mathrm{s}}}{r}\right\}\right\} \\ &+ \frac{\mathrm{i}r_{\mathrm{s}}}{4r}\frac{(\vec{k}\times\vec{\sigma})^I}{m} + \frac{1}{4}\left\{\frac{r_{\mathrm{s}}}{r}(\mathrm{i}\vec{k}\cdot\vec{r}), \frac{p^I}{m}\right\} . \end{aligned} \tag{9.125}$$

In addition to the terms that are already known to the relativistic physics community, which take up the first two lines, we have a gravitational correction to the dipole coupling, which is the first term of the third line. There is also a gravitational correction to the quadrupole coupling (the third term of the third line and the final term of the fourth line). The other two terms, i.e., the second term of the third line and first term of the fourth line, are a gravitational correction to the magnetic coupling. These terms may be used to calculate the relativistic affects of photon emission of a particle in a gravitational potential.

10. CHIRAL FOLDY–WOUTHUYSEN TRANSFORMATION

10.1. ORIENTATION

In [24] a new "Foldy–Wouthuysen" transformation is discussed, which rather then using an iterative process, this proposed method uses an exact transformation, with an Taylor series approximation of a square root at the end, profiting from a deceptively innocuous decoupling of "even" and "odd" terms in the original Hamiltonian, based on "seductive" properties of the Dirac algebra under certain parity-breaking transformations. This transformation was investigated recently in reference [88]. Here, and through the rest of the chapter, we elaborate on the discussion, and provide greater detail with regards to the derivations. The proposed method utilizes the rotation

$$U = U_2 U_1 \tag{10.1}$$

where

$$U_1 = \frac{1}{\sqrt{2}} (1 + J \, \Lambda) \qquad U_2 = \frac{1}{\sqrt{2}} (1 + \beta J) \ , \tag{10.2}$$

and

$$\Lambda = \frac{H}{\sqrt{H^2}} \, , \qquad J = \mathrm{i} \, \gamma^5 \, \beta \, , \tag{10.3}$$

where H is the Hamiltonian that we are trying to transform. For this transformation to work, it is essential that

$$\{H, J\} = 0 \, , \tag{10.4}$$

in which case $J\,H^2 = H^2\,J$, from which it follows $J\sqrt{H^2} = \sqrt{H^2}\,J$. If we use these proposed rotations on a Hamiltonian where $\{H, J\} \neq 0$, then the following proofs do not hold, and we are no longer working with a unitary transform. On that note, we can show that provided the discussed condition is met, then the operator U is unitary,

$$\begin{aligned}
U\,U^{+} =& U_2\,U_1\,U_1^{+}\,U_2^{+} = U_2\frac{1}{\sqrt{2}}\,(1 + J\,\Lambda)\left(\frac{1}{\sqrt{2}}\,(1 + J\,\Lambda)\right)^{+} U_2^{+} \\
=& U_2\frac{1}{2}(1 + J\Lambda)(1 + \Lambda J) = \frac{1}{2}U_2(1 + J\Lambda + \Lambda J + J\Lambda\Lambda J)U_2^{+} \\
=& \frac{1}{2}U_2(2 + J\Lambda + \Lambda J)U_2^{+} = \frac{1}{4}(1 + \beta J)(2 + J\Lambda + \Lambda J)(1 + J\beta) \\
=& \frac{1}{4}(2 + 2\beta J + J\Lambda + \beta\Lambda + \Lambda J + \beta J\Lambda J)(1 + J\beta) \\
=& \frac{1}{4}(2 + 2\beta J + J\Lambda + \beta\Lambda + \Lambda J + \beta J\Lambda J + 2J\beta + 2\beta JJ\beta \\
&+ J\Lambda J\beta + \beta\Lambda J\beta + \Lambda JJ\beta + \beta J\Lambda JJ\beta)\,.
\end{aligned} \tag{10.5}$$

Let us now take note of a few properties

$$J^2 = 1\,, \qquad \beta^2 = 1\,, \qquad J\beta = -\beta J\,, \qquad J\Lambda = -\Lambda J\,, \tag{10.6}$$

in which case (10.5) becomes

$$\begin{aligned}
UU^{+} =& \frac{1}{4}(2 + 2\beta J + J\Lambda + \beta\Lambda - J\Lambda - \beta\Lambda - 2\beta J + 2 - \Lambda\beta + \beta\Lambda J\beta \\
&+ \Lambda\beta - \beta\Lambda J\beta) \\
=& \frac{1}{4}(4) = 1\,.
\end{aligned} \tag{10.7}$$

When this rotation is applied to the Hamiltonian H, which was used to construct the rotation, we find

$$
\begin{aligned}
U\,H\,U^{+} =& U_2 U_1\,H\,U_1^{+}U_2^{+} = \frac{1}{2}U_2(1+J\Lambda)H(1+\Lambda J)U_2^{+} \\
=& \frac{1}{2}U_2(H+J\Lambda H+H\Lambda J+J\Lambda H\Lambda J)U_2^{+} = \frac{1}{2}U_2(H+2J\Lambda H-H)U_2^{+} \\
=& U_2\,J\Lambda H\,U_2^{+} = \frac{1}{2}(1+\beta J)J\Lambda H(1+J\beta) \\
=& \frac{1}{2}(J\Lambda H+\beta\Lambda H+J\Lambda HJ\beta+\beta\Lambda HJ\beta) \\
=& \frac{1}{2}(J\Lambda H+\beta\Lambda H+\Lambda H\beta-\beta\Lambda H\beta J) \\
=& \frac{1}{2}(\sqrt{H^2}+\beta\sqrt{H^2}\beta)\beta+\frac{1}{2}(\sqrt{H^2}-\beta\sqrt{H^2}\beta)J \\
=& \left[\sqrt{H^2}\right]_{\text{even}}\beta+\left\{\sqrt{H^2}\right\}_{\text{odd}}J\,.
\end{aligned}
\tag{10.8}
$$

Given an operator A we can find its even and odd parts, in spinor space, using the definitions

$$
[A]_{\text{even}} \equiv \frac{1}{2}(A+\beta A\beta)\,, \qquad \{A\}_{\text{odd}} \equiv \frac{1}{2}(A-\beta A\beta)\,, \tag{10.9}
$$

thus in the “chiral” rotation is performed by dividing the operator $\sqrt{H^2}$ into its even and odd components in spinor space, and multiplying said components by either β or J (respectively). We note that J is odd in spinor space, meaning that the rotated Hamiltonian give in (10.8) is even. To separate the even and odd parts of $\sqrt{H^2}$, one must expand the operator in terms of momenta. This seems like a simple enough procedure, and it could go a long way in reducing the complexity of the standard Foldy–Wouthuysen transformation. For the rest of the chapter we shall apply the chiral transformation to the Dirac Hamiltonians which we investigated in chapter 9. To somewhat simplify the calculations, we will only keep terms up to the third order in momenta, except for the case of the free particle.

10.2. FREE PARTICLE

Recall that the Dirac Hamiltonian for a free particle is given as

$$H_{\mathrm{F}} = \vec{\alpha} \cdot \vec{p} + \beta m \, . \tag{10.10}$$

Then

$$H_{\mathrm{F}}^2 = (\vec{\alpha} \cdot \vec{p} + \beta m)(\vec{\alpha} \cdot \vec{p} + \beta m) = \vec{\alpha} \cdot \vec{p}\,\vec{\alpha} \cdot \vec{p} + \vec{\alpha} \cdot \vec{p}\beta m + \beta m \vec{\alpha} \cdot \vec{p} + m^2 = \vec{p}^{\,2} + m^2 \, , \tag{10.11}$$

and

$$\sqrt{H_{\mathrm{F}}^2} \approx m + \frac{\vec{p}^{\,2}}{2m} - \frac{\vec{p}^{\,4}}{8m^3} \, . \tag{10.12}$$

From which we calculate

$$\left[\sqrt{H_{\mathrm{F}}^2}\right]_{\mathrm{even}} \beta = \beta \left(m + \frac{\vec{p}^{\,2}}{2m} \right) \, , \tag{10.13}$$

and

$$\left\{\sqrt{H_{\mathrm{F}}^2}\right\}_{\mathrm{odd}} J = 0 \, . \tag{10.14}$$

Thus, expanded to the fourth order in momenta,

$$H_{\mathrm{F}}^{(\mathrm{CFW})} = \beta \left(m + \frac{\vec{p}^{\,2}}{2m} - \frac{\vec{p}^{\,4}}{8m^3} \right) \, , \tag{10.15}$$

where we used the superscript “(CFW)” indicate that the result was derived using the chiral Foldy–Wouthuysen transform. It is important that this not be confused with the superscript “(FW),” which was used to indicate that the result was obtained using the *standard* Foldy–Wouthuysen transform. This matches the result for the free particle using the standard Foldy–Wouthuysen transformation (see (9.15)), and it seems as though performing the transformation is much easier. The chiral method might have potential, but lets see how it works with the remainder of our Hamiltonians.

10.3. DIRAC-COULOMB HAMILTONIAN

To further investigate this method we look at how the rotation effects the Dirac-Coulomb Hamiltonian. We recall that the Hamiltonian is

$$H_{\mathrm{DC}} = \vec{\alpha} \cdot \vec{p} + \beta m - \frac{Z\alpha}{r}\,. \tag{10.16}$$

At this point we notice that $J\,H_{\mathrm{DC}} \neq -H_{\mathrm{DC}}\,J$, and (10.4) is violated. Strictly speaking, the Dirac–Coulomb Hamiltonian does not fulfill a necessary condition for the application of the chiral transform. Despite this shortcoming, we apply the chiral Foldy–Wouthuysen transform to H_{DC}, encouraged by the fact that there are cases in physics and mathematics where necessary conditions for the application of a mathematical method are not fulfilled, yet consistent results are attained. For example, physically consistent results can be attained using asymptotic expansions in a non–asymptotic regime when suitable resummation prescriptions are utilized [112–114].

With the pitfalls in mind, we proceed with the transformation, first squaring the Hamiltonian,

$$H_{\mathrm{DC}}^2 = m^2 + \vec{p}^{\,2} - Z\alpha\,\vec{\alpha} \cdot \left[\vec{p}, \frac{1}{r}\right] - 2\frac{Z\alpha}{r}\vec{\alpha} \cdot \vec{p} - 2\beta m \frac{Z\alpha}{r} + \frac{Z^2\alpha^2}{r^2}\,, \tag{10.17}$$

and then expand the square root,

$$\sqrt{H_{\mathrm{DC}}^2} \approx m + \frac{\vec{p}^{\,2}}{2m} - \frac{1}{2m} Z\alpha\vec{\alpha} \cdot \left[\vec{p}, \frac{1}{r}\right] - \frac{Z\alpha}{mr}\vec{\alpha} \cdot \vec{p} - \beta\frac{Z\alpha}{r} + \cancel{\frac{Z^2\alpha^2}{2mr^2}}\,, \tag{10.18}$$

Where we ignore the final term as it is of the fourth order in momenta. We then find

$$\left[\sqrt{H_{\mathrm{DC}}^2}\right]_{\mathrm{even}} \beta = \beta\left(m + \frac{\vec{p}^{\,2}}{2m}\right) - \frac{Z\alpha}{r} + \beta\frac{Z^2\alpha^2}{2mr^2}\,, \tag{10.19a}$$

$$\left\{\sqrt{H_{\mathrm{DC}}^2}\right\}_{\mathrm{odd}} J = +\frac{Z\alpha}{2mr^3}\beta\vec{\Sigma}\cdot\vec{r} - \mathrm{i}\frac{Z\alpha}{mr}\beta\vec{\Sigma}\cdot\vec{p}. \tag{10.19b}$$

Thus

$$H_{\mathrm{DC}}^{(\mathrm{CFW})} = \beta\left(m + \frac{\vec{p}^{\,2}}{2m}\right) - \frac{Z\alpha}{r} + \beta\frac{Z^2\alpha^2}{2mr^2} + \frac{Z\alpha}{2mr^3}\beta\vec{\Sigma}\cdot\vec{r} - \mathrm{i}\frac{Z\alpha}{mr}\beta\vec{\Sigma}\cdot\vec{p}, \tag{10.20}$$

which bears little resemblance to our result in (9.33), even when accounting for the higher order of the previous calculation. One further observes that there are two major issues with this result, beyond the fact that the result differs from the well known result. The first is that the second to last term is a pseudo–scalar, i.e., a pseudo–vector dotted with a vector, which violates parity, despite the fact that the original Hamiltonian is parity invariant. Secondly, the final term is not Hermitian. However, due to the already stated fact that the starting Hamiltonian H_{DC} does not fulfill the conditions needed to perform the chiral transformation. As a result the transformation used was not unitary. We cannot rule out the effectiveness of this transformation on a smaller class of Hamiltonians. This is contrary to the standard Foldy–Wouthuysen transformation which has a much broader applicability. The failure of the chiral method on the paradigmatic Dirac–Coulomb Hamiltonian indicates severe limits on the range of applicability of the method for practically interesting and phenomenologically important physical systems.

10.4. DIRAC HAMILTONIAN WITH A SCALAR POTENTIAL

The Dirac Hamiltonian with a scalar potential (discussed above in chapter 9.4, see also [88]) is given as

$$H_{\rm SP} = \vec{\alpha} \cdot \vec{p} + \beta \left(m - \frac{\lambda}{r} \right) . \tag{10.21}$$

Then,

$$\begin{aligned} H_{\rm SP}^2 &= \vec{p}^{\,2} + \left(m + \frac{\lambda}{r} \right)^2 + \vec{\alpha} \cdot \vec{p} \beta \left(m - \frac{\lambda}{r} \right) + \beta \left(m - \frac{\lambda}{r} \right) \vec{\alpha} \cdot \vec{p} \\ &= \vec{p}^{\,2} + m^2 - 2m\frac{\lambda}{r} + \cancel{\frac{\lambda^2}{r^2}} + \beta \left(\vec{\alpha} \cdot \vec{p} \frac{\lambda}{r} - \frac{\lambda}{r} \vec{\alpha} \cdot \vec{p} \right) \\ &= \vec{p}^{\,2} + m^2 + \beta \left[\vec{\alpha} \cdot \vec{p}, \frac{\lambda}{r} \right] - 2m\frac{\lambda}{r} . \end{aligned} \tag{10.22}$$

We can then expand the square root of the square,

$$\sqrt{H_{\rm SP}^2} \approx m + \frac{\vec{p}^{\,2}}{2m} + \frac{\lambda\beta}{2m} \left[\vec{\alpha} \cdot \vec{p}, \frac{1}{r} \right] - \frac{\lambda}{r} = m + \frac{\vec{p}^{\,2}}{2m} + \frac{\mathrm{i}\lambda\beta}{2mr^3} \vec{\alpha} \cdot \vec{r} - \frac{\lambda}{r} . \tag{10.23}$$

Then the even and odd parts are

$$\left[\sqrt{H_{\rm SP}^2} \right]_{\rm even} \beta = \beta \left(m + \frac{\vec{p}^{\,2}}{2m} - \frac{\lambda}{r} \right) , \tag{10.24}$$

$$\left\{ \sqrt{H_{\rm SP}^2} \right\}_{\rm odd} J = \frac{\mathrm{i}\lambda\beta}{2mr^3} \vec{\alpha} \cdot \vec{r} J = -\frac{\lambda}{2m} \frac{\vec{\Sigma} \cdot \vec{r}}{r^3} , \tag{10.25}$$

and the transformed Hamiltonian is

$$H_{\rm SP}^{\rm (CFW)} = \beta \left(m + \frac{\vec{p}^{\,2}}{2m} - \frac{\lambda}{r} \right) - \frac{\lambda}{2m} \frac{\vec{\Sigma} \cdot \vec{r}}{r^3} . \tag{10.26}$$

Again we find that there are a number of differences when comparing the results from the standard transformation (see (9.48)) with those of the chiral method. Unlike the Dirac–Coulomb case, the initial Hamiltonian, $H_{\rm SP}$, does meet the requirements imposed by the chiral transform, and the resulting Hamiltonian is indeed Hermitian. However, we again find that the last term breaks parity.

10.5. DIRAC HAMILTONIAN WITH A SCALAR CONFINING POTENTIAL

The Dirac Hamiltonian with a scalar confining potential (as discussed in chapter 9.5 and [88, 107, 108]) is given as

$$H_{\rm LC} = \vec{\alpha}\cdot\vec{p} + \beta\,\left(m + \alpha^2 m^2 r\right)\,. \tag{10.27}$$

Then

$$\begin{aligned} H_{\rm LC}^2 &= \vec{p}^{\,2} + (m+\alpha^2 m^2 r)^2 + \vec{\alpha}\cdot\vec{p}\,\beta(m+\alpha^2 m^2 r) + \beta(m+\alpha^2 m^2 r)\,\vec{\alpha}\cdot\vec{p} && (10.28)\\ &= \vec{p}^{\,2} + m^2 + 2\alpha^2 m^3 r + \alpha^4 m^4 r^2 + \beta\left(-\vec{\alpha}\cdot\vec{p}\,\alpha^2 m^2 r + \alpha^2 m^2 r\,\vec{\alpha}\cdot\vec{p}\right) && (10.29)\\ &= m^2 + \vec{p}^{\,2} + 2\alpha^2 m^3 r + \alpha^4 m^4 r^2 - \beta\alpha^2 m^2\left[\vec{\alpha}\cdot\vec{p}, r\right]\,. && (10.30) \end{aligned}$$

From here we can find the expansion of the square root, expanded about small momenta

$$\begin{aligned} \sqrt{H_{\rm LC}^2} &\approx m + \frac{\vec{p}^{\,2}}{2m} + \alpha^2 m^2 r + \frac{1}{2}\alpha^4 m^3 r^2 - \frac{\beta\alpha^2 m}{2}[\vec{\alpha}\cdot\vec{p}, r] && (10.31)\\ &= m + \frac{\vec{p}^{\,2}}{2m} + \alpha^2 m^2 r + \frac{1}{2}\alpha^4 m^3 r^2 - \frac{\beta\alpha^2 m}{2}\vec{\alpha}\cdot\left(-\frac{\mathrm{i}\,\vec{r}}{r}\right) && (10.32)\\ &= m + \frac{\vec{p}^{\,2}}{2m} + \alpha^2 m^2 r + \frac{1}{2}\alpha^4 m^3 r^2 + \frac{\mathrm{i}\,\alpha^2 m\,\beta}{2\,r}\vec{\alpha}\cdot\vec{r}. && (10.33) \end{aligned}$$

Then the even and odd (in spinor space) parts are

$$\left[\sqrt{H_{\rm LC}^2}\right]_{\rm even}\beta = \beta\left(m + \frac{\vec{p}^{\,2}}{2m} + \alpha^2 m^2 r + \frac{1}{2}\alpha^4 m^3 r^2\right)\,, \tag{10.34}$$

$$\left\{\sqrt{H_{\rm LC}^2}\right\}_{\rm odd} J = \frac{{\rm i}\,\alpha^2 m\,\beta}{2\,r}\vec{\alpha}\cdot\vec{r}\,J = -\frac{\alpha^2 m}{2\,r}\vec{\Sigma}\cdot\vec{r}. \tag{10.35}$$

Thus,

$$H_{\rm LC}^{\rm (CFW)} = \beta\left(m + \frac{\vec{p}^{\,2}}{2m} + \alpha^2 m^2 r + \frac{1}{2}\alpha^4 m^3 r^2\right) - \frac{\alpha^2 m}{2\,r}\vec{\Sigma}\cdot\vec{r}. \tag{10.36}$$

Again, while there are some similarities, we find that when this transformed Hamiltonian is compared to its counterpart, found using the standard transformation (see (9.58)), there are obvious differences. Additionally, we again find that the final term is a parity breaking term.

10.6. DIRAC–EINSTEIN–SCHWARZSCHILD HAMILTONIAN

The chiral Fold–Wouthuysen transformation has been previously performed in [94]. However the formalism was somewhat different, and it is advantageous to perform the calculation using our formalism. Recall that the Dirac–Einstein–Schwarzschild Hamiltonian was previously derived in (8.84) as

$$H_{\rm DS} = \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \left(1 - \frac{r_{\rm s}}{r}\right)\right\} + \beta\,m\left(1 - \frac{r_{\rm s}}{2r}\right)\,. \tag{10.37}$$

However, the calculation is somewhat easier to perform when written in a more compact form, i.e.,

$$H_{\rm DS} = \beta\,m\,w + \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, \mathcal{F}\right\}\,, \tag{10.38}$$

$$w = 1 - \frac{r_{\rm s}}{2r}\,, \quad v = 1 + \frac{r_{\rm s}}{2r}\,, \quad \mathcal{F} = \frac{w}{v} \approx 1 - \frac{r_{\rm s}}{r}\,. \tag{10.39}$$

We begin by first squaring the Hamiltonian,

$$
\begin{aligned}
H_{\mathrm{DS}}^2 = & \left(\beta m\, w + \frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \mathcal{F}\}\right)\left(\beta m\, w + \frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \mathcal{F}\}\right) \\
= & \,(\beta m\, w)\,(\beta m\, w) + (\beta m\, w)\left(\frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \mathcal{F}\}\right) + \left(\frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \mathcal{F}\}\right)(\beta m\, w) \\
& + \left(\frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \mathcal{F}\}\right)\left(\frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \mathcal{F}\}\right) \\
= & \, m^2 w^2 + \frac{m}{2}\left(\beta w\,\{\vec{\alpha}\cdot\vec{p}, \mathcal{F}\} + \{\vec{\alpha}\cdot\vec{p}, \mathcal{F}\}\,\beta w\right) \\
& + \frac{1}{4}\left(\vec{\alpha}\cdot\vec{p}\mathcal{F} + \mathcal{F}\vec{\alpha}\cdot\vec{p}\right)\left(\vec{\alpha}\cdot\vec{p}\mathcal{F} + \mathcal{F}\vec{\alpha}\cdot\vec{p}\right) \\
= & \, m^2 w^2 + \frac{m}{2}\{\beta w, \{\vec{\alpha}\cdot p, \mathcal{F}\}\} + \frac{1}{4}\left(-\mathrm{i}\,\vec{\alpha}\cdot\vec{f} + 2\,\mathcal{F}\,\vec{\alpha}\cdot\vec{p}\right)\left(-\mathrm{i}\,\vec{\alpha}\cdot\vec{f} + 2\,\mathcal{F}\,\vec{\alpha}\cdot\vec{p}\right) \\
= & \, m^2 w^2 + \frac{m}{2}\{\beta w, \{\vec{\alpha}\cdot p, \mathcal{F}\}\} \\
& + \frac{1}{4}\left(-f^2 + 2\mathcal{F}(\vec{\nabla}\cdot\vec{f}) + 4\mathcal{F}\vec{\Sigma}\cdot(\vec{\mathcal{F}}\times\vec{p}) + \mathcal{F}\vec{p}^{\,2}\mathcal{F}\right) \\
= & \, m^2 w^2 + \frac{m}{2}\{\beta w, \{\vec{\alpha}\cdot p, \mathcal{F}\}\} + \frac{1}{2}\mathcal{F}(\vec{\nabla}\cdot\vec{f}) + \mathcal{F}\vec{\Sigma}\cdot(\vec{f}\times\vec{p}) + \mathcal{F}\vec{p}^{\,2}\mathcal{F} - \frac{1}{4}f^2\,.
\end{aligned}
\tag{10.40}
$$

We now focus on the anticommutator term. i.e.,

$$
\begin{aligned}
\{\beta w, \{\vec{\alpha}\cdot\vec{p}, \mathcal{F}\}\} = & \{\beta w, 2\mathcal{F}\vec{\alpha}\cdot\vec{p}\} - \cancel{\mathrm{i}\{\beta w, \vec{\alpha}\cdot\vec{f}\}} = \beta w(2\mathcal{F}\vec{\alpha}\cdot\vec{p}) + (2\mathcal{F}\vec{\alpha}\cdot\vec{p})\beta w \\
= & \, 2\beta w\mathcal{F}\vec{\alpha}\cdot\vec{p} - 2\beta\mathcal{F}\vec{\alpha}\cdot\vec{p}w = 2\beta\mathcal{F}(w\vec{\alpha}\cdot\vec{p} - w\vec{\alpha}\cdot\vec{p} - \vec{\alpha}\cdot(\vec{p}w)) \\
= & \, 2\mathrm{i}\mathcal{F}\beta\vec{\alpha}\cdot\vec{\phi} = -2\mathcal{F}\vec{\Sigma}\cdot\mathrm{i}\gamma^5\beta\vec{\phi} = -2\mathcal{F}\vec{\Sigma}\cdot J\vec{\phi}\,.
\end{aligned}
\tag{10.41}
$$

Note that we have been using the definitions

$$
\vec{f} = \vec{\nabla}\mathcal{F}\,, \qquad \vec{\phi} = \vec{\nabla}w\,. \tag{10.42}
$$

Plugging our result for the double anticommutator into our expression for H_{DS}^2

$$
H_{\mathrm{DS}}^2 = m^2 w^2 + \mathcal{F}\vec{p}^{\,2}\mathcal{F} + \frac{1}{2}\mathcal{F}(\vec{\nabla}\cdot\vec{f}) - \frac{1}{4}f^2 + \mathcal{F}\vec{\Sigma}\cdot([\vec{f}\times\vec{p}] - Jm\vec{\phi})\,. \tag{10.43}
$$

From which we calculate

$$\sqrt{H_{\rm DS}^2} \approx mw + \frac{1}{4m}(v^{-1}\vec{p}^{\,2}F + F\vec{p}^{\,2}v^{-1}) + \frac{1}{2mv}(\vec{\nabla}\cdot\vec{f}) - \frac{1}{8mw}f^2 + \frac{1}{4m}\vec{\Sigma}\cdot(\{v^{-1},[\vec{f}\times\vec{p}]\}) - 2Jv^{-1}m\vec{\phi})\,, \tag{10.44}$$

where we have used the fact that $w^{-1}\,\mathcal{F} = v^{-1}$. The only term in this equation that doesn't commute with β is J, which anticommutes with β. As such

$$\left[\sqrt{H_{\rm DS}^2}\right]_{\rm even}\beta = \beta\left(mw + \frac{1}{4m}(v^{-1}\vec{p}^{\,2}\mathcal{F} + \mathcal{F}\vec{p}^{\,2}v^{-1}) + \frac{1}{2mv}(\vec{\nabla}\cdot\vec{f}) - \frac{1}{8mw}f^2 + \frac{1}{4m}\vec{\Sigma}\cdot(\{v^{-1},[\vec{f}\times\vec{p}]\})\right)\,, \tag{10.45}$$

and

$$\left\{\sqrt{H_{\rm DS}^2}\right\}_{\rm odd} J = -\frac{1}{2m}\vec{\Sigma}\cdot\vec{\phi}\,v^{-1}\,. \tag{10.46}$$

we can now begin identifying all the terms so that our Hamiltonian can be written in terms of the Schwarzschild radius. We begin by identifying

$$\vec{\phi} = \frac{r_{\rm s}}{2r^3}\vec{r}\,, \quad \vec{f} = \frac{r_{\rm s}}{r^3}\vec{r}\,, \quad w^{-1} \approx 1 + \frac{r_{\rm s}}{2r} = v\,, \quad v^{-1} \approx 1 - \frac{r_{\rm s}}{2r} = w\,, \tag{10.47}$$

which we then use to simplify the remaining terms, keeping terms only to the first order in gravity,

$$mw \approx m - m\frac{r_{\rm s}}{2r}\,, \quad v^{-1}\,\vec{p}^{\,2}\,\mathcal{F} + \mathcal{F}\,\vec{p}^{\,2}\,v^{-1} \approx 2\vec{p}^{\,2}\,, \tag{10.48a}$$

$$v^{-1}(\vec{\nabla}\cdot\vec{f}) \approx -\vec{\nabla}^2\frac{r_{\rm s}}{r} = 4\pi\,r_{\rm s}\,\delta^{(3)}(\vec{r})\,, \quad f^2 \approx 0\,, \tag{10.48b}$$

$$\vec{\Sigma}\cdot\vec{\phi}\,v^{-1} \approx \frac{r_{\rm s}}{2r^3}\vec{\Sigma}\cdot\vec{r}\,, \quad \vec{\Sigma}\cdot(v^{-1}[\vec{f}\times\vec{p}] + [\vec{f}\times\vec{p}]v^{-1}) \approx 2\frac{r_{\rm s}}{r^3}\vec{\Sigma}\cdot\vec{L}\,. \tag{10.48c}$$

Putting this all together, we obtain the result

$$H_{\mathrm{DS}}^{(\mathrm{CFW})} = \beta\left(m + \frac{\vec{p}^{\,2}}{2m}\right) - \beta\frac{m\, r_{\mathrm{s}}}{2r} + \beta\frac{2\pi\, r_{\mathrm{s}}}{m}\delta^{(3)}(\vec{r}) + \beta\frac{r_{\mathrm{s}}}{2m}\frac{\vec{\Sigma}\cdot\vec{L}}{r^3} - \frac{r_{\mathrm{s}}}{4m}\frac{\vec{\Sigma}\cdot\vec{r}}{r^3}\,. \quad (10.49)$$

Yet again we find that the result obtained by the chiral transform differs significantly from that obtained using the standard method (see equation (9.76)). While the chiral result does contain similar terms, including the terms for the free particle and the gravitational potential, the prefactors for the gravitational zitterbewegung term and the Fokker precession do not match. Furthermore, the particle–antiparticle symmetry which exists in the standard result, ensuring that both particles and antiparticles interact with gravity in the same way, is broken by the final term. Additionally, as we have seen before, the final term breaks parity.

10.7. DIRAC HAMILTONIAN IN A ROTATING NON-INERTIAL REFERENCE FRAME

We will now use the chiral Foldy–Wouthuysen transformation on the Dirac Hamiltonian in the non-inertial reference frame. Recall that the Hamiltonian is given by (see (8.90)),

$$\begin{aligned} H_{\mathrm{NF}} &= (1 + \vec{a}\cdot\vec{r})\,\beta\, m + \frac{1}{2}\{1 + \vec{a}\cdot\vec{r}, \vec{\alpha}\cdot\vec{p}\} - \vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right) \\ &= \beta\, m + \beta\, m\,\vec{a}\cdot\vec{r} + \vec{\alpha}\cdot\vec{p} + \frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\} - \vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right)\,. \end{aligned} \quad (10.50)$$

As usual we begin the chiral transformation by first squaring the Hamiltonian. As with the traditional Foldy–Wouthuysen transformation, we are only going to keep the acceleration and angular rotation frequency ($\vec{a}$ and $\vec{\omega}$) to the first order. The canceled out terms are of high enough order that we can approximate them to be zero.

$$
\begin{aligned}
H_{\rm NF}^2 &= \beta\, m\left(\beta\, m + \beta\, m\, \vec{a}\cdot\vec{r} + \vec{\alpha}\cdot\vec{p} + \frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\} - \vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right)\right) \\
&\quad + \beta\, m\, \vec{a}\cdot\vec{r}\left(\beta\, m + \cancel{\beta m \vec{a}\cdot\vec{r}} + \vec{\alpha}\cdot\vec{p} + \cancel{\frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}} - \cancel{\vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right)}\right) \\
&\quad + \vec{\alpha}\cdot\vec{p}\left(\beta\, m + \beta\, m\, \vec{a}\cdot\vec{r} + \vec{\alpha}\cdot\vec{p} + \frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\} - \vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right)\right) \\
&\quad + \frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}\left(\beta\, m + \cancel{\beta m \vec{a}\cdot\vec{r}} + \vec{\alpha}\cdot\vec{p} + \cancel{\frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}} - \cancel{\vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right)}\right) \\
&\quad - \vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right)\left(\beta\, m + \cancel{\beta m \vec{a}\cdot\vec{r}} + \vec{\alpha}\cdot\vec{p} + \cancel{\frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}} - \cancel{\vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right)}\right) \\
&= m^2 + m^2\vec{a}\cdot\vec{r} + \cancel{\beta m \vec{\alpha}\cdot\vec{p}} + \cancel{\frac{\beta\, m}{2}\{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}} - \beta\, m\, \vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right) \\
&\quad + m^2\vec{a}\cdot\vec{r} + \beta\, m\, \vec{a}\cdot\vec{r}\, \vec{\alpha}\cdot\vec{p} - \cancel{\beta m \vec{\alpha}\cdot\vec{p}} - \beta\, m\, \vec{\alpha}\cdot\vec{p}\, \vec{a}\cdot\vec{r} + \vec{p}^{\,2} \\
&\quad + \frac{1}{2}\vec{\alpha}\cdot\vec{p}\{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\} - \vec{\alpha}\cdot\vec{p}\,\vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right) - \cancel{\frac{\beta\, m}{2}\{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}} \\
&\quad + \frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}\vec{\alpha}\cdot\vec{p} - \beta\, m\, \vec{\omega}\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right) - \cancel{\vec{\alpha}\cdot\vec{p}\,\vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right)} \\
&= m^2 + 2\, m^2\, \vec{a}\cdot\vec{r} - 2\,\beta\, m\, \vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right) + \beta\, m[\vec{a}\cdot\vec{r}, \vec{\alpha}\cdot\vec{p}] \\
&\quad + \vec{p}^{\,2} + \frac{1}{2}\{\vec{\alpha}\cdot\vec{p}, \{\vec{\alpha}\cdot\vec{p}, \vec{a}\cdot\vec{r}\}\}\,. \qquad (10.51)
\end{aligned}
$$

Using equation (9.98), as well as the result that

$$
[\vec{a}\cdot\vec{r}, \vec{\alpha}\cdot\vec{p}] = \vec{a}\cdot\vec{r}\,\vec{\alpha}\cdot\vec{p} - \vec{\alpha}\cdot\vec{p}\,\vec{a}\cdot\vec{r} = \mathrm{i}\,\vec{a}\cdot\vec{\alpha}\,, \qquad (10.52)
$$

we can simplify the squared Hamiltonian to,

$$
\begin{aligned}
H_{\rm NF}^2 &= m^2(1 + 2\vec{a}\cdot\vec{r}) + \frac{1}{2}\{1 + 2\,\vec{a}\cdot\vec{r}, \vec{p}^{\,2}\} + \vec{\Sigma}\cdot(\vec{a}\times\vec{p}) + \mathrm{i}\beta\, m\,\vec{\alpha}\cdot\vec{a} \\
&\quad - 2\beta\, m\,\vec{\omega}\cdot\left(\vec{L} + \frac{1}{2}\vec{\Sigma}\right)\,. \qquad (10.53)
\end{aligned}
$$

Then the expanded square root is

$$\sqrt{H_{\rm NF}^2} \approx m(1+\vec{a}\cdot\vec{r}) + \frac{\vec{p}^{\,2}}{2\,m} + \frac{1}{2\,m}\{\vec{p}^{\,2},\vec{a}\cdot\vec{r}\} + \frac{1}{2\,m}\vec{\Sigma}\cdot(\vec{a}\times\vec{p}) \\ + \frac{{\rm i}\,\beta}{2}\vec{\alpha}\cdot\vec{a} - \beta\,\vec{\omega}\cdot\left(\vec{L}+\frac{1}{2}\vec{\Sigma}\right)\,, \tag{10.54}$$

from which we find,

$$\left[\sqrt{H_{\rm NF}^2}\right]_{\rm even}\beta = \beta\left(m + \frac{\vec{p}^{\,2}}{2\,m} + \vec{a}\cdot\vec{r}\,m + \frac{1}{2\,m}\{\vec{p}^{\,2},\vec{a}\cdot\vec{r}\} + \frac{1}{2\,m}\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\right)\,, \\ -\vec{\omega}\cdot\left(\vec{L}+\frac{1}{2}\vec{\Sigma}\right) \\ \left\{\sqrt{H_{\rm NF}^2}\right\}_{\rm odd} J = \left(\frac{{\rm i}\,\beta}{2}\vec{\alpha}\cdot\vec{a}\right)\,J = -\frac{1}{2}\vec{\Sigma}\cdot\vec{a}\,. \tag{10.55}$$

Here we used the following the fact that $\vec{\alpha}\cdot\vec{p}$ and $\vec{\omega}\cdot\vec{J}$ commute (see (8.89), as well as the relations

$$ {\rm i}\beta\vec{\alpha}\cdot\vec{a}J = {\rm i}\,\gamma^0\gamma^0\gamma^I a^I\,{\rm i}\,\gamma^5\gamma^0 = -\gamma^I\gamma^5\gamma^0 a^I = -\gamma^5\gamma^0\gamma^I a^I = -\vec{\Sigma}\cdot\vec{a}\,, \tag{10.56}$$

$$\vec{\alpha}\cdot\vec{p}\,J = \gamma^0\gamma^I p^I\,{\rm i}\,\gamma^5\gamma^0 = -{\rm i}\,\gamma^0\gamma^5\gamma^0\gamma^I p^I = -{\rm i}\,\beta\,\vec{\Sigma}\cdot\vec{p}. \tag{10.57}$$

Furthermore, $[\gamma^5,\vec{\Sigma}]=0$, from which it follows

$$\left[\vec{\Sigma}\cdot\vec{p},\left(\vec{L}+\frac{1}{2}\vec{\Sigma}\right)\right] = 0\,. \tag{10.58}$$

We then obtain the chiral Foldy–Wouthuysen transformation of the Dirac Hamiltonian in a rotating non–inertial reference frame as

$$H_{\rm NF}^{\rm (CFW)} = \beta\left(m + \frac{\vec{p}^{\,2}}{2\,m} + m\,\vec{a}\cdot\vec{r} + \frac{1}{2\,m}\{\vec{p}^{\,2},\vec{a}\cdot\vec{r}\} + \frac{1}{2\,m}\vec{\Sigma}\cdot(\vec{a}\times\vec{p})\right) \\ -\frac{1}{2}\vec{\Sigma}\cdot\vec{a} - \vec{\omega}\cdot\left(\vec{L}+\frac{1}{2}\vec{\Sigma}\right)\,. \tag{10.59}$$

Again we see a number of familiar terms from the standard method (see (9.108)), however the prefactors tend not to match. Additionally there is a spurious parity breaking term $(\vec{\sigma} \cdot \vec{a})$. Surprisingly, the Mashhoon term seems to be unaffected to this order. Even so it is clear that the chiral transformation has failed, yet again, to reproduce the result obtained using the standard method (9.108), the latter being obtained after a more tedious calculation than the deceptively easy chiral method would otherwise require.

10.8. ON THE VIOLATION OF PARITY

In the Dirac representation, parity is somewhat more complicated than in non–relativistic quantum mechanics. In the latter, the parity transform is simply $\mathcal{P} : x \to -x$ (and consequently $\vec{p} \to -\vec{p}$). In the Dirac representation we must additionally exchange the left–hand and right–hand spinors, i.e., $\mathcal{P}_D : \psi_\pm(x) \to \psi_\mp(-x)$, and we can write the Dirac parity operator as $\mathcal{P}_D = \gamma^0 \mathcal{P}$, while stressing that $\mathcal{P}_D$ is the *Dirac* parity operator, and $\mathcal{P}$ serves only change the coordinates. Then, for example, we can show that the free Dirac Hamiltonian is invariant under parity,

$$\begin{aligned} \mathcal{P}_D \, H_{\rm F} \, \mathcal{P}_D^{-1} &= \gamma^0 \mathcal{P} \, \left(\vec{\alpha} \cdot \vec{p} + \beta \, m\right) \mathcal{P} \gamma^0 = \gamma^0 \left(-\vec{\alpha} \cdot \vec{p} + \beta \, m\right) \gamma^0 \\ &= \vec{\alpha} \cdot \vec{p} + \beta \, m = H_{\rm F} \, . \end{aligned} \tag{10.60}$$

Indeed, using this methodology, we can show that all the initial Hamiltonians, i.e., *before they are transformed*, are parity invariant. Consequently, under a parity transformation $\Lambda \to \Lambda$, and one can trivially show that $J \to -J$, where Λ and J are defined as in (10.3). We then find that the chiral rotation as defined in (10.1) and (10.2) is *not* parity invariant, i.e.,

$$\begin{aligned}
U_1^{(\mathcal{P})} &= \mathcal{P}_D\, U_1\, \mathcal{P}_D^{-1} = \frac{1}{\sqrt{2}}\,(1 - J\Lambda) \neq U_1\,, \\
U_2^{(\mathcal{P})} &= \mathcal{P}_D\, U_2\, \mathcal{P}_D^{-1} = \frac{1}{\sqrt{2}}\,(1 - \beta J) \neq U_2\,, \\
U^{(\mathcal{P})} &= \mathcal{P}_D\, U\, \mathcal{P}_D^{-1} = \mathcal{P}_D\, U_2\, \mathcal{P}_D^{-1}\, \mathcal{P}_D\, U_1 \mathcal{P}_D^{-1} \\
&= U_2^{(\mathcal{P})} U_1^{(\mathcal{P})} = \frac{1}{2}\,(1 - \beta J - J\Lambda + \beta\Lambda) \neq U\,,
\end{aligned} \tag{10.61}$$

where we have used the superscript "$(\mathcal{P})$" to indicate that the operator is transformed under parity. If we let $H' = U\, H\, U^{+}$, where H fulfills (10.4), and is parity invariant, then

$$\begin{aligned}
\mathcal{P}_D\, H'\, \mathcal{P}_D^{-1} &= \mathcal{P}_D\, U\, H\, U^{+}\, \mathcal{P}_D^{-1} = \mathcal{P}_D\, U\, \mathcal{P}_D^{-1}\, \mathcal{P}_D\, H\, \mathcal{P}_D^{-1}\, \mathcal{P}_D\, U^{+} \mathcal{P}_D^{-1} \\
&= U^{(\mathcal{P})} H \left(U^{(\mathcal{P})}\right)^{+} = U_2^{(\mathcal{P})} U_1^{(\mathcal{P})}\, H \left(U_1^{(\mathcal{P})}\right)^{+} \left(U_2^{(\mathcal{P})}\right)^{+} \\
&= \frac{1}{2} U_2^{(\mathcal{P})}\,(1 - J\Lambda)\, H\, (1 - \Lambda J) \left(U_2^{(\mathcal{P})}\right)^{+} \\
&= \frac{1}{2} U_2^{(\mathcal{P})}\,(\cancel{H} - J\Lambda H - H\Lambda J + \cancel{J\Lambda H\Lambda J}) \left(U_2^{(\mathcal{P})}\right)^{+} \\
&= U_2^{(\mathcal{P})}\,(-J\Lambda H) \left(U_2^{(\mathcal{P})}\right)^{+} = -\frac{1}{2}\,(1 - \beta J)\, J\Lambda H\, (1 - J\beta) \\
&= -\frac{1}{2}\,(J\Lambda H - \beta\Lambda H - J\Lambda H J\beta + \beta\Lambda H J\beta) \\
&= \frac{1}{2}\left(\sqrt{H^2} + \beta\sqrt{H^2}\beta\right)\beta - \frac{1}{2}\left(\sqrt{H^2} - \beta\sqrt{H^2}\beta\right) J \\
&= \left[\sqrt{H^2}\right]_{\text{even}} \beta - \left\{\sqrt{H^2}\right\}_{\text{odd}} J\,.
\end{aligned} \tag{10.62}$$

By comparing this result to equation (10.8), it is clear that the only way the chiral transform will result in a parity invariant Hamiltonian H', is if $\sqrt{H^2}$ does not contain any odd terms, i.e., $\left\{\sqrt{H^2}\right\}_{\text{odd}} = 0$, as is the case for the free Dirac Hamiltonian. Rather than being a condition for the implementation of the chiral method, this demonstrates how parity may be *accidentally* conserved. Since the rotations utilized by the proposed chiral Foldy–Wouthuysen transform do not conserve parity, the physical interpretation of quantum mechanical operators that result are dubious at best.

11. (PARTIAL) CONCLUSIONS

The main results of part II of this thesis can be summarized as follows: (i) We perform the standard Foldy–Wouthuysen transformation on the Dirac–Einstein–Schwarzschild Hamiltonian (9.76), and find that the interpretation is rather straightforward. The mathematical structure is in fact similar to the structure of the transformed Dirac–Coulomb Hamiltonian, which has been extensively studied [22, 23, 60, 88, 116–118]. The second term is instantly recognizable as the gravitational potential $(-GmM/r)$, when the definition of $r_{\rm s}$ is applied. We additionally find a gravitational version of the zitterbewegung term, as well as a gravitational spin–orbit coupling term, otherwise known as Fokker precession. The Fokker precession is in full agreement with the classical result [115], which in turn has been confirmed by Gravity Probe B [119]. Additionally there is particle–antiparticle symmetry, ensuring that both particles and antiparticles are affected the same by gravity. This is of course in contrast to the Dirac–Coulomb case, in which fields that attract particles repel antiparticles. This is to be expected, as particles and antiparticles have opposite charges, while their masses are identical, both in terms of the gravitational and inertial mass. (ii) We then find the corrections, to the fourth order in momenta, of a Dirac Hamiltonian with a scalar potential. Both the particles and antiparticles are attracted to the center (see equation (9.48)). We find a surprising $\{\vec{p}^{\,2}, 1/r\}$ term, despite the similarities to the Dirac–Coulomb Hamiltonian, which is absent of any such term. (iii) We also calculate the relativistic corrections, again to the fourth order in momenta, of the Dirac Hamiltonian with a scalar confining potential (9.58). This transformed Hamiltonian also exhibits an anticommutator as a kinetic correction, i.e., $\{\vec{p}^{\,2}, r\}$. (iv) We additionally perform the standard transformation on the Dirac Hamiltonian in a rotating non-inertial frame, finding a compact representation up to the fourth order in momenta (9.108), and verify that the Mashhoon term [103] is unaltered up to the fourth order in momenta. It is also worth noting that as with the Dirac–Einstein–Schwarzschild Hamiltonian, all three of these Dirac Hamiltonians, (9.48), (9.58), and (9.108), have particle–antiparticle symmetry, thus both the particles and antiparticles behave identically in these potentials. (v) Finally we apply the rotations used to transform the Dirac–Einstein–Schwarzschild Hamiltonian to the gravitationally coupled transition current (9.125), finding that in addition to the known corrections there is an additional gravitational kinetic correction, as well as gravitational corrections to the magnetic coupling.

While the Dirac Hamiltonian in a rotating non-inertial frame is an interesting equation in its own right, it also serves to contrast the rest of the performed transformations, which otherwise only require two sweeps of the Foldy–Wouthuysen program, as an example of a more complicated Hamiltonian, requiring three iterations of the standard Foldy–Wouthuysen program for the calculation. This level of complexity illustrates why there is a desire to find a simpler, and possibly exact, methodology to reveal the relativistic corrections. One such proposal is the chiral Foldy–Wouthuysen transform [24], which seeks to simplify the procedure by decoupling the particle and antiparticle degrees of freedom utilizing some deceptively appealing properties of Dirac algebra. Rather than including the need to iterate the procedure, it instead seeks to perform the transformation in a single series of well defined steps, approximating $\sqrt{H^2}$ (where H is the Dirac Hamiltonian we wish to decouple) using a Taylor series expansion about small momenta. While the procedure seems promising, quickly, if accidentally, solving for the free Dirac Hamiltonian, it falls apart when applied to other Hamiltonians. Chief among these failures is the Dirac–Coulomb Hamiltonian, which does not meet the requirements imposed by the chiral method. It is not a good sign when a proposed procedure fails to work on one of the most important examples of a generalized Dirac Hamiltonian.

Setting aside the failings of the chiral transform, in terms of its limited applicability, the procedure fails to produce results in agreement with a number of Hamiltonians transformed via the standard method. This is somewhat surprising, as one would surmise that given a Hamiltonian which fulfills the condition $\{H, J\} = 0$, the resulting transformation is unitary, as is that of the standard Foldy–Wouthuysen transform. If both transforms are unitary, why then are we getting conflicting results? The answer lies in the definition of the chiral rotation U (10.1), which breaks parity and alters one of the fundamental symmetries of the Hamiltonian it is applied to. As a simple example of a unitary transform that breaks parity, let us consider the unitary transformation $U = \exp(\mathrm{i}\,\vec{A}\cdot\vec{r})$, where $\vec{A}$ is a constant vector, applied to the free Schrödinger Hamiltonian $H = \vec{p}^{\,2}/(2m)$. Then $H' = U\,H\,U^{+} = (\vec{p}-\vec{A})^2/(2m)$, which has a term proportional to $\vec{A}\cdot\vec{p}$, which breaks parity. Thus a parity breaking term is introduced by a unitary transformation which does not conserve parity. Furthermore, as shown in [88], the chiral method also changes the physical interpretation of the spin matrix $\vec{\Sigma}$. Thus, despite the elegance and simplicity of the chiral method, the more complicated, and more widely applicable, standard Foldy–Wouthuysen transformation is the more reliable choice when finding the relativistic corrections resulting from generalized Dirac Hamiltonians.

Part III

Pseudo–Hermiticity and Ultrarelativistic Decoupling

12. INTRODUCTION

In the current part of the thesis we strive to unify the approaches of parts I and II, by answering the question: Are there pseudo–Hermitian variants of Dirac–Hamiltonians whose eigenvalues can be approximated using a decoupling transform? As it turns out, superluminal particles [25] constitute a class of pseudo–Hermitian Dirac particles [43]. Unfortunately, the Foldy–Wouthuysen program, which we used to great effect in chapter 9, is used to find the nonrelativistic limit, and is therefore unsuited to deal with tachyons (superluminal particles). Inspired by the Foldy–Wouthuysen transformation, we develop an expansion which follows a fundamentally different paradigm, in which the particles and antiparticles are decoupled in the ultrarelativistic limit [93]. While the Foldy–Wouthuysen transformation is performed in the limit where the mass term dominates the kinetic and potential terms, the ultrarelativistic decoupling transform is performed in the limit where the kinetic term dominates. Such a transform is best described in the helicity basis (see chapter 23 of [121]). This is especially clear when considering the massless limit for the cases of the free Dirac Hamiltonians (both sub- and superluminal), which approach the Weyl equation (chapter 2.4.2 of [60]), which is known to describe spin-1/2 particles traveling exactly at the speed of light.

As illustrated in figure 12.1 neither sub- nor superluminal particles can break the light–speed barrier, i.e., tachyons are forbidden from slowing down to, or below, c, as infinite energy would be required to slow down to c. One might be inclined to wonder if the usual Foldy–Wouthuysen program might be used to find the corrections to the “low energy limit” (which would correspond to the nonrelativistic limit of a tardyon) of a tachyon. Again, by considering figure 12.1, we realize that as the energy of a superluminal particle goes to zero, its speed goes to infinity. Thus the

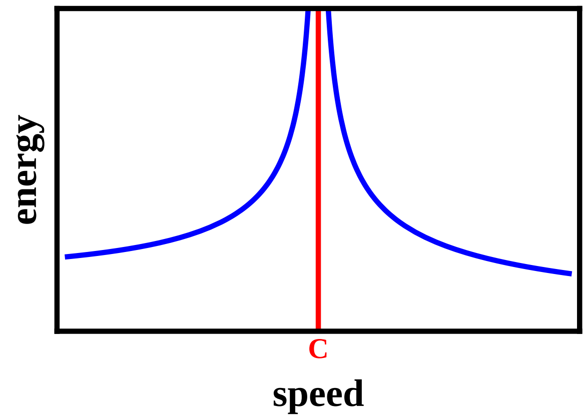


Figure 12.1: Here we present a plot of the energy vs. speed of a free relativistic particle, for both the sub- (to the left of c) and superluminal (to the right of c) case. When $v < c$ we use the equation $E = m\,c^2/\sqrt{1-(v/c)^2}$, while when $v > c$ we use the equation $E = m\,c^2/\sqrt{(v/c)^2-1}$. In both cases, as the speed of the particle approaches the speed of light (c), its energy goes to infinity, indicating that there is a speed barrier at c which the particle cannot cross. Furthermore, we note that for the superluminal case, as the energy approaches zero, the speed goes to infinity.

"low energy" limit is manifestly non–physical. Thus we are left to consider how to decouple the particle and antiparticle degrees of freedom in the "high energy," i.e., ultrarelativistic, limit.

This part is organized as follows: In chapter 13 we briefly discuss the free tachyonic Dirac equation, and in chapter 14 we investigate the pseudo–Hermitian character of said particles. In chapter 15 we generalize the tachyonic Dirac equation to include curved space, specifically the tachyonic Dirac equation coupled to a gravitational center. In chapter 16 we draw inspiration from our work with the Foldy–Wouthuysen program and construct an exact ultrarelativistic decoupling transform for the free tardyon and tachyon. Finally, in chapter 17 we investigate the perturbative version of the ultrarelativistic decoupling transformation, and find a somewhat surprising result concerning tachyons in curved space. Lastly, some conclusions are drawn in chapter 18.

13. FREE TACHYONIC DIRAC EQUATION

While we are working with units such that $c = 1$, for the first part of this chapter we write c into our equations explicitly, as it will serve to enhance the discussion. As discussed in appendix C, the Dirac equation is a result of the liniarization (with respect to the time derivative) of the Klein–Gordon equation. The Klein–Gordon equation is arrived at using the identification that, when moving from the classical to the quantum level, $E \to \mathrm{i}\partial_t$ and $\vec{p} \to -\mathrm{i}\vec{\nabla}$. This is of course contingent on the dispersion relation $E^2 = c^2\vec{p}^{\,2} + c^4 m_1^2$. Furthermore, this equation is Lorentz invariant, meaning that we can perform a Lorentz boost such that the speed of the particle exceeds the speed of light ($v > c$), and retain the same dispersion relation. On the surface, it would appear that there is no discernible difference between the dispersion relation for sub– and superluminal particles, however upon closer examination it becomes clear that there is in fact a difference.

On the classical level, the relativistic spatial momentum is given as

$$\vec{p} = \frac{m_1\,\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}}\,. \tag{13.1}$$

We then plug this into our dispersion relation to obtain

$$E^2 = c^2 \frac{m_1^2\,v^2}{1 - \frac{v^2}{c^2}} + c^4\,m_1 = c^4\,m_1^2 \left(\frac{\frac{v^2}{c^2}}{1 - \frac{v^2}{c^2}} + 1 \right) = \frac{m_1^2\,c^4}{1 - \left(\frac{v}{c}\right)^2}\,, \tag{13.2}$$

giving us the equation energy:

$$E = \frac{m_1 c^2}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}\,. \tag{13.3}$$

This result applies to both tardyons and tachyons. In the tardyonic case ($v < c$) the denominator is real, and the implicit assumption that the mass is real holds. On the other hand, for the tachyonic case ($v > c$), we find that the denominator is imaginary. This creates a bit of a problem, as we require that the energy is real valued. The only way to resolve this discrepancy is the relation that for a superluminal object, the mass term is imaginary, i.e.,

$$m_1 = \mathrm{i}\, m \tag{13.4}$$

where m is real. With this realization in hand, one method of obtaining the superluminal Dirac equation is to simply recall the subluminal Dirac Hamiltonian (C.34),

$$H = \vec{\alpha} \cdot \vec{p} + \beta\, m \,,$$

and have $m \to \mathrm{i}\, m$, yielding

$$H_1 = \vec{\alpha} \cdot \vec{p} + \mathrm{i}\, \beta\, m \,, \tag{13.5}$$

which results in a perfectly acceptable result. However, this is not the only possible result. Instead, let us consider the dispersion relation in the context of (13.4), i.e., the tachyonic dispersion relation is

$$E^2 = \vec{p}^{\,2} - m^2 \,. \tag{13.6}$$

By taking the square root of (13.6) we arrive at the Lorentz invariant equation

$$E = \sqrt{\vec{p}^{\,2} - m^2} \,. \tag{13.7}$$

When we shift to the quantum level, equation (13.6) becomes

$$\left(\partial_t^2 - \vec{\nabla}^2 - m^2\right)\phi(t,\vec{r}) = 0\,, \tag{13.8}$$

the superluminal Klein–Gordon equation, which can be rewritten as

$$-\partial_t^2\phi(t,\vec{r}) = \left(-\vec{\nabla}^{\,2} - m^2\right)\phi(t,\vec{r}) = \left(\vec{p}^{\,2} - m^2\right)\phi(t,\vec{r})\,, \tag{13.9}$$

which we can interpret as

$$H_{\rm i}^2 = \vec{p}^{\,2} - m^2\,. \tag{13.10}$$

It is simple to check that the Dirac Hamiltonian given in (13.5) satisfies this definition,

$$\begin{aligned} H_1^2 &= (\vec{\alpha}\cdot\vec{p} - {\rm i}\,\beta\,m)\,(\vec{\alpha}\cdot\vec{p} - {\rm i}\,\beta\,m) = (\vec{\alpha}\cdot\vec{p})^2 - {\rm i}\,m\,\{\vec{\alpha},\beta\}\cdot\vec{p} + {\rm i}^2\,\beta^2\,m^2 \\ &= \vec{p}^{\,2} - m^2 = H_{\rm i}^2\,, \end{aligned} \tag{13.11}$$

where we used the identities $(\vec{\alpha}\cdot\vec{p})^2 = \vec{p}^{\,2}$, $\{\alpha^I,\beta\} = 0$ and $\beta^2 = 0$ (see appendix C). We also find that the Dirac Hamiltonian

$$H_2 = \vec{\alpha}\cdot\vec{p} + \beta\,\gamma^5\,m\,, \tag{13.12}$$

satisfies (13.10) as well,

$$\begin{aligned} H_2^2 &= \left(\vec{\alpha}\cdot\vec{p} + \beta\,\gamma^5\,m\right)\left(\vec{\alpha}\cdot\vec{p} + \beta\,\gamma^5\,m\right) = (\vec{\alpha}\cdot\vec{p})^2 + \left\{\vec{\alpha},\beta\,\gamma^5\right\}\cdot\vec{p}\,m + \left(\beta\,\gamma^5\right)^2\,m^2 \\ &= \vec{p}^{\,2} - m^2 = H_{\rm i}^2\,. \end{aligned} \tag{13.13}$$

To obtain this result we again used $(\vec{\alpha}\cdot\vec{p})^2 = \vec{p}^{\,2}$, as well as

$$\left\{\alpha^I, \beta\,\gamma^5\right\} = \left\{\gamma^0\,\gamma^I, \gamma^0\,\gamma^5\right\} = \gamma^0\,\gamma^I\,\gamma^0\,\gamma^5 + \gamma^0\,\gamma^5\,\gamma^0\,\gamma^5 = \gamma^5\,\gamma^I - \gamma^5\,\gamma^I = 0\,, \quad (13.14)$$

and

$$\left(\beta\,\gamma^5\right)^2 = \gamma^0\,\gamma^5\,\gamma^0\,\gamma^5 = -1\,, \quad (13.15)$$

which stems from the relations $\{\gamma^0\,,\gamma^5\} = 0$ and $(\gamma^0)^2 = (\gamma^5)^2 = 1$ Thus we are left with two viable options for the superluminal free Dirac Hamiltonian. These Hamiltonians correspond to the Dirac equations

$$\left(\mathrm{i}\gamma^A\partial_A - \mathrm{i}\,m\right)\phi_1(x) = 0\,, \quad \left(\mathrm{i}\gamma^A\partial_A - \gamma^5\,m\right)\phi_2(x) = 0\,, \quad (13.16)$$

where we use the somewhat unconventional notation of Roman characters, keeping in line with chapter 8.5. As shown in appendix A of [43], there exists a transform which takes H_1 in H_2. For our purposes, we will be using H_2 to perform our calculations, and we now define the tachyonic free Dirac Hamiltonian as

$$H_{\mathrm{TF}} = H_2 = \vec{\alpha}\cdot\vec{p} + \beta\,\gamma^5\,m\,. \quad (13.17)$$

14. A PSEUDO–HERMITIAN DIRAC HAMILTONIAN

14.1. ORIENTATION

With the tachyonic free Dirac Hamiltonian in hand, we quickly realize that it is not in fact Hermitian ($H_{\rm TF} \neq H^+_{\rm TF}$) due to the fact that $(\beta\,\gamma^5)^+ = \gamma^5\,\beta = -\beta\,\gamma^5$. Instead we find that the Hamiltonian is actually γ^5–Hermitian [43],

$$\gamma^5\,H^+_{\rm TF}\,\gamma^5 = \gamma^5\left(\vec{\alpha}\cdot\vec{p} + \beta\,\gamma^5\,m\right)^+\gamma^5 = \vec{\alpha}\cdot\vec{p} + \beta\,\gamma^5\,m = H_{\rm TF}\,. \qquad (14.1)$$

In short, the tachyonic free Dirac equation (and as we shall see in chapter 15, the gravitationally coupled tachyonic Dirac equation) are γ^5–Hermitian, and adhere to the same properties as the subluminal pseudo–Hermitian Hamiltonians, as discussed in chapter 2.

Additionally we find that the tachyonic Dirac equation is invariant under both $\mathcal{C}_D\mathcal{P}_D$ and $\mathcal{T}_D$, where $\mathcal{C}_D$ is the Dirac charge conjugation operator, $\mathcal{P}_D$ is the Dirac parity operator, and $\mathcal{T}_D$ is the Dirac time reversal operator. Under these considerations, it follows that the superluminal Dirac equation is $\mathcal{CPT}$–symmetric. Despite the time we spend on showing that the superluminal free Dirac equation is $\mathcal{CPT}$–symmetric, it is far simpler to use γ^5–Hermiticity. This results from the requirement that the parity, time reversal, and charge conjugation operations use the Dirac equation, rather than the Dirac Hamiltonian. When we perform the Foldy–Wouthuysen transformation, we do so using the Hamiltonian form, and we shall again use said form when performing the ultrarelativistic decoupling transformation in chapter 16. Still it may be instructive to delve into the particulars of $\mathcal{CPT}$–symmetry.

While defining all the operators, we will be working with subluminal example cases.

Note: all Dirac γ matrices appearing in this chapter are flat–space matrices, and we forgo the inclusion of the tilde.

14.2. SOME PROPERTIES

Let us go over some relevant, easily verified, properties of the Dirac γ matrices,

$$\gamma^0 = \left(\gamma^0\right)^T = \left(\gamma^0\right)^+ = \left(\gamma^0\right)^* \,, \quad \gamma^1 = -\left(\gamma^1\right)^T = -\left(\gamma^1\right)^+ = \left(\gamma^1\right)^* \,, \tag{14.2a}$$

$$\gamma^2 = \left(\gamma^2\right)^T = -\left(\gamma^2\right)^+ = -\left(\gamma^2\right)^* \,, \quad \gamma^3 = -\left(\gamma^3\right)^T = -\left(\gamma^3\right)^+ = \left(\gamma^3\right)^* \,, \tag{14.2b}$$

$$\gamma^5 = \left(\gamma^5\right)^T = \left(\gamma^5\right)^+ = \left(\gamma^5\right)^* \,, \tag{14.2c}$$

From these properties we can additionally deduce that

$$\gamma^0 \left(\gamma^A\right)^+ \gamma^0 = \gamma^A \,, \quad \gamma^0 \left(\gamma^5\right)^+ \gamma^0 = -\gamma^5 \,, \tag{14.3}$$

where $A = 0, 1, 2, 3$. We also take this opportunity to define

$$\overline{\psi} \equiv \psi^+ \gamma^0 \,. \tag{14.4}$$

We are now ready to investigate the different operators.

14.3. CHARGE CONJUGATION

There seems to be at least two ways that one can define the charge conjugation operator $\mathcal{C}_D$ (we are using a unconventional notation to distinguish from the nonrelativistic case). The first is the way in which it is defined in [60], while the second is found in [23]. Let us begin with the more involved definition.

It is well know that in relativistic quantum mechanics, we use the covariant coupling $\mathrm{i}\partial_B \rightarrow \mathrm{i}\partial_B - e\,A_B$, where $\vec{A}$ is the vector potential, giving us the electromagnetically coupled Dirac equation [23]

$$\left[\gamma^B \left(\mathrm{i}\partial_B - e\,A_B\right) - m\right] \psi = 0\,. \tag{14.5}$$

Taking the adjoint (transpose and complex conjugation) of this equation we find

$$\psi^+ \left[\left(\gamma^B\right)^+ \left(-\mathrm{i}\overleftarrow{\partial}_B - e\,A_B\right) - m\right] = 0\,. \tag{14.6}$$

Insertion of the γ^0 matrix (multiplying the equation by γ^0 on the right, and using the fact that $\left(\gamma^0\right)^2 = 1$) we obtain

$$\left(\psi^+\gamma^0\right)\gamma^0 \left[\left(\gamma^B\right)^+ \left(-\mathrm{i}\overleftarrow{\partial}_B - e\,A_B\right) - m\right]\gamma^0 = \overline{\psi}\left[\gamma^B \left(-\mathrm{i}\overleftarrow{\partial}_B - e\,A_B\right) - m\right] = 0\,, \tag{14.7}$$

where we used (14.3). We now take the transpose to obtain

$$\left[\left(\gamma^B\right)^T \left(-\mathrm{i}\partial_B - e\,A_B\right) - m\right] \overline{\psi}^{\,T} = 0\,. \tag{14.8}$$

We now introduce the charge conjugation matrix C, with the defining properties

$$C\left(\gamma^B\right)^T C^{-1} = -\gamma^B\,. \tag{14.9}$$

One possible choice for this matrix is $C = \mathrm{i}\gamma^2\gamma^0$. Then $C^{-1} = C^T = \mathrm{i}\gamma^0\gamma^2$, i.e.,

$$C\,C^T = \mathrm{i}\gamma^2\gamma^0\,\mathrm{i}\gamma^0\gamma^2 = -\gamma^2\gamma^2 = 1\,. \tag{14.10}$$

We now verify that this choice of C fulfills (14.8),

$$C\ \left(\gamma^0\right)^T\ C^{-1} = \mathrm{i}\gamma^2\gamma^0\ \left(\gamma^0\right)\ \mathrm{i}\gamma^0\gamma^2 = -\gamma^2\gamma^0\gamma^2 = -\gamma^0\,, \tag{14.11}$$

$$C\ \left(\gamma^1\right)^T\ C^{-1} = \mathrm{i}\gamma^2\gamma^0\ \left(-\gamma^1\right)\ \mathrm{i}\gamma^0\gamma^2 = -\gamma^2\gamma^1\gamma^2 = -\gamma^1\,, \tag{14.12}$$

$$C\ \left(\gamma^2\right)^T\ C^{-1} = \mathrm{i}\gamma^2\gamma^0\ \left(\gamma^2\right)\ \mathrm{i}\gamma^0\gamma^2 = \gamma^2\gamma^2\gamma^2 = -\gamma^2\,, \tag{14.13}$$

$$C\ \left(\gamma^3\right)^T\ C^{-1} = \mathrm{i}\gamma^2\gamma^0\ \left(-\gamma^3\right)\ \mathrm{i}\gamma^0\gamma^2 = -\gamma^2\gamma^3\gamma^2 = -\gamma^3\,. \tag{14.14}$$

Then by multiplying (14.8) on the left by C, we find

$$C\left[\left(\gamma^B\right)^T\left(-\mathrm{i}\partial_B - e\,A_B\right) - m\right]C^{-1}\left(C\overline{\psi}^{\,T}\right) = \left[\gamma^B\left(\mathrm{i}\,\partial_B + e\,A_B\right) - m\right]\psi^{\mathcal{C}} = 0\,, \tag{14.15}$$

where $\psi^{\mathcal{C}} = C\overline{\psi}^{\,T}$. Notice that we have almost recovered the original form of the equation. The difference being that now the charge term is added rather than subtracted, hence "charge conjugation."

Alternately, we begin with the electromagnetically coupled Dirac equation

$$\left[\gamma^B\left(\mathrm{i}\partial_B - e\,A_B\right) - m\right]\psi\,, \tag{14.16}$$

and we complex conjugate the equation, yielding

$$\left[\left(\gamma^B\right)^*\left(-\mathrm{i}\partial_B - e\,A_B\right) - m\right]\psi^*\,, \tag{14.17}$$

and introduce the alternate charge conjugation matrix C_{alt}, which is defined by the property

$$C_{\mathrm{alt}}\ \left(\gamma^B\right)^*\ C_{\mathrm{alt}}^{-1} = -\gamma^B\,. \tag{14.18}$$

A possible choice is $C_{\text{alt}} = \mathrm{i}\gamma^2$. Let us verify this choice:

$$C_{\text{alt}}\left(\gamma^0\right)^* C_{\text{alt}}^{-1} = \mathrm{i}\gamma^2\left(\gamma^0\right)\mathrm{i}\gamma^2 = -\gamma^2\gamma^0\gamma^2 = \gamma^2\gamma^2\gamma^0 = -\gamma^0\,, \tag{14.19a}$$

$$C_{\text{alt}}\left(\gamma^0\right)^* C_{\text{alt}}^{-1} = \mathrm{i}\gamma^2\left(\gamma^1\right)\mathrm{i}\gamma^2 = 0\gamma^2\gamma^1\gamma^2 = \gamma^2\gamma^2\gamma^1 = -\gamma^1\,, \tag{14.19b}$$

$$C_{\text{alt}}\left(\gamma^2\right)^* C_{\text{alt}}^{-1} = \mathrm{i}\gamma^2\left(-\gamma^2\right)\mathrm{i}\gamma^2 = \gamma^2\gamma^2\gamma^2 = -\gamma^2\,, \tag{14.19c}$$

$$C_{\text{alt}}\left(\gamma^3\right)^* C_{\text{alt}}^{-1} = \mathrm{i}\gamma^2\left(\gamma^3\right)\mathrm{i}\gamma^2 = -\gamma^2\gamma^3\gamma^2 = \gamma^2\gamma^2\gamma^3 = -\gamma^3\,. \tag{14.19d}$$

The by multiplying (14.17) on the left by C_{alt} we find

$$C_{\text{alt}}\left[\gamma^B\left(-\mathrm{i}\partial_B - e\,A_B\right) - m\right] C_{\text{alt}}^{-1}\left(C_{\text{alt}}\psi^*\right) = \left[\gamma^B\left(\mathrm{i}\partial_B + e\,A_B\right) - m\right]\psi_{\text{alt}}^{\mathcal{C}} = 0\,, \tag{14.20}$$

where $\psi_{\text{alt}}^{\mathcal{C}} = C_{\text{alt}}\psi^*$. We note that the alternate method seems to produce the same result as the first method, save for a possible difference in the wave–functions. In fact, it is possible to show that the wave–functions which result from each of the methods are equivalent,

$$\begin{aligned}\psi^{\mathcal{C}} &= C\overline{\psi}^{\,T} = \mathrm{i}\gamma^2\gamma^0\left(\psi^+\gamma^0\right)^T = \mathrm{i}\gamma^2\gamma^0\left(\gamma^0\right)^T\left(\psi^+\right)^T \\ &= \mathrm{i}\gamma^2\gamma^0\gamma^0\left(\left(\psi^*\right)^T\right)^T = \mathrm{i}\gamma^2\psi^* = C_{\text{alt}}\psi^* = \psi_{\text{alt}}^{\mathcal{C}}\,.\end{aligned} \tag{14.21}$$

Thus the two different methods lead to the exact same result. We have use two slightly different methods to perform the charge conjugation operation. The first method is rather involved, and requires that we take the adjoint, multiply by γ^0, take the transpose, and finally perform a matrix multiplication. The alternative method seems to simplify this process, and requires only that a matrix multiplication is performed following a complex conjugation. Both seem to be viable options.

14.4. TIME REVERSAL

Again there are two approaches to the time reversal operation. The first is found in [60], while the second may be found in [23].

We begin with the the Dirac equation for a free particle,

$$\left(\mathrm{i}\gamma^B \partial_B - m\right) \psi\left(x\right) = 0\,. \tag{14.22}$$

We then take the adjoint, giving us

$$\psi^+\left(x\right) \left(-\mathrm{i}\left(\gamma^B\right)^+ \overleftarrow{\partial}_B - m\right) = 0\,. \tag{14.23}$$

As we did with the charge conjugation, we multiply the equation by γ^0 on the right, and use the fact that $\left(\gamma^0\right)^2 = 1$, leading to

$$\left(\psi^+\left(x\right)\gamma^0\right)\gamma^0 \left(-\mathrm{i}\left(\gamma^B\right) \overleftarrow{\partial}_B - m\right)\gamma^0 = \overline{\psi}\left(x\right)\left(-\mathrm{i}\gamma^B \overleftarrow{\partial}_B - m\right) = 0\,, \tag{14.24}$$

where we again used (14.3). We now take the transpose to find

$$\left(-\mathrm{i}\left(\gamma^B\right)^T \partial_B - m\right) \overline{\psi}^T\left(x\right) = 0\,. \tag{14.25}$$

We now introduce the time reversal matrix T, which is defined by the properties

$$T\left(\gamma^0\right)^T T^{-1} = \gamma^0\,, \quad T\left(\gamma^i\right)^T T^{-1} = -\gamma^i\,. \tag{14.26}$$

We now explicitly show that a possible choice for T is $T = \mathrm{i}\gamma^5\gamma^2$:

$$T\left(\gamma^0\right)^T T^{-1} = \mathrm{i}\gamma^5\gamma^2\,\left(\gamma^0\right)\mathrm{i}\gamma^2\gamma^5 = -\gamma^5\gamma^2\,\gamma^0\,\gamma^2\gamma^5 = \gamma^0\,, \tag{14.27}$$

$$T\left(\gamma^1\right)^T T^{-1} = \mathrm{i}\gamma^5\gamma^2\,\left(-\gamma^1\right)\mathrm{i}\gamma^2\gamma^5 = \gamma^5\gamma^2\,\gamma^1\,\gamma^2\gamma^5 = -\gamma^1\,, \tag{14.28}$$

$$T\left(\gamma^2\right)^T T^{-1} = \mathrm{i}\gamma^5\gamma^2\,\left(\gamma^2\right)\mathrm{i}\gamma^2\gamma^5 = -\gamma^5\gamma^2\,\gamma^2\,\gamma^2\gamma^5 = -\gamma^2\,, \tag{14.29}$$

$$T\left(\gamma^3\right)^T T^{-1} = \mathrm{i}\gamma^5\gamma^2\,\left(-\gamma^3\right)\mathrm{i}\gamma^2\gamma^5 = -\gamma^5\gamma^2\,\gamma^3\,\gamma^2\gamma^5 = -\gamma^3\,. \tag{14.30}$$

By applying the time reversal matrix T to (14.25) we find

$$T\left(-\mathrm{i}\left(\gamma^B\right)^T\partial_B - m\right)T^{-1}\left(T\overline{\psi}^{\,T}(x)\right) = \left(-\mathrm{i}\gamma^0\partial_0 + \mathrm{i}\gamma^i\partial_i - m\right)\left(T\overline{\psi}^{\,T}(x)\right) = 0\,. \tag{14.31}$$

Finally we let $t \to -t$, i.e., $x = (t, \vec{x}) \to x_{\mathcal{T}} = (-t, \vec{x})$ (this also means that $\partial_0 \to -\partial_0$), yielding

$$\left(\mathrm{i}\gamma^B\partial_B - m\right)\left(T\overline{\psi}^{\,T}(x_{\mathcal{T}})\right) = \left(\mathrm{i}\gamma^B\partial_B - m\right)\psi^{\mathcal{T}}(x)\;, \tag{14.32}$$

where $\psi^{\mathcal{T}}(x) = T\overline{\psi}^{T}(x_{\mathcal{T}})$. Here we find that the Dirac equation retains its form when it undergoes time reversal.

Using the alternate approach, we again start with the free Dirac equation

$$\left(\mathrm{i}\gamma^B\partial_B - m\right)\psi(x) = 0\,, \tag{14.33}$$

and we complex conjugate

$$\left(-\mathrm{i}\left(\gamma^B\right)^*\partial_B - m\right)\psi^*(x) = 0\,. \tag{14.34}$$

We then introduce the alternative time reversal matrix T_{alt} which is defined by the properties

$$T_{\text{alt}} \left(\gamma^0\right)^* T_{\text{alt}}^{-1} = \gamma^0 \,, \quad T_{\text{alt}} \left(\gamma^i\right)^* T_{\text{alt}}^{-1} = -\gamma^i \,. \tag{14.35}$$

One possible choice is $T_{\text{alt}} = \mathrm{i}\gamma^5\gamma^2\gamma^0$. Explicitly:

$$T_{\text{alt}} \left(\gamma^0\right)^* T_{\text{alt}}^{-1} = \mathrm{i}\gamma^5\gamma^2\gamma^0 \left(\gamma^0\right) \mathrm{i}\gamma^0\gamma^2\gamma^5 = -\gamma^5\gamma^2\gamma^0\,\gamma^0\,\gamma^0\gamma^2\gamma^5 = \gamma^0 \,, \tag{14.36}$$

$$T_{\text{alt}} \left(\gamma^1\right)^* T_{\text{alt}}^{-1} = \mathrm{i}\gamma^5\gamma^2\gamma^0 \left(\gamma^1\right) \mathrm{i}\gamma^0\gamma^2\gamma^5 = -\gamma^5\gamma^2\gamma^0\,\gamma^1\,\gamma^0\gamma^2\gamma^5 = -\gamma^1 \,, \tag{14.37}$$

$$T_{\text{alt}} \left(\gamma^2\right)^* T_{\text{alt}}^{-1} = \mathrm{i}\gamma^5\gamma^2\gamma^0 \left(-\gamma^2\right) \mathrm{i}\gamma^0\gamma^2\gamma^5 = \gamma^5\gamma^2\gamma^0\,\gamma^2\,\gamma^0\gamma^2\gamma^5 = -\gamma^2 \,, \tag{14.38}$$

$$T_{\text{alt}} \left(\gamma^3\right)^* T_{\text{alt}}^{-1} = \mathrm{i}\gamma^5\gamma^2\gamma^0 \left(\gamma^3\right) \mathrm{i}\gamma^0\gamma^2\gamma^5 = -\gamma^5\gamma^2\gamma^0\,\gamma^3\,\gamma^0\gamma^2\gamma^5 = -\gamma^3 \,. \tag{14.39}$$

We then apply the alternative time reversal matrix to (14.34),

$$T_{\text{alt}} \left(-\mathrm{i} \left(\gamma^B\right)^* \partial_B - m\right) T_{\text{alt}}^{-1} \left(T_{\text{alt}}\psi^* \left(x\right)\right) = \left(-\mathrm{i}\gamma^0\partial_0 + \mathrm{i}\gamma^i\partial_i - m\right) \left(T_{\text{alt}}\psi^*(x)\right) = 0 \,. \tag{14.40}$$

We again let $x \to x_{\mathcal{T}}$, giving us

$$\left(\mathrm{i}\gamma^B\partial_B - m\right) \psi_{\text{alt}}^{\mathcal{T}}(x) = 0 \,, \tag{14.41}$$

where $\psi_{\text{alt}}^{\mathcal{T}}(x) = T_{\text{alt}}\psi^*(x_{\mathcal{T}})$. Notice that

$$\begin{aligned} \psi^{\mathcal{T}}\left(x\right) &= T\overline{\psi}^{\,T}(x_{\mathcal{T}}) = \mathrm{i}\gamma^5\gamma^2 \left(\psi^+(x_{\mathcal{T}})\gamma^0\right)^T = \mathrm{i}\gamma^5\gamma^2 \,\left(\gamma^0\right)^T \left(\psi^+(x_{\mathcal{T}})\right)^T \\ &= \mathrm{i}\gamma^5\gamma^2\gamma^0 \,\left(\left(\psi^*(x_{\mathcal{T}})\right)^T\right)^T = T_{\text{alt}}\psi^*(x_{\mathcal{T}}) = \psi_{\text{alt}}^{\mathcal{T}}(x) \,. \end{aligned} \tag{14.42}$$

Again the two different methods lead to the exact same result.

14.5. PARITY

We already talked about the parity operation in chapter 10.8 but for completeness, let us apply it to the free Dirac equation,

$$\left(\mathrm{i}\gamma^B\partial_B - m\right)\psi(x) = 0\,. \tag{14.43}$$

Under parity $x = (t, \vec{x}) \to x_{\mathcal{P}} = (t, -\vec{x})$, i.e.,

$$\left(\mathrm{i}\gamma^0\partial_0 - \mathrm{i}\gamma^i\partial_i - m\right)\psi(x_{\mathcal{P}}) = 0\,. \tag{14.44}$$

We then introduce the parity matrix P, defined by the properties

$$P\gamma^0P^{-1} = \gamma^0\,, \quad P\gamma^iP^{-1} = -\gamma^i\,. \tag{14.45}$$

A clear candidate is $P = \gamma^0$. Thus

$$P\left(\mathrm{i}\gamma^0\partial_0 - \mathrm{i}\gamma^i\partial_i - m\right)P^{-1}\left(P\psi(x_{\mathcal{P}})\right) = \left(\mathrm{i}\gamma^B\partial_B - m\right)\psi^{\mathcal{P}}(x) = 0\,, \tag{14.46}$$

where $\psi^{\mathcal{P}}(x) = P\psi(x_{\mathcal{P}})$. As with the time reversal operation, we find that the form of the Dirac equation is unchanged under the parity operation. Unlike the charge conjugation and time reversal operations, there seems to be only the one definition for the Dirac parity operation.

14.6. $\mathcal{C}, \mathcal{P}, \mathcal{T}$ FOR THE TACHYONIC DIRAC HAMILTONIAN

Although we have looked at two separate definitions of both the charge conjugation and time reversal operations, we have seen that both methods lead to the same result, at least in the subluminal case. We will not show it here, but the same result is again achieved when examining the superluminal case. We use the first definition of both in the section. It is easily verified that when we apply the charge conjugation, time reversal and parity matrices to the γ^5 matrix we obtain

$$C\gamma^5 C^{-1} = \gamma^5 \,, \quad T\gamma^5 T^{-1} = -\gamma^5 \,, \quad P\gamma^5 P^{-1} = -\gamma^5 \,. \tag{14.47}$$

We know that the Dirac equation for a free tachyon is given as (13.16)

$$\left(\mathrm{i}\gamma^B \partial_B - \gamma^5\, m\right) \psi(x) = 0 \,. \tag{14.48}$$

The initial steps for both charge conjugation and time dilation are identical, and begin by taking the adjoint of the equation, i.e.,

$$\psi^+(x) \left(-\mathrm{i} \left(\gamma^B\right)^+ \overleftarrow{\partial}_B - \gamma^5 m\right) = 0 \,, \tag{14.49}$$

we then multiply by γ^0 on the right and use the fact that $(\gamma^0)^2 = 1$ to find

$$\left(\psi^+(x)\gamma^0\right) \gamma^0 \left(-\mathrm{i} \left(\gamma^B\right)^+ \overleftarrow{\partial}_B - \gamma^5 m\right) \gamma^0 = \overline{\psi}(x) \left(-\mathrm{i}\gamma^B \overleftarrow{\partial}_B + \gamma^5 m\right) = 0 \,, \tag{14.50}$$

and finally take the transpose, yielding

$$\left(-\mathrm{i} \left(\gamma^B\right)^T \partial_B + \gamma^5 m\right) \overline{\psi}^{\,T}(x) \,. \tag{14.51}$$

We then apply the charge conjugation matrix to this equation, and find

$$C \left(-\mathrm{i} \left(\gamma^B\right)^T \partial_B + \gamma^5 m\right) C^{-1} \left(C \overline{\psi}^{\,T}(x)\right) = \left(\mathrm{i}\gamma^B \partial_B + \gamma^5 m\right) \psi^{\mathcal{C}}(x) \,. \tag{14.52}$$

Thus, under charge conjugation the mass term of the tachyonic Dirac equation reverses sign. We can now complete the time reversal operation by applying the time reversal matrix to (14.51), giving us

$$T\left(-\mathrm{i}\left(\gamma^B\right)^T\partial_B+\gamma^5 m\right)T^{-1}\left(T\overline{\psi}^{\,T}(x)\right)=\left(-\mathrm{i}\gamma^0\partial_0+\mathrm{i}\gamma^i\partial_i-\gamma^5 m\right)\left(T\overline{\psi}^{\,T}(x)\right)$$
$$=\left(\mathrm{i}\gamma^B\partial_B-\gamma^5 m\right)\psi^{\mathcal{T}}(x)=0\,, \qquad (14.53)$$

where again $\psi^{\mathcal{T}}(x)=T\overline{\psi}^{\,T}(x_{\mathcal{T}})$. As with the subluminal free particle, we find that the superluminal free Dirac equation is unaffected by the time reversal operation. Finally we investigate the effects of the parity operation, in which we begin by letting $x\to x_{\mathcal{P}}$, i.e.,

$$\left(\mathrm{i}\gamma^0\partial_0-\mathrm{i}\gamma^i\partial_i-\gamma^5 m\right)\psi(x_{\mathcal{P}})=0\,, \qquad (14.54)$$

we now apply the parity matrix and find

$$P\left(\mathrm{i}\gamma^0\partial_0-\mathrm{i}\gamma^i\partial_i-\gamma^5 m\right)P^{-1}\left(P\psi(x_{\mathcal{P}})\right)=\left(\mathrm{i}\gamma^B\partial_B+\gamma^5 m\right)\psi^{\mathcal{P}}(x)=0\,, \qquad (14.55)$$

where $\psi^{\mathcal{P}}(x)=P\psi(x_{\mathcal{P}})$. As with the charge conjugation transform, we find that the parity transform switches the sign of the mass term. Thus under a charge conjugation and parity transform, the sign of the mass term will switch twice, restoring the original equation, and the free tachyonic Dirac equation is then $\mathcal{C}_D\mathcal{P}_D$–symmetric. Additionally, we have seen that the equation is invariant under the time reversal operation, thus the equation is $\mathcal{T}_D$ symmetric. Thus, the superluminal Dirac equation is $\mathcal{CPT}$–symmetric. This is a somewhat important result, as the violation of $\mathcal{CPT}$–symmetry would otherwise imply that the equation is *not* Lorentz invariant [127].

15. GRAVITATIONALLY COUPLED TACHYON

Having worked through the derivation of the Dirac–Einstein–Schwarzschild Hamiltonian in chapter 8.5, we may now consider the gravitationally coupled tachyonic Dirac Hamiltonian with ease. As we have seen, the Dirac equation for a free tachyon in flat space-time is given by equation (13.16)

$$\left(\mathrm{i}\widetilde{\gamma}^A\partial_A - \widetilde{\gamma}^5\, m\right)\psi = 0\,. \tag{15.1}$$

As with a tardyon, the key observation is that the coupling to the gravitational field is given by the covariant derivative $\partial_A \to \nabla_\mu = \partial_\mu - \Gamma_\mu$ (8.9), where the spin connection matrix Γ_μ is defined in (8.50), and the replacement of the flat-space-time Dirac matrices by their curved-space-time counterparts. Thus the Dirac equation for a tachyon in curved space is

$$\left(\mathrm{i}\overline{\gamma}^\mu\nabla_\mu - \overline{\gamma}^5 m\right)\psi = 0\,. \tag{15.2}$$

Following the same procedure as in chapter 8.5, we quickly come to the equation

$$\mathrm{i}(\overline{\gamma}^0)^2\partial_0\psi = \left((\overline{\gamma}^0\overline{\gamma}^j p^j + \mathrm{i}\overline{\gamma}^0\overline{\gamma}^\mu\Gamma_\mu + \overline{\gamma}^0\overline{\gamma}^5 m\right)\psi\,, \tag{15.3}$$

where the $\overline{\gamma}^\mu$ matrices and Γ_μ are all the same as in the subluminal case, thus we can rewrite the equation in the form $\mathrm{i}\partial_t\psi = H\psi$, where

$$H = \frac{w}{v}\vec{\alpha}\cdot\vec{p} + \frac{\vec{\alpha}\cdot(\vec{p}\,w)}{2v} + \frac{w\vec{\alpha}\cdot(\vec{p}\,v)}{v^2} + \beta\overline{\gamma}^5 m\, w\,. \tag{15.4}$$

We then scale as the Hamiltonian as we did in (8.80), (8.81), and (8.82), obtaining

$$H' = \frac{1}{2}\{\vec{\alpha}\cdot\vec{p},\mathcal{F}\} + \beta\overline{\gamma}^5\, w\,, \quad \mathcal{F} = \frac{w}{v}\,. \tag{15.5}$$

The final step is to approximate for small r_s, just like we did in equation (8.83). However, for reasons that will become obvious later, we are here going to keep terms up to the second order in r_s. We recall that (see chapter 8.5)

$$w = \frac{1-\frac{r_\mathrm{s}}{4r}}{1+\frac{r_\mathrm{s}}{4r}}\,, \quad v = \left(1+\frac{r_\mathrm{s}}{4r}\right)^2\,, \quad \mathcal{F} = \frac{w}{v}\,. \tag{15.6}$$

By expanding about a small r_s, up to the second order, we find

$$w = 1 - \frac{r_\mathrm{s}}{2r} + \frac{r_\mathrm{s}^2}{8r^2} + \mathcal{O}(r_\mathrm{s}^3)\,, \quad v = 1 + \frac{r_\mathrm{s}}{2r} + \frac{r_\mathrm{s}^2}{16r^2} + \mathcal{O}(r_\mathrm{s}^3)\,, \tag{15.7a}$$

$$\mathcal{F} = 1 - \frac{r_\mathrm{s}}{r} + \frac{9r_\mathrm{s}^2}{16r^2} + \mathcal{O}(r_\mathrm{s}^3)\,, \tag{15.7b}$$

resulting in the equation

$$H_\mathrm{tg} = \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, 1 - \frac{r_\mathrm{s}}{r} + \frac{9\,r_\mathrm{s}^2}{16r^2}\right\} + \beta\,\overline{\gamma}^5\, m\left(1 - \frac{r_\mathrm{s}}{2r} + \frac{r_\mathrm{s}^2}{8r^2}\right)\,. \tag{15.8}$$

At this point we are familiar with all the terms in the Hamiltonian, except for the curved–space $\overline{\gamma}^5$ matrix. According to equation (18) of [91] the flat- and curved–space γ^5 matrices are generalized as

$$\gamma_5 = (-g)^{-\frac{1}{2}}(1/4!)\epsilon^{\alpha\beta\gamma\delta}\gamma_\alpha\gamma_\beta\gamma_\gamma\gamma_\delta\,, \quad (\gamma_5)^2 = -1\,, \tag{15.9}$$

where I have written the equation exactly written by Brill and Wheeler in [91]. Notice that using this definition we see that the square of γ_5 is -1, while we want to work in a system where $(\gamma_5)^2 = +1$. To accomplish this we simply multiply (15.9) by the imaginary unit i, bringing the definition into our notation as

$$\widetilde{\gamma}_5 = \frac{\mathrm{i}}{4!}\frac{\epsilon^{\alpha\beta\gamma\delta}}{\sqrt{-\det\widetilde{g}_{\mu\nu}}}\widetilde{\gamma}_\alpha\widetilde{\gamma}_\beta\widetilde{\gamma}_\delta\widetilde{\gamma}_\gamma\,, \quad \widetilde{\gamma}^5 = \frac{\mathrm{i}}{4!}\frac{\epsilon_{\alpha\beta\gamma\delta}}{\sqrt{-\det\widetilde{g}^{\mu\nu}}}\widetilde{\gamma}^\alpha\widetilde{\gamma}^\beta\widetilde{\gamma}^\delta\widetilde{\gamma}^\gamma\,, \tag{15.10a}$$

$$\overline{\gamma}_5 = \frac{\mathrm{i}}{4!}\frac{\epsilon^{\alpha\beta\gamma\delta}}{\sqrt{-\det\overline{g}_{\mu\nu}}}\overline{\gamma}_\alpha\overline{\gamma}_\beta\overline{\gamma}_\delta\overline{\gamma}_\gamma\,, \quad \overline{\gamma}^5 = \frac{\mathrm{i}}{4!}\frac{\epsilon_{\alpha\beta\gamma\delta}}{\sqrt{-\det\overline{g}^{\mu\nu}}}\overline{\gamma}^\alpha\overline{\gamma}^\beta\overline{\gamma}^\delta\overline{\gamma}^\gamma\,, \tag{15.10b}$$

where $(\widetilde{\gamma}_5)^2 = (\widetilde{\gamma}^5)^2 = 1$ and $\epsilon = \widetilde{\epsilon}$ is the flat-space Levi-Civita tensor in all cases, defined according to equation (2.19) in chapter 4 of [120] as

$$\epsilon_{\alpha\beta\gamma\delta} = \epsilon^{\alpha\beta\gamma\delta} = \begin{cases} +1 & \text{if } \alpha,\beta,\gamma,\delta \text{ is an even permutation of } 0,1,2,3 \\ -1 & \text{if } \alpha,\beta,\gamma,\delta \text{ is an odd permutation of } 0,1,2,3 \\ 0 & \text{otherwise} \end{cases}\,, \tag{15.11}$$

where we specialized to the four dimensional case. We know that in flat-space the metric is $\widetilde{g}_{\mu\nu} = \widetilde{g}^{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$, this $\sqrt{-\det\widetilde{g}_{\mu\nu}} = \sqrt{-\det\widetilde{g}^{\mu\nu}} = 1$, thus

$$\widetilde{\gamma}_5 = \frac{\mathrm{i}}{4!}\epsilon^{\alpha\beta\delta\gamma}\widetilde{\gamma}_\alpha\widetilde{\gamma}_\beta\widetilde{\gamma}_\delta\widetilde{\gamma}_\gamma\,, \quad \widetilde{\gamma}^5 = \frac{\mathrm{i}}{4!}\epsilon_{\alpha\beta\delta\gamma}\widetilde{\gamma}^\alpha\widetilde{\gamma}^\beta\widetilde{\gamma}^\delta\widetilde{\gamma}^\gamma\,. \tag{15.12}$$

Now, for a diagonal metric (such as both our flat-space and our curved-space metrics) $\{\gamma_\alpha, \gamma_\beta\} = \{\gamma^\alpha, \gamma^\beta\} = 0$ provided $\alpha \neq \beta$, which in the case of our equation for the gamma 5 matrices is guaranteed by the Levi-Civita tensor. Thus

$$\widetilde{\gamma}_5 = \frac{\mathrm{i}}{4!}\epsilon^{\alpha\beta\delta\gamma}\epsilon_{0123}\widetilde{\gamma}_\alpha\widetilde{\gamma}_\beta\widetilde{\gamma}_\delta\widetilde{\gamma}_\gamma = \frac{\mathrm{i}}{4!}\delta^{\alpha\beta\delta\gamma}_{0123}\widetilde{\gamma}_\alpha\widetilde{\gamma}_\beta\widetilde{\gamma}_\delta\widetilde{\gamma}_\gamma = \mathrm{i}\widetilde{\gamma}_{[0}\widetilde{\gamma}_1\widetilde{\gamma}_2\widetilde{\gamma}_{3]} = \mathrm{i}\widetilde{\gamma}_0\widetilde{\gamma}_1\widetilde{\gamma}_2\widetilde{\gamma}_3\,, \tag{15.13}$$

where we used $\epsilon_{0123} = \pm 1$ (0123 is an even permutation) in the first step, the identity $\delta^{\alpha\beta\delta\gamma}_{\mu\nu\rho\sigma} = \epsilon^{\alpha\beta\delta\gamma}\epsilon_{\mu\nu\rho\sigma}$ (see equation (2.20) in chapter 4 of [120]) in the second step, and the fact that the Dirac γ matrices anticommute in the final step. Similarly

$$\widetilde{\gamma}^5 = \mathrm{i}\widetilde{\gamma}^0\widetilde{\gamma}^1\widetilde{\gamma}^2\widetilde{\gamma}^3\,. \tag{15.14}$$

Due to the fact that our curved space metric is diagonal as well, this same process can be easily applied to the curved space Dirac $\widetilde{\gamma}^5$ matrices, yielding

$$\overline{\gamma}_5 = \frac{\mathrm{i}}{\sqrt{-\det \overline{g}_{\mu\nu}}} \overline{\gamma}_0 \overline{\gamma}_1 \overline{\gamma}_2 \overline{\gamma}_3 \,, \quad \overline{\gamma}^5 = \frac{\mathrm{i}}{\sqrt{-\det \overline{g}^{\mu\nu}}} \overline{\gamma}^0 \overline{\gamma}^1 \overline{\gamma}^2 \overline{\gamma}^3 \,. \tag{15.15}$$

This is valid for any Dirac γ^5 matrix provided the relevant metric is diagonal. Let us explicitly write out the metric, its inverse and the relation between the flat- and curved–spacetime Dirac γ matrices,

$$[\overline{g}_{\mu\nu}] = \begin{pmatrix} w^2 & 0 & 0 & 0 \\ 0 & -v^2 & 0 & 0 \\ 0 & 0 & -v^2 & 0 \\ 0 & 0 & 0 & -v^2 \end{pmatrix} \,, \quad [\overline{g}^{\mu\nu}] = \begin{pmatrix} \frac{1}{w^2} & 0 & 0 & 0 \\ 0 & -\frac{1}{v^2} & 0 & 0 \\ 0 & 0 & -\frac{1}{v^2} & 0 \\ 0 & 0 & 0 & -\frac{1}{v^2} \end{pmatrix} \,, \tag{15.16}$$

$$\overline{\gamma}_0 = w \, \widetilde{\gamma}_0 \,, \quad \overline{\gamma}_i = v \, \widetilde{\gamma}_i \,, \quad \overline{\gamma}^0 = \frac{1}{w} \widetilde{\gamma}^0 \,, \quad \overline{\gamma}^i = \frac{1}{v} \widetilde{\gamma}^i \,, \tag{15.17}$$

thus

$$\det \overline{g}_{\mu\nu} = -\, w^2 v^6 \,, \quad \det \overline{g}^{\mu\nu} = -\frac{1}{w^2 v^6} \,, \tag{15.18a}$$

$$\sqrt{-\det \overline{g}_{\mu\nu}} = w \, v^3 \,, \quad \sqrt{-\det \overline{g}^{\mu\nu}} = \frac{1}{w \, v^3} \,, \tag{15.18b}$$

$$\overline{\gamma}_0 \overline{\gamma}_1 \overline{\gamma}_2 \overline{\gamma}_3 = w v^3 \widetilde{\gamma}_0 \widetilde{\gamma}_1 \widetilde{\gamma}_2 \widetilde{\gamma}_3 \,, \quad \overline{\gamma}^0 \overline{\gamma}^1 \overline{\gamma}^2 \overline{\gamma}^3 = \frac{1}{w v^3} \widetilde{\gamma}^0 \widetilde{\gamma}^1 \widetilde{\gamma}^2 \widetilde{\gamma}^3 \,. \tag{15.18c}$$

We then plug these results into (15.15) giving us

$$\overline{\gamma}_5 = \mathrm{i} \widetilde{\gamma}_0 \widetilde{\gamma}_1 \widetilde{\gamma}_2 \widetilde{\gamma}_3 = \widetilde{\gamma}_5 \,, \quad \overline{\gamma}^5 = \mathrm{i} \widetilde{\gamma}^0 \widetilde{\gamma}^1 \widetilde{\gamma}^2 \widetilde{\gamma}^3 = \widetilde{\gamma}^5 \,. \tag{15.19}$$

Thus the tachyonic gravitationally coupled Dirac Hamiltonian becomes

$$H_{\mathrm{tg}} = \frac{1}{2} \left\{ \vec{\alpha} \cdot \vec{p}, 1 - \frac{r_{\mathrm{s}}}{r} + \frac{9 \, r_{\mathrm{s}}^2}{16 r^2} \right\} + \beta \, \widetilde{\gamma}^5 \, m \left(1 - \frac{r_{\mathrm{s}}}{2r} + \frac{r_{\mathrm{s}}^2}{8 r^2} \right) \,. \tag{15.20}$$

Similarly, we will also need the Dirac–Einstein–Schwarzschild Hamiltonian to the second order in $r_{\rm s}$. Using our approximations from (15.7), as well as the exact result for the Dirac–Einstein–Schwarzschild Hamiltonian (8.82), we quickly find that to the second order in $r_{\rm s}$

$$H_{\rm ds} = \frac{1}{2}\left\{\vec{\alpha}\cdot\vec{p}, 1 - \frac{r_{\rm s}}{r} + \frac{9\,r_{\rm s}^2}{16r^2}\right\} + \beta\,m\left(1 - \frac{r_{\rm s}}{2r} + \frac{r_{\rm s}^2}{8r^2}\right)\,. \tag{15.21}$$

Here we anticipate the need for this slight generalization of (8.84) (see chapter 17.1).

16. EXACT ULTRARELATIVISTIC DECOUPLING TRANSFORM

16.1. INITIAL ROTATION INTO THE WEYL BASIS

We start from the Dirac representation of the γ matrices, both for the tardyonic as well as tachyonic Dirac Hamiltonians, and first rotate into a different matrix representation (the helicity basis, see chapter 23 of [121]), before carrying out the decoupling transformation. For simplicity, we shall consider the free tachyonic Hamiltonian here, first. All the considerations in this chapter are trivially generalized to the gravitationally coupled Hamiltonians. Recall that the tachyonic free particle Hamiltonian is

$$H_{\rm TF} = \vec{\alpha} \cdot \vec{p} + \beta \, \gamma^5 \, m \, . \tag{16.1}$$

Also note that the Hamiltonian is not Hermitian, but is instead γ^5-Hermitian. We consider three possible initial rotations, which shift the Hamiltonian into the Weyl representation, and prepare it for the ultrarelativistic transform. These rotations are all unitary, and are

$$U_{\rm A} = \frac{1}{\sqrt{2}} \left(\beta + \gamma^5 \right) \, , \tag{16.2}$$

$$U_{\rm B} = \frac{1}{\sqrt{2}} \left(\beta - \gamma^5 \right) \, , \tag{16.3}$$

$$U_{\rm C} = \frac{1}{\sqrt{2}} \left(1 - \beta \gamma^5 \right) \, . \tag{16.4}$$

Let us investigate how all three of them transform the Hamiltonian:

$$H_A = U_{\rm A} \, H_{\rm TF} \, U_{\rm A}^+ = \beta \vec{\Sigma} \cdot \vec{p} - \beta \, \gamma^5 \, m \, , \tag{16.5}$$

$$H_B = U_{\rm B} \, H_{\rm TF} \, U_{\rm B}^+ = -\beta \, \vec{\Sigma} \cdot \vec{p} - \beta \, \gamma^5 \, m \, . \tag{16.6}$$

Now, to calculate H_C we note that $U_{\rm C} = \beta\, U_{\rm B}$, thus

$$H_C = U_{\rm C}\, H_{\rm TF}\, U_{\rm C}^{+} = \beta U_{\rm B}\, H_{\rm TF}\, U_{\rm B}^{+} \beta = \beta\, H_B\, \beta = -\beta\, \vec{\Sigma} \cdot \vec{p} + \beta\, \gamma^5\, m\,. \tag{16.7}$$

This final rotation ($U_{\rm C}$) gives us the form we want, as it has the plus sign in front of the $\beta\,\gamma^5$ term.

It is known that the Weyl equation describes a massless spin–1/2 particle, and splits into two equations which describe left–handed and right–handed spinors (see chapter 23 of [121] and page 87 of [60]),

$$\mathrm{i}\, \partial_t\, \psi_L = H_L\, \psi_L\,, \quad H_L = -\vec{\sigma} \cdot \vec{p}\,, \tag{16.8}$$

$$\mathrm{i}\, \partial_t\, \psi_R = H_R\, \psi_R\,, \quad H_R = \vec{\sigma} \cdot \vec{p}\,. \tag{16.9}$$

Under parity a left–handed spinor transforms into a right–handed spinor, and as such the Weyl equations break parity. However, by "stacking" the helicity spinors, one can construct the spinor solutions to the Dirac equation [122]. The massless Weyl spinors (equations (16.8) and (16.9)) correspond to the ultrarelativistic limit for a massive particle, plus correction terms. We can then expect that our initial rotation will necessarily break parity, i.e., the γ^5–Hermiticity will be broken.

We define the initial rotation U_1 as

$$U_1 = U_{\rm C} = \frac{1}{\sqrt{2}} \left(1 - \beta\gamma^5\right)\,, \tag{16.10}$$

and the resulting transformed Hamiltonian for the free tachyon is

$$H'_{\rm TF} = H_C = \beta\mathcal{E} + \beta\, \gamma^5\, m\,, \tag{16.11}$$

where

$$\mathcal{E} = -\vec{\Sigma} \cdot \vec{p}\,, \tag{16.12}$$

is the energy operator for a left–handed neutrino.

We also note that the transformed Hamiltonian is no longer γ^5–Hermitian, instead it is β– or γ^0–Hermitian. Given a pseudo–Hermitian Hamiltonian H, which is $\mathcal{A}$–Hermitian, then the transformed Hamiltonian $H' = U\,H\,U^+$ (where U is unitary) will be $\mathcal{B}$–Hermitian, where $\mathcal{B} = U\,\mathcal{A}\,U^+$, i.e.,

$$\begin{aligned}(H')^+ = \left(U\,H\,U^+\right)^+ &= U\,H^+\,U^+ = U\,\mathcal{A}\,H\,\mathcal{A}^{-1}\,U^+ \\ &= U\,\mathcal{A}\,U^+\,U\,H\,U^+\,U\,\mathcal{A}^{-1}\,U^+ = \mathcal{B}\,H'\,\mathcal{B}^{-1}\,.\end{aligned} \tag{16.13}$$

Now, U_1 is unitary, and $U_1\,\gamma^5\,U_1 = -\beta$, thus $H'_{\rm TF}$ is β–Hermitian. We also note that $\beta = \gamma^5_{\rm W}$, where $\gamma^5_{\rm W}$ is the Dirac γ^5 matrix in the Weyl representation, thus in the Weyl representation $H'_{\rm TF}$ is $\gamma^5_{\rm W}$–Hermitian.

16.2. TARDYONIC FREE PARTICLE

Inspired by the exact Foldy–Wouthuysen transform for a free particle, we apply the same methodology to perform an exact ultrarelativistic transform. The free Dirac–Hamiltonian (for a tardyon) is given as (C.34)

$$H_{\rm F} = \vec{\alpha} \cdot \vec{p} + \beta\,m\,. \tag{16.14}$$

We must first transform the Hamiltonian into the Weyl basis, as we want the energy ($\mathcal{E}$) term to be along the diagonal. Applying the initial rotation U_1 given in equation (16.10) we find

$$\begin{aligned}
H'_{\mathrm{F}} &= U_1\, H_{\mathrm{F}}\, U_1^{+} = \frac{1}{2}\left(1-\beta\gamma^5\right)\left(\vec{\alpha}\cdot\vec{p}+\beta\, m\right)\left(1-\gamma^5\beta\right) \\
&= \frac{1}{2}\left(1-\beta\gamma^5\right)\left(\vec{\alpha}\cdot\vec{p}+\beta\, m-\beta\,\gamma^5\vec{\alpha}\cdot\vec{p}+\gamma^5\, m\right) \\
&= \frac{1}{2}\left(\vec{\alpha}\cdot\vec{p}+\beta\, m-\beta\,\gamma^5\vec{\alpha}\cdot\vec{p}+\gamma^5\, m-\beta\,\gamma^5\vec{\alpha}\cdot\vec{p}+\gamma^5\, m-\vec{\alpha}\cdot\vec{p}-\beta\, m\right) \\
&= \beta\mathcal{E}+\gamma^5\, m\,,
\end{aligned} \tag{16.15}$$

where $\mathcal{E}$ is defined in (16.12). We then apply a second rotation U_{F} where

$$U_{\mathrm{F}} = \mathrm{e}^{\mathrm{i}\, S_{\mathrm{F}}}\,, \quad S_{\mathrm{F}} = -\mathrm{i}\beta\gamma^5\frac{m}{\mathcal{E}}\theta\,. \tag{16.16}$$

Notice S_{F} is Hermitian, ensuring that the transform U_{F} is unitary. We can rewrite the transform as

$$\begin{aligned}
U_{\mathrm{F}} &= \exp\left(\beta\gamma^5\frac{m}{\mathcal{E}}\theta\right) = \sum_{n=0}^{\infty}\frac{1}{n!}\left(\beta\gamma^5\frac{m}{\mathcal{E}}\theta\right) \\
&= \sum_{k=0}^{\infty}\frac{1}{(2k)!}\left(\beta\gamma^5\frac{m}{\mathcal{E}}\theta\right)^{2k}+\sum_{k=0}^{\infty}\frac{1}{(2k+1)!}\left(\beta\gamma^5\frac{m}{\mathcal{E}}\theta\right)^{2k+1} \\
&= \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}\left(\frac{m}{|\vec{p}|}\theta\right)^{2k}+\beta\gamma^5\frac{|\vec{p}|}{\mathcal{E}}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}\left(\frac{m}{|\vec{p}|}\theta\right)^{2k+1} \\
&= \cos\left(\frac{m}{|\vec{p}|}\theta\right)+\beta\gamma^5\frac{|\vec{p}|}{\mathcal{E}}\sin\left(\frac{m}{|\vec{p}|}\theta\right)\,.
\end{aligned} \tag{16.17}$$

Then applying it to H'_{F} we find

$$
\begin{aligned}
U_{\mathrm{F}}\, H'_{\mathrm{F}}\, U_{\mathrm{F}}^{+} &= \left[\cos\Theta + \beta\gamma^5 \frac{|\vec{p}|}{m}\sin\Theta\right]\left[\beta\mathcal{E} + \gamma^5\, m\right]\left[\cos\Theta - \beta\gamma^5 \frac{|\vec{p}|}{m}\sin\Theta\right] \\
&= \left[\beta\mathcal{E} + \gamma^5\, m\right]\left[\cos\Theta - \beta\gamma^5 \frac{|\vec{p}|}{m}\sin\Theta\right]^2 \\
&= \left[\beta\mathcal{E} + \gamma^5\, m\right]\left[\exp\left(-\beta\gamma^5 \frac{m}{\mathcal{E}}\theta\right)\right]^2 = \left[\beta\mathcal{E} + \gamma^5\, m\right]\exp\left(-2\beta\gamma^5 \frac{m}{\mathcal{E}}\theta\right) \\
&= \left[\beta\mathcal{E} + \gamma^5\, m\right]\left[\cos 2\Theta - \beta\gamma^5 \frac{|\vec{p}|}{\mathcal{E}}\sin 2\Theta\right] \\
&= \beta\mathcal{E}\, \cos 2\Theta - \gamma^5\mathcal{E}\frac{|\vec{p}|}{\mathcal{E}}\sin 2\Theta + \gamma^5 m\cos 2\Theta + \beta m\frac{|\vec{p}|}{\mathcal{E}}\sin 2\Theta \\
&= \gamma^5\, m\, \cos 2\Theta\left[1 - \frac{|\vec{p}|}{m}\tan 2\Theta\right] + \beta\left[\mathcal{E}\, \cos 2\Theta + m\frac{|\vec{p}|}{\mathcal{E}}\, \sin 2\Theta\right], \qquad (16.18)
\end{aligned}
$$

where $\Theta = (m/\,|\vec{p}|)\theta$. In order to eliminate the odd terms (the terms with the γ^5 prefactor), we choose θ such that

$$
\tan 2\Theta = \tan\left(2\frac{m}{|\vec{p}|}\theta\right) = \frac{m}{|\vec{p}|}\,. \qquad (16.19)
$$

Therefore

$$
\sin 2\Theta = \sin\left(2\frac{m}{|\vec{p}|}\theta\right) = \frac{m}{\sqrt{\vec{p}^2 + m^2}}\,, \quad \cos 2\Theta = \cos\left(2\frac{m}{|\vec{p}|}\theta\right) = \frac{|\vec{p}|}{\sqrt{\vec{p}^2 + m^2}}\,, \qquad (16.20)
$$

and we find

$$
\begin{aligned}
\mathcal{H}_{\mathrm{F}} = U_{\mathrm{F}}\, H\, U_{\mathrm{F}}^{+} &= \beta\left[\mathcal{E}\frac{|\vec{p}|}{\sqrt{\vec{p}^2 + m^2}} + m\frac{|\vec{p}|}{\mathcal{E}}\frac{m}{\sqrt{\vec{p}^2 + m^2}}\right] \\
&= \beta\frac{\mathcal{E}}{|\vec{p}|}\left[\frac{\vec{p}^2}{\sqrt{\vec{p}^2 + m^2}} + \cancelto{1}{\frac{\vec{p}^2}{\mathcal{E}^2}}\, \frac{m^2}{\sqrt{\vec{p}^2 + m^2}}\right] \\
&= \beta\frac{\mathcal{E}}{|\vec{p}|}\frac{\vec{p}^2 + m^2}{\sqrt{\vec{p}^2 + m^2}} = \beta\frac{\mathcal{E}}{|\vec{p}|}\sqrt{\vec{p}^2 + m^2}\,. \qquad (16.21)
\end{aligned}
$$

In explicit matrix form this is

$$\mathcal{H}_{\rm F} = \begin{pmatrix} -\frac{\vec{\sigma}\cdot\vec{p}}{|\vec{p}|}\sqrt{\vec{p}^{\,2}+m^2} & 0 \\ 0 & \frac{\vec{\sigma}\cdot\vec{p}}{|\vec{p}|}\sqrt{\vec{p}^{\,2}+m^2} \end{pmatrix} . \tag{16.22}$$

In this notation it is clear that the Hamiltonian has been successfully separated into a set of left– and right–handed Hamiltonians.

With this result in hand, we would like to note that we are not the first to consider this type of transform [123–125]. Like our ultrarelativistic decoupling transformation, the so called Cini–Touschek transformation [123] was designed to decouple the particle and antiparticle degrees of freedom. However, rather than rotating into a more appropriate basis, this method eliminates the even elements, leaving only the off diagonal odd elements. This is in sharp contrast to our method, which brings the rotated Hamiltonian to diagonal form.

16.3. TACHYONIC FREE PARTICLE

From (16.11) we already know that the free particle Dirac–Hamiltonian (for a tachyon), rotated to the Weyl representation is given as

$$H'_{\rm TF} = \beta\mathcal{E} + \beta\,\gamma^5\,m\,. \tag{16.23}$$

where $\mathcal{E}$ is the energy operator for a left–handed neutrino as given in (16.12). The exact transform $U_{\rm TF}$ is then

$$U_{\rm TF} = {\rm e}^{{\rm i}\,S_{\rm TF}}\,, \quad S_{\rm TF} = -{\rm i}\gamma^5\frac{m}{\mathcal{E}}\theta\,. \tag{16.24}$$

Notice that $S_{\rm TF}$ is *not* Hermitian, but is instead β–Hermitian, as discussed at the end of chapter 16.1, $H'_{\rm TF}$ is β–Hermitian, and we want to preserve that property of the Hamiltonian. Since $S_{\rm TF}$ is β–Hermitian i.e.,

$$S_{\rm TF}^{+} = {\rm i}\gamma^5 \frac{m}{\mathcal{E}}\theta = \beta\beta{\rm i}\gamma^5 \frac{m}{\mathcal{E}}\theta = \beta\left(-{\rm i}\gamma^5 \frac{m}{\mathcal{E}}\theta\right)\beta = \beta\, S_{\rm TF}\,\beta\,, \tag{16.25}$$

we ensure that the operator $U_{\rm TF}$ is β–unitary, i.e.,

$$U_{\rm TF}^{-1} = \beta\, U_{\rm TF}^{+}\,\beta\,, \quad \text{or} \quad U_{\rm TF}^{+}\beta\, U_{\rm TF} = \beta\,. \tag{16.26}$$

which will ensure that $H''_{\rm TF} = U_{\rm TF} H'_{\rm TF}\, U_{\rm TF}^{-1}$ is unitary. We can now rewrite $U_{\rm TF}$ as

$$\begin{aligned}
U_{\rm TF} = \exp\left(\gamma^5 \frac{m}{\mathcal{E}}\theta\right) &= \sum_{n=0}^{\infty}\frac{1}{n!}\left(\gamma^5 \frac{m}{\mathcal{E}}\theta\right)^n \\
&= \sum_{k=0}^{\infty}\frac{1}{(2k)!}\left(\gamma^5 \frac{m}{\mathcal{E}}\theta\right)^{2k} + \sum_{k=0}^{\infty}\frac{1}{(2k+1)!}\left(\gamma^5 \frac{m}{\mathcal{E}}\theta\right)^{2k+1} \\
&= \sum_{k=0}^{\infty}\frac{1}{(2k)!}\left(\frac{m}{|\vec{p}|}\theta\right)^{2k} + \gamma^5\frac{|\vec{p}|}{\mathcal{E}}\sum_{k=0}^{\infty}\frac{1}{(2k+1)!}\left(\frac{m}{|\vec{p}|}\theta\right)^{2k+1} \\
&= \cosh\left(\frac{m}{|\vec{p}|}\theta\right) + \gamma^5\frac{|\vec{p}|}{\mathcal{E}}\sinh\left(\frac{m}{|\vec{p}|}\theta\right)\,.
\end{aligned} \tag{16.27}$$

Thus

$$\begin{aligned}
\mathcal{H}''_{\rm TF} =& U_{\rm TF}\, H'_{\rm TF}\, U_{\rm TF}^{-1} \\
=& \left[\cosh\Theta + \gamma^5\frac{|\vec{p}|}{\mathcal{E}}\sinh\Theta\right]\left[\beta\mathcal{E}r + \beta\,\gamma^5\, m\right]\left[\cosh\Theta - \gamma^5\frac{|\vec{p}|}{\mathcal{E}}\sinh\Theta\right] \\
=& \left[\beta\mathcal{E} + \beta\,\gamma^5\, m\right]\left[\cosh\Theta - \gamma^5\frac{|\vec{p}|}{\mathcal{E}}\sinh\Theta\right]^2 = \left[\beta\mathcal{E} + \beta\,\gamma^5\, m\right]\exp\left(2\gamma^5\frac{m}{\mathcal{E}}\theta\right) \\
=& \beta\left[\mathcal{E} + \gamma^5\, m\right]\left[\cosh 2\Theta - \gamma^5\frac{|\vec{p}|}{\mathcal{E}}\sinh 2\Theta\right] \\
=& \beta\left[\mathcal{E}\cosh 2\Theta - \gamma^5\,|\vec{p}|\sinh 2\Theta + \gamma^5\, m\cosh 2\Theta - \frac{|\vec{p}|}{\mathcal{E}}m\sinh 2\Theta\right] \\
=& \beta\left[\gamma^5\, m\cosh 2\Theta\left[1 - \frac{|\vec{p}|}{m}\tanh 2\Theta\right] + \mathcal{E}\cosh 2\Theta - \frac{|\vec{p}|}{\mathcal{E}}m\sinh 2\Theta\right]\,,
\end{aligned} \tag{16.28}$$

where $\Theta = (m/|\vec{p}|)\theta$. Then to eliminate the odd terms (γ^5 prefactor) we choose θ such that

$$\tanh\Theta = \tanh\left(2\frac{m}{|\vec{p}|}\theta\right) = \frac{m}{|\vec{p}|}\,, \tag{16.29}$$

in which case

$$\sinh 2\Theta = \sinh\left(2\frac{m}{|\vec{p}|}\theta\right) = \frac{m}{\sqrt{\vec{p}^{\,2} - m^2}}\,, \tag{16.30a}$$

$$\cosh 2\Theta = \cosh\left(2\frac{m}{|\vec{p}|}\theta\right) = \frac{|\vec{p}|}{\sqrt{\vec{p}^{\,2} - m^2}}\,. \tag{16.30b}$$

Then

$$\begin{aligned}
\mathcal{H}_{\mathrm{TF}} &= \beta\left[\mathcal{E}\frac{|\vec{p}|}{\sqrt{\vec{p}^{\,2} - m^2}} - \frac{|\vec{p}|}{\mathcal{E}}m\frac{m}{\sqrt{\vec{p}^{\,2} - m^2}}\right] \\
&= \beta\frac{\mathcal{E}}{|\vec{p}|}\left[\frac{\vec{p}^{\,2}}{\sqrt{\vec{p}^{\,2} - m^2}} - \cancelto{1}{\frac{\vec{p}^{\,2}}{\mathcal{E}^2}}\frac{m^2}{\sqrt{\vec{p}^{\,2} - m^2}}\right] \\
&= \beta\frac{\mathcal{E}}{|\vec{p}|}\frac{\vec{p}^{\,2} - m^2}{\sqrt{\vec{p}^{\,2} - m^2}} = \beta\frac{\mathcal{E}}{|\vec{p}|}\sqrt{\vec{p}^{\,2} - m^2}\,.
\end{aligned} \tag{16.31}$$

Thus we have performed the exact ultrarelativistic decoupling transform on both the sub- and superluminal free particles. In performing the transformation we managed to eliminate the odd parts, and fully decouple the particles and antiparticles.

17. GENERAL ULTRARELATIVISTIC DECOUPLING TRANSFORM

17.1. DIRAC–EINSTEIN–SCHWARZSCHILD HAMILTONIAN

We now want to apply the ultrarelativistic decoupling transform to the Dirac–Einstein–Schwarzschild Hamiltonian. As with the Foldy–Wouthuysen transformation, this Hamiltonian is too complex to perform an exact transform (i.e., to all orders in momenta). Instead we must use a perturbative method. In performing this calculation we will keep all terms up to the second order in $r_{\rm s}$, and to the first order in $1/\mathcal{E}$. Prior to the perturbative transform, we must perform the exact initial rotation into the Weyl basis. Transforming (15.21) using U_1 (16.10) yields

$$H'_{\rm ds} = \beta\mathcal{E} - \frac{\beta}{2}\left\{\mathcal{E}, \frac{r_{\rm s}}{r}\right\} + \beta\frac{9}{32}\left\{\mathcal{E}, \frac{r_{\rm s}^2}{r^2}\right\} + \mathcal{O}_{\rm ds}\,, \tag{17.1}$$

where we have already introduced the odd term

$$\mathcal{O}_{\rm ds} = \gamma^5\, m\left(1 - \frac{r_{\rm s}}{2r} + \frac{r_{\rm s}^2}{8r^2}\right)\,. \tag{17.2}$$

We can now construct the operator

$$S_{\rm ds} = -\,{\rm i}\frac{\beta}{4}\left\{\mathcal{O}_{\rm ds}, \frac{1}{\mathcal{E}}\right\} = -{\rm i}\frac{m}{2}\left(\beta\gamma^5\frac{1}{\mathcal{E}} - \beta\gamma^5\frac{1}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\} + \beta\gamma^5\frac{1}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\}\right)\,. \tag{17.3}$$

The perturbative calculation is then performed in much the same way as the Foldy–Wouthuysen program, utilizing the approximation

$$H''_{\rm ds} = H'_{\rm ds} + \frac{{\rm i}^1}{1!}[S_{\rm ds}, H_{\rm ds}] + \frac{{\rm i}^2}{2!}[S_{\rm ds}, [S_{\rm rs}, H'_{\rm ds}]] + \dots\,, \tag{17.4}$$

which is again a series of nested commutators. As before, the canceled terms are of a sufficiently high order, so that they can be neglected. The single commutator is then

$$\begin{aligned}[S_{\rm ds}, H'_{\rm ds}] = [S_{\rm ds}, \beta\mathcal{E}] - \frac{1}{2}\left[S_{\rm ds}, \beta\left\{\mathcal{E}, \frac{r_{\rm s}}{r}\right\}\right] + \frac{9}{32}\left[S_{\rm ds}, \left\{\mathcal{E}, \frac{r_{\rm s}^2}{r^2}\right\}\right] \\ + \left[S_{\rm ds}, \gamma^5\, m\right] - \frac{1}{2}\left[S_{\rm ds}, \gamma^5\, m\frac{r_{\rm s}}{r}\right] + \frac{1}{8}\left[S_{\rm ds}, \gamma^5\, m\frac{r_{\rm s}^2}{r^2}\right]\,, \end{aligned} \tag{17.5}$$

giving us six terms to calculate. Let us begin by calculating the first term

$$\begin{aligned}[S_{\rm ds}, \beta\mathcal{E}] &= -\frac{\rm i}{4}\left[\beta\left\{\mathcal{O}_{\rm ds}, \frac{1}{\mathcal{E}}\right\}, \beta\mathcal{E}\right] = \frac{\rm i}{4}\left\{\mathcal{E}, \left\{\frac{1}{\mathcal{E}}, \mathcal{O}_{\rm ds}\right\}\right\} = \frac{\rm i}{4}\left\{\mathcal{E}, \frac{1}{\mathcal{E}}\mathcal{O}_{\rm ds} + \mathcal{O}_{\rm ds}\frac{1}{\mathcal{E}}\right\} \\ &= \frac{\rm i}{4}\left(\mathcal{E}\frac{1}{\mathcal{E}}\mathcal{O}_{\rm ds} + \frac{1}{\mathcal{E}}\mathcal{O}_{\rm ds}\mathcal{E} + \mathcal{E}\mathcal{O}_{\rm ds}\frac{1}{\mathcal{E}} + \mathcal{O}_{\rm ds}\frac{1}{\mathcal{E}}\mathcal{E}\right) \\ &= \frac{\rm i}{4}\left(2\mathcal{O}_{\rm ds} + \frac{1}{\mathcal{E}}\mathcal{O}_{\rm ds}\mathcal{E} + \mathcal{E}\mathcal{O}_{\rm ds}\frac{1}{\mathcal{E}}\right) \\ &= \frac{\rm i}{4}\left(2\mathcal{O}_{\rm ds} + \left(\mathcal{O}_{\rm ds} - \frac{1}{\mathcal{E}}\left[\mathcal{E}, \mathcal{O}_{\rm ds}\right]\right) + \left(\mathcal{O}_{\rm ds} + \left[\mathcal{E}, \mathcal{O}_{\rm ds}\right]\frac{1}{\mathcal{E}}\right)\right) \\ &= \frac{\rm i}{4}\left(4\mathcal{O}_{\rm ds} + \frac{1}{\mathcal{E}}\mathcal{E}\left[\mathcal{E}, \mathcal{O}_{\rm ds}\right]\frac{1}{\mathcal{E}} - \frac{1}{\mathcal{E}}\left[\mathcal{E}, \mathcal{O}_{\rm ds}\right]\right) \\ &= \frac{\rm i}{4}\left(4\mathcal{O}_{\rm ds} + \frac{1}{\mathcal{E}}\left[\mathcal{E}, \left[\mathcal{E}, \mathcal{O}_{\rm ds}\right]\right]\frac{1}{\mathcal{E}} + \frac{1}{\mathcal{E}}\left[\mathcal{E}, \mathcal{O}_{\rm ds}\right] - \frac{1}{\mathcal{E}}\left[\mathcal{E}, \mathcal{O}_{\rm ds}\right]\right) \\ &= {\rm i}\mathcal{O}_{\rm ds} + \cancel{\frac{\rm i}{4}\frac{1}{\mathcal{E}}\left[\mathcal{E}, \left[\mathcal{E}, \mathcal{O}_{\rm ds}\right]\right]\frac{1}{\mathcal{E}}}\,. \end{aligned} \tag{17.6}$$

Notice that we approximated the second term in the final expression to zero, despite the fact that the double commutator has two instances of the operator $\mathcal{E}$. It may appear as if the second term is proportional to 1 (in terms of inverse powers of $\mathcal{E}$). This is in fact not the case, as the commutators ensure that the contained instances of $\mathcal{E}$ will not operate on the wave–function, which would otherwise generate energy terms. We are working in the high energy limit, and we only get these dominant terms when the operator $\mathcal{E}$ operates on the wave–function, not when it operates on any other term. Thus

$$\frac{1}{\mathcal{E}}\left[\mathcal{E},\left[\mathcal{E},\mathcal{O}_{\mathrm{ds}}\right]\right]\frac{1}{\mathcal{E}}=\mathcal{O}\left(\frac{1}{\mathcal{E}^2}\right)\to 0\,. \tag{17.7}$$

We now turn our attention to the second term,

$$\begin{aligned}\left[S_{\mathrm{ds}},\beta\left\{\mathcal{E},\frac{r_{\mathrm{s}}}{r}\right\}\right] &= -\mathrm{i}\frac{m}{2}\left[\beta\gamma^5\frac{1}{\mathcal{E}}-\beta\gamma^5\frac{1}{4}\left\{\frac{1}{\mathcal{E}},\frac{r_{\mathrm{s}}}{r}\right\}+\cancel{\frac{\gamma^5}{16}\left\{\frac{1}{\mathcal{E}},\frac{r_{\mathrm{s}}^2}{r^2}\right\}},\beta\left\{\mathcal{E},\frac{r_{\mathrm{s}}}{r}\right\}\right]\\ &=\frac{\mathrm{i}}{2}\gamma^5\,m\left(\left\{\frac{1}{\mathcal{E}},\left\{\mathcal{E},\frac{r_{\mathrm{s}}}{r}\right\}\right\}-\frac{1}{4}\left\{\left\{\frac{1}{\mathcal{E}},\frac{r_{\mathrm{s}}}{r}\right\},\left\{\mathcal{E},\frac{r_{\mathrm{s}}}{r}\right\}\right\}\right)\,.\end{aligned} \tag{17.8}$$

Here it is beneficial to show that for a function $f=f(r)$,

$$\begin{aligned}\frac{1}{\mathcal{E}}\,f\,\mathcal{E}+\mathcal{E}\,f\,\frac{1}{\mathcal{E}} &= \left(f-\frac{1}{\mathcal{E}}\left[\mathcal{E},f\right]\right)+\left(f+\left[\mathcal{E},f\right]\frac{1}{\mathcal{E}}\right)=2f+\frac{1}{\mathcal{E}}\mathcal{E}\left[\mathcal{E},f\right]-\frac{1}{\mathcal{E}}\left[\mathcal{E},f\right]\\ &=2f+\cancel{\frac{1}{\mathcal{E}}\left[\mathcal{E},\left[\mathcal{E},f\right]\right]\frac{1}{\mathcal{E}}}+\frac{1}{\mathcal{E}}\left[\mathcal{E},f\right]-\frac{1}{\mathcal{E}}\left[\mathcal{E},f\right]=2f\,,\end{aligned} \tag{17.9}$$

and we can use this identity to show another,

$$\begin{aligned}\left\{\frac{1}{\mathcal{E}},\left\{\mathcal{E},f\right\}\right\} &= \left\{\frac{1}{\mathcal{E}},\mathcal{E}\,f+f\,\mathcal{E}\right\}=\frac{1}{\mathcal{E}}\mathcal{E}\,f+\mathcal{E}\,f\,\frac{1}{\mathcal{E}}+\frac{1}{\mathcal{E}}\,f\,\mathcal{E}+f\,\mathcal{E}\frac{1}{\mathcal{E}}\\ &=2f+\left(\mathcal{E}\,f\,\frac{1}{\mathcal{E}}+\frac{1}{\mathcal{E}}\,f\,\mathcal{E}\right)=4f\,,\end{aligned} \tag{17.10}$$

in our approximation. The latter identity will be used to calculate the first anticommutator of the second term for us, leaving only the second anticommutator,

$$\begin{aligned}\left\{\left\{\frac{1}{\mathcal{E}},\frac{r_{\mathrm{s}}}{r}\right\},\left\{\mathcal{E},\frac{r_{\mathrm{s}}}{r}\right\}\right\} &= \left\{\frac{1}{\mathcal{E}}\frac{r_{\mathrm{s}}}{r}+\frac{r_{\mathrm{s}}}{r}\frac{1}{\mathcal{E}},\mathcal{E}\frac{r_{\mathrm{s}}}{r}+\frac{r_{\mathrm{s}}}{r}\mathcal{E}\right\}\\ &=2\frac{r_{\mathrm{s}}^2}{r^2}+\left(\frac{1}{\mathcal{E}}\frac{r_{\mathrm{s}}^2}{r^2}\mathcal{E}+\mathcal{E}\frac{r_{\mathrm{s}}^2}{r^2}\frac{1}{\mathcal{E}}\right)+\left\{\frac{r_{\mathrm{s}}}{r},\frac{1}{\mathcal{E}}\frac{r_{\mathrm{s}}}{r}\mathcal{E}\right\}+\left\{\frac{r_{\mathrm{s}}}{r},\mathcal{E}\frac{r_{\mathrm{s}}}{r}\frac{1}{\mathcal{E}}\right\}\\ &=4\frac{r_{\mathrm{s}}^2}{r^2}+\left\{\frac{r_{\mathrm{s}}}{r},\left(\frac{1}{\mathcal{E}}\frac{r_{\mathrm{s}}}{r}\mathcal{E}+\mathcal{E}\frac{r_{\mathrm{s}}}{r}\frac{1}{\mathcal{E}}\right)\right\}\\ &=4\frac{r_{\mathrm{s}}^2}{r^2}+\left\{\frac{r_{\mathrm{s}}}{r},2\frac{r_{\mathrm{s}}}{r}\right\}=8\frac{r_{\mathrm{s}}^2}{r^2}\,,\end{aligned} \tag{17.11}$$

where we used (17.9) twice. Applying (17.10) and (17.11) to (17.8) we find

$$\left[S_{\mathrm{ds}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}}{r}\right\}\right] = \frac{\mathrm{i}}{2}\gamma^5\, m\left(4\frac{r_{\mathrm{s}}}{r} - 2\frac{r_{\mathrm{s}}^2}{r^2}\right) = \mathrm{i}\gamma^5\, m\left(2\frac{r_{\mathrm{s}}}{r} - \frac{r_{\mathrm{s}}^2}{r^2}\right). \tag{17.12}$$

Now for the third term, we have

$$\begin{aligned}\left[S_{\mathrm{ds}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}\right] &= -\mathrm{i}\frac{m}{2}\left[\beta\gamma^5\frac{1}{\mathcal{E}} - \cancel{\beta\gamma^5\frac{1}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\}} + \cancel{\beta\gamma^5\frac{1}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}\right] \\ &= \frac{\mathrm{i}}{2}\gamma^5\, m\left\{\frac{1}{\mathcal{E}}, \left\{\mathcal{E}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}\right\} = 2\mathrm{i}\gamma^5\, m\frac{r_{\mathrm{s}}^2}{r^2}, \end{aligned} \tag{17.13}$$

where we used (17.10) in the last step. We now calculate the fourth term,

$$\begin{aligned}\left[S_{\mathrm{ds}}, \gamma^5\, m\right] &= -\mathrm{i}\frac{m}{2}\left[\beta\gamma^5\frac{1}{\mathcal{E}} - \beta\gamma^5\frac{1}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} + \beta\gamma^5\frac{1}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}, \gamma^5\, m\right] \\ &= -\mathrm{i}\beta\, m^2\left(\frac{1}{\mathcal{E}} - \frac{1}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} + \frac{1}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}\right). \end{aligned} \tag{17.14}$$

Now the fifth term reads as

$$\begin{aligned}\left[S_{\mathrm{ds}}, \gamma^5\, m\frac{r_{\mathrm{s}}}{r}\right] &= -\mathrm{i}\frac{m}{2}\left[\beta\gamma^5\frac{1}{\mathcal{E}} - \beta\gamma^5\frac{1}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} + \cancel{\beta\gamma^5\frac{1}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}}, \gamma^5\, m\frac{r_{\mathrm{s}}}{r}\right] \\ &= -\mathrm{i}\beta\, m^2\left(\frac{1}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} - \frac{1}{8}\left\{\frac{r_{\mathrm{s}}}{r}, \left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\}\right\}\right) \\ &= -\mathrm{i}\beta\, m^2\left(\frac{1}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} - \frac{1}{8}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\} - \frac{1}{4}\frac{r_{\mathrm{s}}}{r}\frac{1}{\mathcal{E}}\frac{r_{\mathrm{s}}}{r}\right). \end{aligned} \tag{17.15}$$

Finally the sixth term,

$$\begin{aligned}\left[S_{\mathrm{ds}}, \gamma^5\, m\frac{r_{\mathrm{s}}^2}{r^2}\right] &= -\mathrm{i}\frac{m}{2}\left[\beta\gamma^5\frac{1}{\mathcal{E}} - \cancel{\beta\gamma^5\frac{1}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\}} + \cancel{\beta\gamma^5\frac{1}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}}, \gamma^5\, m\frac{r_{\mathrm{s}}^2}{r^2}\right] \\ &= -\frac{\mathrm{i}}{2}\beta\, m^2\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}. \end{aligned} \tag{17.16}$$

Here we also note that

$$
\begin{aligned}
[S_{\rm ds}, \mathcal{O}_{\rm ds}] &= \left[S_{\rm ds}, \gamma^5\, m\right] - \frac{1}{2}\left[S_{\rm ds}, \gamma^5\, m\frac{r_{\rm s}}{r}\right] + \frac{1}{8}\left[S_{\rm ds}, \gamma^5\, m\frac{r_{\rm s}^2}{r^2}\right] \\
&= -\,\mathrm{i}\beta\, m^2 \left(\frac{1}{\mathcal{E}} - \frac{1}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\} + \frac{3}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} + \frac{1}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) . \qquad (17.17)
\end{aligned}
$$

Using these results, we find the single commutator to be

$$
\begin{aligned}
[S_{\rm ds}, H'_{\rm ds}] &= -\,\mathrm{i}\left(-\mathcal{O}_{\rm ds} + \gamma^5\, m\frac{r_{\rm s}}{r} - \gamma^5\frac{m}{2}\frac{r_{\rm s}^2}{r^2} - \gamma^5\frac{9\,m}{16}\frac{r_{\rm s}^2}{r^2} + \beta\frac{m^2}{\mathcal{E}} - \frac{1}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\}\right. \\
&\quad \left. +\beta\frac{3\,m^2}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} + \beta\frac{m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) \\
&= -\,\mathrm{i}\left(-\mathcal{O}_{\rm ds} + \gamma^5\, m\frac{r_{\rm s}}{r} - \gamma^5\frac{17\,m}{16}\frac{r_{\rm s}^2}{r^2} + \beta\frac{m^2}{\mathcal{E}} - \frac{1}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\}\right. \\
&\quad \left. +\beta\frac{3\,m^2}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} + \beta\frac{m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) . \qquad (17.18)
\end{aligned}
$$

The last four terms are of order $(1/\mathcal{E})$, so we can ignore them when calculating the double commutator. Thus

$$
[S_{\rm ds}, [S_{\rm ds}, H'_{\rm ds}]] = \mathrm{i}\left([S_{\rm ds}, \mathcal{O}_{\rm ds}] - \left[S_{\rm ds}, \gamma^5\, m\frac{r_{\rm s}}{r}\right] + \frac{17}{16}\left[S_{\rm ds}, \gamma^5\, m\frac{r_{\rm s}^2}{r^2}\right]\right) , \qquad (17.19)
$$

where all the commutators are familiar from (17.17), (17.15), and (17.16), respectively. Thus

$$
\begin{aligned}
[S_{\rm ds}, [S_{\rm ds}, H'_{\rm ds}]] &= -\left(-\beta\frac{m^2}{\mathcal{E}} + \beta\frac{m^2}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\} - \beta\frac{3\,m^2}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\}\right. \\
&\quad -\beta\frac{m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r} + \beta\frac{m^2}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\} - \beta\frac{m^2}{8}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} \\
&\quad \left. -\beta\frac{m^2}{4}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r} - \beta\frac{17\,m^2}{32}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\}\right) \\
&= -\left(-\beta\frac{m^2}{\mathcal{E}} + \beta\, m^2\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\}\right. \\
&\quad \left. -\beta\,\frac{27\,m^2}{32}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} - \beta\frac{3\,m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) . \qquad (17.20)
\end{aligned}
$$

We do not need to calculate the triple commutator, as is will be of high enough order that we can neglect it. Then, after the first transformation, the Hamiltonian reads as

$$\begin{aligned} H''_{\rm ds} =& \beta\mathcal{E} - \frac{\beta}{2}\left\{\mathcal{E}, \frac{r_{\rm s}}{r}\right\} + \beta\frac{9}{32}\left\{\mathcal{E}, \frac{r_{\rm s}^2}{r^2}\right\} + \mathcal{O}_{\rm ds} + \left(-\mathcal{O}_{\rm ds} + \gamma^5\, m\frac{r_{\rm s}}{r} - \gamma^5\frac{17\, m}{16}\frac{r_{\rm s}^2}{r^2}\right. \\ & \left. +\beta\frac{m^2}{\mathcal{E}} - \frac{1}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\} + \beta\frac{3\, m^2}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} + \beta\frac{m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) \\ & + \frac{1}{2}\left(-\beta\frac{m^2}{\mathcal{E}} + \beta\, m^2\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\} - \beta\,\frac{27\, m^2}{32}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} - \beta\frac{3\, m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) \\ =& \beta\left(\mathcal{E} + \frac{m^2}{2\mathcal{E}} - \frac{1}{2}\left\{\mathcal{E}, \frac{r_{\rm s}}{r}\right\} + \frac{9}{32}\left\{\mathcal{E}, \frac{r_{\rm s}}{r}\right\} - \frac{15\, m}{64}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} - \frac{m^2}{16}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) + \mathcal{O}'_{\rm ds}\,, \end{aligned} \tag{17.21}$$

where

$$\mathcal{O}'_{\rm ds} = \gamma^5\, m\frac{r_{\rm s}}{r} - \gamma^5\frac{17\, m}{16}\frac{r_{\rm s}^2}{r^2}\,. \tag{17.22}$$

The second transform is then performed using

$$S'_{\rm ds} = -{\rm i}\frac{\beta}{4}\left\{\mathcal{O}'_{\rm ds}, \frac{1}{\mathcal{E}}\right\} = -{\rm i}\frac{m}{4}\left(\beta\gamma^5\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\} - \beta\gamma^5\frac{17}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\}\right)\,. \tag{17.23}$$

Ignoring the higher–order terms, the single commutator is then

$$[S'_{\rm ds}, H''_{\rm ds}] = [S'_{\rm ds}, \beta\mathcal{E}] - \frac{1}{2}\left[S'_{\rm ds}, \beta\left\{\mathcal{E}, \frac{r_{\rm s}}{r}\right\}\right] + [S'_{\rm ds}, \mathcal{O}'_{\rm ds}]\,. \tag{17.24}$$

Using the same argument as (17.6) we find

$$[S'_{\rm ds}, \beta\mathcal{E}] = {\rm i}\mathcal{O}'_{\rm ds}\,. \tag{17.25}$$

The second term is

$$\begin{aligned}\left[S'_F, \beta\left\{\mathcal{E}, \frac{r_\mathrm{s}}{r}\right\}\right] =& -\mathrm{i}\frac{m}{4}\left[\beta\gamma^5\left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}}{r}\right\} - \cancel{\beta\gamma^5\frac{17}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}^2}{r^2}\right\}}, \beta\left\{\mathcal{E}, \frac{r_\mathrm{s}}{r}\right\}\right] \\ =& \,\mathrm{i}\gamma^5\frac{m}{4}\left\{\left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}}{r}\right\}, \left\{\mathcal{E}, \frac{r_\mathrm{s}}{r}\right\}\right\} = 2\mathrm{i}\gamma^5\, m\frac{r_\mathrm{s}^2}{r^2}\,,\end{aligned} \tag{17.26}$$

where we used (17.11) in the last step. The final term is

$$\begin{aligned}[S'_\mathrm{ds}, \mathcal{O}'_F] =& -\mathrm{i}\frac{m}{4}\left[\beta\gamma^5\left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}}{r}\right\} - \cancel{\beta\gamma^5\frac{17}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}^2}{r^2}\right\}}, \gamma^5\, m\frac{r_\mathrm{s}}{r} - \cancel{\frac{17\, m}{16}\frac{r_\mathrm{s}^2}{r^2}}\right] \\ =& -\frac{\mathrm{i}}{4}\beta\left\{\frac{r_\mathrm{s}}{r}, \left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}}{r}\right\}\right\} = -\frac{\mathrm{i}}{4}\beta\, m^2\left(\left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}^2}{r^2}\right\} + 2\frac{r_\mathrm{s}}{r}\frac{1}{\mathcal{E}}\frac{r_\mathrm{s}}{r}\right)\,.\end{aligned} \tag{17.27}$$

Thus

$$[S'_\mathrm{ds}, H''_\mathrm{ds}] = -\mathrm{i}\left(-\mathcal{O}'_\mathrm{ds} + \gamma^5\, m\frac{r_\mathrm{s}^2}{r^2} + \beta\frac{m^2}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}^2}{r^2}\right\} + \beta\frac{m^2}{2}\frac{r_\mathrm{s}}{r}\frac{1}{\mathcal{E}}\frac{r_\mathrm{s}}{r}\right)\,. \tag{17.28}$$

By inspection it is clear that the only term to contribute to the double commutator will be the first term, i.e.,

$$[S'_\mathrm{ds}, [S'_\mathrm{ds}, H''_\mathrm{ds}]] = \mathrm{i}\,[S'_\mathrm{ds}, \mathcal{O}'_\mathrm{ds}] = -\left(-\beta\frac{m^2}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}^2}{r^2}\right\} - \beta\frac{m^2}{2}\frac{r_\mathrm{s}}{r}\frac{1}{\mathcal{E}}\frac{r_\mathrm{s}}{r}\right)\,, \tag{17.29}$$

where we used the result of (17.27). Thus after two transformations the Hamiltonian is

$$\begin{aligned}H'''_\mathrm{ds} =& \,\beta\left(\mathcal{E} + \frac{m^2}{2\mathcal{E}} - \frac{1}{2}\left\{\mathcal{E}, \frac{r_\mathrm{s}}{r}\right\} + \frac{9}{32}\left\{\mathcal{E}, \frac{r_\mathrm{s}}{r}\right\} - \frac{15\, m}{64}\left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}^2}{r^2}\right\} - \frac{m^2}{16}\frac{r_\mathrm{s}}{r}\frac{1}{\mathcal{E}}\frac{r_\mathrm{s}}{r}\right) + \mathcal{O}'_\mathrm{ds} \\ &+ \left(-\mathcal{O}'_\mathrm{ds} + \gamma^5\, m\frac{r_\mathrm{s}^2}{r^2} + \beta\frac{m^2}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}^2}{r^2}\right\} + \beta\frac{m^2}{2}\frac{r_\mathrm{s}}{r}\frac{1}{\mathcal{E}}\frac{r_\mathrm{s}}{r}\right) \\ &+ \frac{1}{2}\left(-\beta\frac{m^2}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_\mathrm{s}^2}{r^2}\right\} - \beta\frac{m^2}{2}\frac{r_\mathrm{s}}{r}\frac{1}{\mathcal{E}}\frac{r_\mathrm{s}}{r}\right)\end{aligned}$$

$$
\begin{aligned}
&= \beta \left(\mathcal{E} + \frac{m^2}{2\mathcal{E}} - \frac{1}{2} \left\{ \mathcal{E}, \frac{r_{\rm s}}{r} \right\} + \frac{9}{32} \left\{ \mathcal{E}, \frac{r_{\rm s}}{r} \right\} \right. \\
&\quad \left. - \frac{7\, m^2}{64} \left\{ \frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2} \right\} - \frac{3\, m^2}{16} \frac{r_{\rm s}}{r} \frac{1}{\mathcal{E}} \frac{r_{\rm s}}{r} \right) + \mathcal{O}''_{\rm ds} \,,
\end{aligned}
\tag{17.30}
$$

where

$$
\mathcal{O}''_{\rm ds} = \gamma^5 \, m \frac{r_{\rm s}^2}{r^2} \,. \tag{17.31}
$$

A third transformation is required, and will serve to eliminate the odd term, without contributing any further terms. Thus after three iterations of the ultrarelativistic transform we find

$$
\mathcal{H}_{\rm ds} = \beta \left(\mathcal{E} + \frac{m^2}{2\mathcal{E}} - \frac{1}{2} \left\{ \mathcal{E}, \frac{r_{\rm s}}{r} \right\} + \frac{9}{32} \left\{ \mathcal{E}, \frac{r_{\rm s}}{r} \right\} - \frac{7\, m^2}{64} \left\{ \frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2} \right\} + \frac{3\, m^2}{16} \frac{r_{\rm s}}{r} \frac{1}{\mathcal{E}} \frac{r_{\rm s}}{r} \right) \,, \tag{17.32}
$$

for the subluminal (tardyonic) gravitationally coupled high–energy Dirac particle.

17.2. TACHYONIC GRAVITATIONALLY COUPLED PARTICLE

We now apply the ultrarelativistic transform to a gravitationally coupled tachyon, keeping terms to the second order in gravity, i.e., $(r_{\rm s})^2$. Utilizing our initial rotation U_1 to transform (15.20) into the Weyl basis we obtain

$$
H'_{\rm tg} = \beta \mathcal{E} - \frac{\beta}{2} \left\{ \mathcal{E}, \frac{r_{\rm s}}{r} \right\} + \frac{9}{32} \beta \left\{ \mathcal{E}, \frac{r_{\rm s}^2}{r^2} \right\} + \mathcal{O}_{\rm tg} \,, \tag{17.33}
$$

where we have already collected the odd terms

$$
\mathcal{O}_{\rm tg} = \beta \, \gamma^5 \, m \left(1 - \frac{r_{\rm s}}{2r} + \frac{r_{\rm s}^2}{8r^2} \right) \,. \tag{17.34}
$$

We can then construct the operator

$$S_{\mathrm{tg}} = -\mathrm{i}\frac{\beta}{4}\left\{\mathcal{O}_{\mathrm{tg}}, \frac{1}{\mathcal{E}}\right\} = -\frac{\mathrm{i}}{2}m\left(\frac{\gamma^5}{\mathcal{E}} - \frac{\gamma^5}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} + \frac{\gamma^5}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}\right), \tag{17.35}$$

and begin the transformation. Again we are keeping terms to the second order in gravity, and to the first order in the inverse of $\mathcal{E}$. Canceled terms are of high enough order that they can be neglected. We begin with the single commutator

$$\begin{aligned}\left[S_{\mathrm{tg}}, H'_{\mathrm{tg}}\right] = {}& [S_{\mathrm{tg}}, \beta\mathcal{E}] - \frac{1}{2}\left[S_{\mathrm{tg}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}}{r}\right\}\right] + \frac{9}{32}\left[S_{\mathrm{tg}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}\right] \\ & + \left[S_{\mathrm{tg}}, \beta\gamma^5 m\right] - \frac{1}{2}\left[S_{\mathrm{tg}}, \beta\gamma^5 m\frac{r_{\mathrm{s}}}{r}\right] + \frac{1}{8}\left[S_{\mathrm{tg}}, \beta\gamma^5 m\frac{r_{\mathrm{s}}^2}{r^2}\right],\end{aligned} \tag{17.36}$$

giving us six terms to calculate. Based on our work in the previous chapter the first term is trivially found to be

$$[S_{\mathrm{tg}}, \beta\mathcal{E}] = \mathrm{i}\mathcal{O}_{\mathrm{tg}}\,, \tag{17.37}$$

for the exact same reason as in equation (17.6). We then turn our attention to the second term,

$$\begin{aligned}\left[S_{\mathrm{tg}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}}{r}\right\}\right] = {}& -\mathrm{i}\frac{m}{2}\left[\frac{\gamma^5}{\mathcal{E}} - \frac{\gamma^5}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} + \cancel{\frac{\gamma^5}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}}{r}\right\}\right] \\ = {}& \frac{\mathrm{i}}{2}\beta\gamma^5 m\left(\left\{\frac{1}{\mathcal{E}}, \left\{\mathcal{E}, \frac{r_{\mathrm{s}}}{r}\right\}\right\} - \frac{1}{4}\left\{\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\}, \left\{\mathcal{E}, \frac{r_{\mathrm{s}}}{r}\right\}\right\}\right) \\ = {}& \mathrm{i}\beta\gamma^5 m\left(2\frac{r_{\mathrm{s}}}{r} - \frac{r_{\mathrm{s}}^2}{r^2}\right),\end{aligned} \tag{17.38}$$

where we used (17.10) and (17.11). We can now move on to the third term

$$\begin{aligned}
\left[S_{\mathrm{tg}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}^{2}}{r^{2}}\right\}\right] &= -\mathrm{i}\frac{m}{2}\left[\frac{\gamma^{5}}{\mathcal{E}} - \cancel{\frac{\gamma^{5}}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\}} + \cancel{\frac{\gamma^{5}}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^{2}}{r^{2}}\right\}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}^{2}}{r^{2}}\right\}\right] \\
&= \frac{\mathrm{i}}{2}\beta\gamma^{5}\, m\left\{\frac{1}{\mathcal{E}}, \left\{\mathcal{E}, \frac{r_{\mathrm{s}}^{2}}{r^{2}}\right\}\right\} \\
&= 2\mathrm{i}\beta\gamma^{5}\, m\frac{r_{\mathrm{s}}^{2}}{r^{2}}\,,
\end{aligned} \tag{17.39}$$

where we used (17.10) in the last step. We can calculate the fourth term

$$\begin{aligned}
\left[S_{\mathrm{tg}}, \beta\gamma^{5}\, m\right] &= -\mathrm{i}\frac{m}{2}\left[\frac{\gamma^{5}}{\mathcal{E}} - \frac{\gamma^{5}}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} + \frac{\gamma^{5}}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^{2}}{r^{2}}\right\}, \beta\gamma^{5}\, m\right] \\
&= \mathrm{i}\beta\, m^{2}\left(\frac{1}{\mathcal{E}} - \frac{1}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} + \frac{1}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^{2}}{r^{2}}\right\}\right)\,.
\end{aligned} \tag{17.40}$$

Then the fifth term

$$\begin{aligned}
\left[S_{\mathrm{tg}}, \beta\gamma^{5}\, m\frac{r_{\mathrm{s}}}{r}\right] &= -\mathrm{i}\frac{m}{2}\left[\frac{\gamma^{5}}{\mathcal{E}} - \frac{\gamma^{5}}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} + \cancel{\frac{\gamma^{5}}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^{2}}{r^{2}}\right\}}, \beta\gamma^{5}\, m\frac{r_{\mathrm{s}}}{r}\right] \\
&= \mathrm{i}\beta\, m^{2}\left(\frac{1}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} - \frac{1}{8}\left\{\frac{r_{\mathrm{s}}}{r}, \left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\}\right\}\right) \\
&= \mathrm{i}\beta\, m^{2}\left(\frac{1}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} - \frac{1}{8}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^{2}}{r^{2}}\right\} - \frac{1}{4}\frac{r_{\mathrm{s}}}{r}\frac{1}{\mathcal{E}}\frac{r_{\mathrm{s}}}{r}\right)\,.
\end{aligned} \tag{17.41}$$

Finally the sixth term,

$$\begin{aligned}
\left[S_{\mathrm{tg}}, \beta\gamma^{5}\, m\frac{r_{\mathrm{s}}^{2}}{r^{2}}\right] &= -\mathrm{i}\frac{m}{2}\left[\frac{\gamma^{5}}{\mathcal{E}} - \cancel{\frac{\gamma^{5}}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\}} + \cancel{\frac{\gamma^{5}}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^{2}}{r^{2}}\right\}}, \beta\gamma^{5}\, m\frac{r_{\mathrm{s}}^{2}}{r^{2}}\right] \\
&= \mathrm{i}\beta\,\frac{m^{2}}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^{2}}{r^{2}}\right\}\,.
\end{aligned} \tag{17.42}$$

We now take a moment, and note that the last three terms can be combined to form the commutator of the operator $S_{\rm tg}$ and the odd part $\mathcal{O}_{\rm tg}$, i.e.,

$$
\begin{aligned}
[S_{\rm tg}, \mathcal{O}_{\rm tg}] = & \left[S_{\rm tg}, \beta\gamma^5\, m\right] - \frac{1}{2}\left[S_{\rm tg}, \beta\gamma^5\, m\frac{r_{\rm s}}{r}\right] + \frac{1}{8}\left[S_{\rm tg}, \beta\gamma^5\, m\frac{r_{\rm s}^2}{r^2}\right] \\
= & \,\mathrm{i}\beta\, m^2\left(\frac{1}{\mathcal{E}} - \frac{1}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\} + \frac{3}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} + \frac{1}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right)\,. \qquad (17.43)
\end{aligned}
$$

The single commutator is then

$$
\begin{aligned}
[S_{\rm tg}, H_{\rm tg}] = & -\mathrm{i}\left(-\mathcal{O}_{\rm tg} + \beta\gamma^5\, m\frac{r_{\rm s}}{r} - \beta\gamma^5\, m\frac{1}{2}\frac{r_{\rm s}^2}{r^2} - \beta\gamma^5\, m\frac{9}{16}\frac{r_{\rm s}^2}{r^2}\right. \\
& \left. -\beta\frac{m^2}{\mathcal{E}} + \beta\frac{m^2}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\} - \beta\frac{3\,m^2}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} - \beta\frac{m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) \\
= & -\mathrm{i}\left(-\mathcal{O}_{\rm tg} + \beta\gamma^5\, m\frac{r_{\rm s}}{r} - \beta\gamma^5\frac{17\, m}{16}\frac{r_{\rm s}^2}{r^2} - \beta\frac{m^2}{\mathcal{E}} + \beta\frac{m^2}{2}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}}{r}\right\}\right. \\
& \left. -\beta\frac{3\,m^2}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} - \beta\frac{m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right)\,. \qquad (17.44)
\end{aligned}
$$

The last four terms are of order $(1/\mathcal{E})$, so we can ignore them when calculating the double commutator. Thus

$$
\left[S_{\rm tg}, \left[S_{\rm tg}, H'_{\rm tg}\right]\right] = \mathrm{i}\left([S_{\rm tg}, \mathcal{O}_{\rm tg}] - \left[S_{\rm tg}, \beta\gamma^5\, m\frac{r_{\rm s}}{r}\right] + \frac{17}{16}\left[S_{\rm tg}, \beta\gamma^5\, m\frac{r_{\rm s}^2}{r^2}\right]\right)\,, \qquad (17.45)
$$

where we have seen all of these terms before, i.e., in equations (17.43), (17.41), and (17.42), respectively. Thus we quickly find that

$$
\begin{aligned}
\left[S_{\rm tg},\left[S_{\rm tg},H'_{\rm tg}\right]\right] = & -\Bigg(\beta\frac{m^2}{\mathcal{E}} - \beta\frac{m^2}{2}\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}}{r}\right\} + \beta\frac{3m^2}{16}\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}^2}{r^2}\right\} \\
& +\beta\frac{m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r} - \beta\frac{m^2}{2}\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}}{r}\right\} + \beta\frac{m^2}{8}\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}^2}{r^2}\right\} \\
& +\beta\frac{m^2}{4}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r} + \beta\frac{17\,m^2}{32}\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}^2}{r^2}\right\}\Bigg) \\
= & -\left(\beta\frac{m^2}{\mathcal{E}} - \beta\, m^2\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}}{r}\right\} + \beta\frac{27\,m^2}{32}\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}^2}{r^2}\right\} + \beta\frac{3\,m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) .
\end{aligned}
\tag{17.46}
$$

After the first transformation the Hamiltonian is

$$
\begin{aligned}
H''_{\rm tg} = & \beta\mathcal{E} - \frac{\beta}{2}\left\{\mathcal{E},\frac{r_{\rm s}}{r}\right\} + \beta\frac{9}{32}\left\{\mathcal{E},\frac{r_{\rm s}^2}{r^2}\right\} + \mathcal{O}_{\rm tg} + \Bigg(-\mathcal{O}_{\rm tg} + \beta\gamma^5\, m\frac{r_{\rm s}}{r} - \beta\gamma^5\frac{17\,m}{16}\frac{r_{\rm s}^2}{r^2} \\
& -\beta\frac{m^2}{\mathcal{E}} + \beta\frac{m^2}{2}\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}}{r}\right\} - \beta\frac{3\,m^2}{16}\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}^2}{r^2}\right\} - \beta\frac{m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\Bigg) \\
& + \frac{1}{2}\left(\beta\frac{m^2}{\mathcal{E}} - \beta\, m^2\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}}{r}\right\} + \beta\frac{27\,m^2}{32}\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}^2}{r^2}\right\} + \beta\frac{3\,m^2}{8}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) \\
= & \beta\Bigg(\mathcal{E} - \frac{m^2}{2\mathcal{E}} - \frac{1}{2}\left\{\mathcal{E},\frac{r_{\rm s}}{r}\right\} + \frac{9}{32}\left\{\mathcal{E},\frac{r_{\rm s}^2}{r^2}\right\} \\
& + \frac{15\,m^2}{64}\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}^2}{r^2}\right\} + \frac{m^2}{16}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\Bigg) + \mathcal{O}'_{\rm tg} ,
\end{aligned}
\tag{17.47}
$$

where

$$
\mathcal{O}'_{\rm tg} = \beta\gamma^5\, m\frac{r_{\rm s}}{r} - \beta\gamma^5\frac{17\,m}{16}\frac{r_{\rm s}^2}{r^2} . \tag{17.48}
$$

We need to perform a second transformation, with

$$
S'_{\rm tg} = -{\rm i}\frac{\beta}{4}\left\{\mathcal{O}'_{\rm tg},\frac{1}{\mathcal{E}}\right\} = -{\rm i}\frac{m}{4}\left(\gamma^5\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}}{r}\right\} - \gamma^5\frac{17}{16}\left\{\frac{1}{\mathcal{E}},\frac{r_{\rm s}^2}{r^2}\right\}\right) . \tag{17.49}
$$

Ignoring all the terms outside our sphere of interest, we find

$$\left[S'_{\mathrm{tg}}, H''_{\mathrm{tg}}\right] = \left[S'_{\mathrm{tg}}, \beta\mathcal{E}\right] - \frac{1}{2}\left[S'_{\mathrm{tg}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}}{r}\right\}\right] + \left[S'_{\mathrm{tg}}, \mathcal{O}'_{\mathrm{tg}}\right] . \tag{17.50}$$

By an argument identical to that found in (17.6), we find that

$$\left[S'_{\mathrm{tg}}, \mathcal{O}'_{\mathrm{tg}}\right] = \mathrm{i}\mathcal{O}'_{\mathrm{tg}} . \tag{17.51}$$

We then turn our attention to the second term

$$\begin{aligned}\left[S'_{\mathrm{tg}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}}{r}\right\}\right] &= -\mathrm{i}\frac{m}{4}\left[\gamma^5\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} - \cancel{\gamma^5\frac{17}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}}, \beta\left\{\mathcal{E}, \frac{r_{\mathrm{s}}}{r}\right\}\right] \\ &= \mathrm{i}\beta\gamma^5\frac{m}{4}\left\{\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\}, \left\{\mathcal{E}, \frac{r_{\mathrm{s}}}{r}\right\}\right\} = 2\mathrm{i}\beta\gamma^5\, m\frac{r_{\mathrm{s}}^2}{r^2} ,\end{aligned} \tag{17.52}$$

where we used (17.11) for the last step. We now calculate the final term from equation (17.50),

$$\begin{aligned}\left[S'_{\mathrm{tg}}, \mathcal{O}'_{\mathrm{tg}}\right] &= -\mathrm{i}\frac{m}{4}\left[\gamma^5\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\} - \cancel{\gamma^5\frac{17}{16}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\}}, \beta\gamma^5\, m\frac{r_{\mathrm{s}}}{r} - \cancel{\beta\gamma^5\,\frac{17\, m\, r_{\mathrm{s}}^2}{16\ \ r^2}}\right] \\ &= \frac{\mathrm{i}}{4}\beta\, m^2\left\{\frac{r_{\mathrm{s}}}{r}, \left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}}{r}\right\}\right\} = \frac{\mathrm{i}}{4}\beta\, m^2\left(\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\} + 2\frac{r_{\mathrm{s}}}{r}\frac{1}{\mathcal{E}}\frac{r_{\mathrm{s}}}{r}\right) .\end{aligned} \tag{17.53}$$

Thus

$$\left[S'_{\mathrm{tg}}, H''_{\mathrm{tg}}\right] = -\mathrm{i}\left(-\mathcal{O}'_{\mathrm{tg}} + \beta\gamma^5\, m\frac{r_{\mathrm{s}}^2}{r^2} - \beta\frac{m^2}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\mathrm{s}}^2}{r^2}\right\} - \beta\frac{m^2}{2}\frac{r_{\mathrm{s}}}{r}\frac{1}{\mathcal{E}}\frac{r_{\mathrm{s}}}{r}\right) . \tag{17.54}$$

By inspection it is clear that only the first term will contribute to the double commutator, i.e.,

$$\left[S'_{\mathrm{tg}}, \left[S'_{\mathrm{tg}}, H''_{\mathrm{tg}}\right]\right] = \mathrm{i}\left[S'_{\mathrm{tg}}, \mathcal{O}'_{\mathrm{tg}}\right] , \tag{17.55}$$

and we already calculated this commutator (17.53), thus

$$\left[S'_{\rm tg}, \left[S'_{\rm tg}, H''_{\rm tg}\right]\right] = -\left(\beta\frac{m^2}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} + \beta\frac{m^2}{2}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) . \tag{17.56}$$

Thus after two transformations the Hamiltonian is

$$\begin{aligned} H'''_{\rm tg} &= \beta\left(\mathcal{E} - \frac{m^2}{2\mathcal{E}} - \frac{1}{2}\left\{\mathcal{E}, \frac{r_{\rm s}}{r}\right\} + \frac{9}{32}\left\{\mathcal{E}, \frac{r_{\rm s}^2}{r^2}\right\} + \frac{15\,m^2}{64}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} + \frac{m^2}{16}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) \\ &\quad + \mathcal{O}'_{\rm tg} + \left(-\mathcal{O}'_{\rm tg} + \beta\gamma^5\, m\frac{r_{\rm s}^2}{r^2} - \beta\frac{m^2}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} - \beta\frac{m^2}{2}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) \\ &\quad + \frac{1}{2}\left(\beta\frac{m^2}{4}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} + \beta\frac{m^2}{2}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) \\ &= \beta\left(\mathcal{E} - \frac{m^2}{2\mathcal{E}} - \frac{1}{2}\left\{\mathcal{E}, \frac{r_{\rm s}}{r}\right\} + \frac{9}{32}\left\{\mathcal{E}, \frac{r_{\rm s}^2}{r^2}\right\}\right. \\ &\quad \left. + \frac{7\,m^2}{64}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} - \frac{3\,m^2}{16}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) + \mathcal{O}''_{\rm tg} , \end{aligned} \tag{17.57}$$

where

$$\mathcal{O}''_{\rm tg} = \beta\gamma^5\, m\frac{r_{\rm s}^2}{r^2} . \tag{17.58}$$

A third transformation is required, and will serve only to eliminate the remaining odd part. Thus after three ultrarelativistic decoupling transforms, we finally find for the tachyonic gravitationally coupled Dirac Hamiltonian, the expansion

$$\mathcal{H}_{\rm tg} = \beta\left(\mathcal{E} - \frac{m^2}{2\mathcal{E}} - \frac{1}{2}\left\{\mathcal{E}, \frac{r_{\rm s}}{r}\right\} + \frac{9}{32}\left\{\mathcal{E}, \frac{r_{\rm s}^2}{r^2}\right\} + \frac{7\,m^2}{64}\left\{\frac{1}{\mathcal{E}}, \frac{r_{\rm s}^2}{r^2}\right\} - \frac{3\,m^2}{16}\frac{r_{\rm s}}{r}\frac{1}{\mathcal{E}}\frac{r_{\rm s}}{r}\right) . \tag{17.59}$$

Which is similar in structure to the tardyonic example, $\mathcal{H}_{\rm ds}$, with only sign differences. The first sign difference is the correction term $-m^2/(2\mathcal{E})$, and is a result of the tachyonic dispersion relation. Additionally the final two term have opposite signs when compared to the final two terms of $\mathcal{H}_{\rm ds}$, indicating a difference in the gravitational interaction of tachyons and tardyons. However, the leading order gravitational terms are identical ($\frac{1}{2}\left\{\mathcal{E}, \frac{r_{\rm s}}{r}\right\}$), indicating that both tardyons, and somewhat surprisingly, tachyons are attracted by a gravitational center, in the high–energy limit, in the same sense as light beams are gravitationally lensed by heavy stars (see appendix D.6).

A few interesting properties come to light when one considers both of the ultra-relativistic decoupled, gravitationally coupled Hamiltonians (the subluminal (17.32) and the superluminal (17.59)). First we note that both equations have a β prefactor, which indicates that in both cases we have particle–antiparticle symmetry. Second, when we compare the two equations we find that all the identical terms are proportional to $\mathcal{E}$, as defined in (16.12), and are independent of the mass. On the other hand, the terms with opposite signs are proportional to the inverse of $\mathcal{E}$, and all have a factor of mass squared, m^2 (note that in both cases m is a real, positive number). Thus we conclude that the kinetic terms of the equations account for the similarities, while the gravitational mass terms distinguish the two Hamiltonians.

18. (PARTIAL) CONCLUSIONS

In chapters 16 and 17, we introduce the ultrarelativistic decoupling transformation, which can be applied to tachyons as well as tardyons. The underlying procedure is more complicated than that of the Foldy–Wouthuysen transformation, as we must first ensure that we are working in a basis which is suited to the ultrarelativistic decoupling transform, the Weyl basis for example. It may be necessary to transform into such a basis. When dealing with pseudo–Hermitian Hamiltonians, such a transform may change the type of pseudo–Hermiticity obeyed by the Hamiltonian. In the examples given, the Hamiltonians are transformed from γ^5–Hermitian to β–Hermitian. Although one can say the Hamiltonian went from γ^5–Hermitian in the Dirac representation, to γ^5_{W}–Hermitian in the Weyl basis (see appendix A of [93]).

Like the Foldy–Wouthuysen transformation, we find that an exact ultrarelativistic decoupling transform exists for free tachyons as well as free tardyons (see chapter 16). However, it seems as if that is as far as the applicability of the exact transform goes. For more complicated Hamiltonians, a perturbative method must be employed, as we demonstrated with both the gravitationally coupled tardyon and tachyon.

Surprisingly, we have found that in the high energy limit tachyons are attracted by gravity. This is in sharp contrast to the classical result, in which tachyons are repulsed by gravity. The disparity may result from the fact that in the high energy limit, the tachyons are light–like, i.e., the momentum term dominates the mass term, and as is demonstrated by the observable effects, light is attracted by gravity. This means that the light barrier does not define the transition between particles being attracted by and repulsed by gravity. This actually is obvious because the light cone is approached for both high-energy tardyons as well as high-energy tachyons;

the dispersion relation becomes $E = \sqrt{\vec{p}^2 + m^2} \approx |\vec{p}|$ or $E = \sqrt{\vec{p}^2 - m^2} \approx |\vec{p}|$ for $|\vec{p}| \gg m$, respectively. Luxons (photons) traveling exactly at the speed of light are known to be gravitationally lensed (see section 6.3 of [126]).

Part IV

Conclusions

Through the course of our investigations we have examined, and utilized, a number of approximation methods in quantum mechanics. These range from numerical approximation of nonrelativistic quantum mechanics, to analytic, ultrarelativistic approximations for both tardyons and tachyons. In all cases, we transform Hamiltonians into an intuitively more understandable form, where the physical degrees of freedom are better displayed, and the operators obtain a more intuitive interpretation. This is perhaps most obvious when we consider the nonrelativistic corrections to the Dirac–Einstein–Schwarzschild Hamiltonian (9.76), obtained via a Foldy–Wouthuysen transformation. The resulting nonrelativistic approximation lends itself quite nicely to physical interpretation. This is partially due to the fact that the nonrelativistic approximation to the gravitational Hamiltonian has a similar structure when compared to the well known nonrelativistic limit of the Dirac–Coulomb Hamiltonian [23, 60].

In part I we begin by investigating three classes of Hamiltonians, Hermitian, pseudo–Hermitian, and $\mathcal{PT}$–symmetric. We work to determine what, if any, relation they all have to each other. Rather than creating a new, independent class of Hamiltonians, pseudo–Hermiticity extends the definition of Hermiticity [1]. Then by the very nature of pseudo–Hermiticity, it is clear that Hermitian operators must be a subset pseudo–Hermitian operators (i.e., a Hermitian Hamiltonian is "1–Hermitian" in the sense of the definition given in equation (2.13), setting $\eta = 1$ equal to the unit operator). By considering the example cases of the real and imaginary cubic anharmonic oscillators (equations (2.39) and (3.1), respectively) we are quickly able to determine that neither $\mathcal{PT}$–symmetry nor Hermiticity is a subset of the other. The real cubic anharmonic oscillator is Hermitian but not $\mathcal{PT}$–symmetric, while the imaginary cubic anharmonic oscillator is not Hermitian, but is $\mathcal{PT}$–symmetric. By the transitive relation it is then clear that the set of pseudo–Hermitian operators is not a subset of the set $\mathcal{PT}$–symmetric operators. We finally to consider whether or not the set of $\mathcal{PT}$–symmetric operators is a subset of the pseudo–Hermitian operators. The "obvious identification" of $\mathcal{PT}$–symmetry being equivalent to $\mathcal{P}$–Hermiticity, comes from the consideration of Hamiltonians of the form $H = \vec{p}^{\,2}/(2m) + V$, in which case $\mathcal{T}\,H\,\mathcal{T} = H^{+}$, and H will then be $\mathcal{P}$–Hermitian. By looking at examples, the simplest being a trivial model ("toy") Hamiltonian, consisting of only the momentum operator, $H_p = p = -\mathrm{i}\,\partial_x$, we find that $\mathcal{PT}$–symmetry does not imply $\mathcal{P}$–Hermiticity. However, all the examples we consider are pseudo–Hermitian in some way. As such, it is clear that the set of $\mathcal{PT}$–symmetric operators is not a subset of the $\mathcal{P}$–Hermitian operators, but may be a subset of the pseudo–Hermitian operators. The two concepts are clearly related, and constitute viable alternatives to Hermiticity.

Additionally, we investigate Hermitizing transforms, which map an exact $\mathcal{PT}$–symmetric Hamiltonian onto a Hermitian Hamiltonian [19, 31–34], order by order in an expansion parameter. The Hermitizing transformation conserves the eigenvalues, and would otherwise seem to suggest that $\mathcal{PT}$–symmetry and Hermiticity are "equivalent." However, the calculation of these transforms is perturbative in nature, and generally leads to a much more complicated, non–local Hermitian Hamiltonian [15, 19]. Under such a transform, the original $\mathcal{PT}$–symmetric Hamiltonian is generally easier to work with. Moreover, in chapter 2.4 we show that the transformation is necessarily non–unitary, and fails to conserve parity. By analytically calculating the metric of a $\mathcal{PT}$–symmetric Hamiltonian, it has been shown that $p^{\,2}+\mathrm{i}\,x^3$ cannot be similar to any Hermitian Hamiltonian [21]. We also consider the differences between the physical interpretation of a $\mathcal{PT}$–symmetric Hamiltonian (a system in which the gain and loss terms are in equilibrium) and a Hermitian Hamiltonian (a closed system). Finally, by numerically calculating a set of $\mathcal{PT}$–symmetric wave–functions (see chapter 3), we develope an intuitive picture which is incompatible with that of a Hermitian wave–function. We are left to conclude that $\mathcal{PT}$–symmetry and pseudo–Hermiticity are independent concepts.

In chapter 3 we aim to develop an intuitive picture of $\mathcal{PT}$–symmetric eigenstates, which bear some similarities to eigenstates of a Hermitian Hamiltonian. While the $\mathcal{PT}$–symmetric wave–functions correspond to manifestly complex potentials, we find that the modulus of the potential, which tends to infinity as $x \to \pm\infty$, confines the wave–function, much like one would expect in the "classically allowed region" (see figure 3.3). While nodes can be used to enumerate Hermitian wave–functions, $\mathcal{PT}$–symmetric wave–functions do not have any complex zeroes. However, the modulus of these wave–functions do have local minima where we would expect to see nodes based on the Hermitian picture, as reported in figures 3.4, 3.5, and 3.7. This allows for a possible method of enumerating the $\mathcal{PT}$–symmetric wave–functions.

Despite the similarities, there are some rather stark differences as well. While the local minima may provide an opportunity to enumerate the wave–functions, the fact that there are no complex zeroes is an indication that $\mathcal{PT}$–symmetry is independent of Hermiticity. Furthermore, there are an infinite number of both real and imaginary zeroes, as reported in figure 3.6. Finally, we see that Hermitian Hamiltonians are governed by the concavity condition. Due to the fact that the $\mathcal{PT}$–symmetric potential is complex, no such condition can be imposed on the associated wave–functions. These differences serve to distinguish $\mathcal{PT}$–symmetry from Hermiticity, and ramify our conclusion that it constitutes an independent concept.

In order to obtain these results, we use an easily scalable matrix diagonalization algorithm which is specially suited for densely populated complex symmetric matrices. The algorithm diagonalizes an input matrix in two steps. First, it tridiagonalizes the input matrix, and then it diagonalizes the resulting tridiagonal matrix using an implicit shift. The user can implement the algorithm using the no–shift option (which is not recommended), or using a linear, quadratic, or cubic shift. Numerical evidence suggests that the cubic and quadratic shifts are in turn the most efficient, depending on the structure of the matrix to be diagonalized.

Finally, let us compare to routines within publicly accessible libraries (e.g., `LAPACK` [61]) which often act as black boxes, without a detailed discussion of the algorithmic steps on which they are based. For typical applications (matrices around rank 500), we find that our `HTDQLS` routine is somewhat faster than `LAPACK`'s routine `ZGEEVX`. Furthermore, it can be challenging to alter `LAPACK`'s precision, while our algorithm was written with transparency and ease of scalability in mind. It doesn't matter how fast a `LAPACK` routine can calculate the eigenvalues and/or eigenvectors if more then 16 digit precision is required.

In part II we investigate a number of generalized Dirac Hamiltonians using the standard Foldy–Wouthuysen transformation [22] and in doing so we obtain five new results along with some rather well known results, e.g., the nonrelativistic limit of the free Dirac particle, and the nonrelativistic limit of the Dirac–Coulomb Hamiltonian. We find the nonrelativistic limit of the Dirac–Einstein–Schwarzschild Hamiltonian (9.76) and find that the resulting Hamiltonian has a similar structure to the well known result of the transformed Dirac–Coulomb Hamiltonian [23, 60] (also see equation (9.33)). The leading terms are the usual kinetic corrections for a free particle, while the second is instantly recognizable as the gravitational potential. The associated β prefactor ensures that *both* particles and antiparticles are attracted by gravity. Additionally there is the gravitational analog to the zitterbewegung (Darwin) term, as well as gravitational spin–orbit coupling, otherwise known as Fokker precession, which is in full agreement with the classical result, which has in turn been confirmed by Gravity Probe B [119]. Overall, there is particle–antiparticle symmetry, ensuring that both particles and antiparticle behave the same when in a gravitational potential. This is in contrast to the result obtained when the chiral Foldy–Wouthuysen transform is applied, which would otherwise imply that particles and antiparticles behave differently when in a gravitational field (see chapter 10.6). We find the corrections, up to the fourth order in momenta, of the Dirac Hamiltonian with a scalar potential (9.48). This Hamiltonian exhibits a surprising $\{\vec{p}^2, 1/r\}$ term, despite the similarities of the initial Hamiltonain and the untransformed Dirac–Coulomb Hamiltonian. Again to the forth order in momenta, we find the relativistic corrections to the Dirac Hamiltonian with a scalar confining potential (9.58). Again we find that the transformed Hamiltonian exhibits an anticommutator term as a kinetic correction. We find a compact representation of the Foldy–Wouthuysen transformed Dirac Hamiltonian in a rotating non–inertial frame (9.108). We confirm that the Mashhoon term [103] is unaffected by the transformation up to the fourth order in the momenta. Finally, we apply the rotations from the Dirac–Einstein–Schwarzschild Hamiltonian to the gravitationally coupled transition current (9.125). In addition to the known corrections terms, there is an additional gravitational kinetic correction, as well as gravitational corrections to the magnetic coupling.

In chapter 10 we apply the chiral Foldy–Wouthuysen transformation [24] to the same set of generalized Dirac Hamiltonians investigated in chapter 9. The chiral method utilizes some interesting properties of Dirac algebra to decouple the particle and antiparticle degrees of freedom. Additionally, for the chiral transform to be unitary, the input Hamiltonian must anticommute with the chiral operator $J = \mathrm{i}\,\gamma^5 \beta$ (see chapter 10.1). We find that the results obtained using the standard and chiral transformation agree ("accidentally") for the free Dirac Hamiltonian (chapter 10.2). In all other cases, the chiral transform introduces spurious parity breaking terms (chapters 10.3–10.7). Perhaps this is not so surprising when considering the Dirac–Coulomb Hamiltonian, which does not anticommute with J (chapter 10.3), and as such the chiral transform is not unitary. The remainder of the generalized Dirac–Hamiltonians we consider (chapters 10.4–10.7) do anticommute with J, so the chiral transform is unitary. Both the standard and chiral Foldy–Wouthuysen transformation utilize unitary transforms (in specific cases), and therefore produce Hermitian Hamiltonians connected by a unitary transform. The results should then be equivalent. Yet the results contain conflicting terms and fulfill different symmetry relations. This is due to the chiral transformation U as defined in equations (10.1) and (10.2), which breaks parity, altering the fundamental symmetries of the Hamiltonian (see chapter 10.8). Despite the "seductive" elegance of the chiral Foldy–Wouthuysen transformation, we find that the standard Foldy–Wouthuysen transformation is a more reliable choice when decoupling the particle and antiparticle degrees of freedom in the nonrelativistic limit.

Finally, in part III we aim to gain a better understanding of Dirac Hamiltonians in the high-energy limit, including Dirac Hamiltonians of pseudo–Hermitian form. The results profit from the preparations described in part I and II, where the concepts of pseudo–Hermiticity and the Foldy–Wouthuysen transformation (nonrelativistic decoupling transformation) have been described. We find results for the ultrarelativistic decoupling of the free Dirac Hamiltonian, the gravitationally coupled Dirac Hamiltonian, and their pseudo–Hermitian variants, i.e., tachyons. Inspired by the Foldy–Wouthuysen transformation, we develop an ultrarelativistic decoupling transform. An exact variation is used on the set of free particles, while a perturbative approach is required for the more complicated example of gravitationally coupled particles. Surprisingly, we find that while tachyons and tardyons are affected differently by the gravitational source, both are attracted by gravity in the hight energy limit. This result contradicts the classical theory, in which tachyons are repulsed by gravitational fields [30]. This result does not imply that all tachyons are attracted by gravity, but it applies first and foremost to the high-energy limit, where particles travel close to the speed of light. The result otherwise implies that in the high-energy limit, the light barrier does not necessarily define a transition region between particles being attracted to or repulsed by gravity. By comparing the subluminal ultrarelativistic corrections to the gravitational interaction (17.32) to the superluminal ultrarelativistic corrections to the gravitational coupling (17.59) we notice that the expressions are almost identical, save for the signs of some of the higher-order terms. We note that the terms proportional to the operator $\mathcal{E}$, with no mass terms, are identical, while the terms proportional to the inverse of $\mathcal{E}$, all of which carry a m^2 dependence, carry the opposite sign. This tells us that the kinetic part of both the tardyon and tachyon interact with gravity in the exact same way, while the mass terms give rise to repulsive interactions for tachyons when the tardyon term would otherwise be attractive.

APPENDIX A

EXPLICIT FORTRAN IMPLEMENTATION OF HTDQLS AND HTDQRS

A.1. IMPLEMENTATION OF HTDQLS

A.1.1. Control Sequence. The subroutine CS allows the user to choose between computing only eigenvalues, or computing both the eigenvalues and eigenvectors (which will take longer). If JOBZ='N' then only eigenvalues are computed, while if JOBZ='V' then both eigenvalues and eigenvectors are computed. The eigenvector located in the i^{th} column of A corresponds to the eigenvalue in the i^{th} position of D. There is also the option to sort the eigenvalues and eigenvectors. If the logical SORTFLAG is true then the sorting will take place. Conversely, there will not be any sorting if it is false.

```
SUBROUTINE HTDQLS(JOBZ, N, A, D, Z, SORTFLAG, SHIFTMODE)
IMPLICIT NONE
CHARACTER JOBZ
INTEGER N
COMPLEX*32 A(N, N), Z(N, N), D(N)
LOGICAL SORTFLAG
INTEGER SHIFTMODE
COMPLEX*32 E(0:N)
CALL COPYM(N, A, Z)
IF (JOBZ .EQ. 'N') THEN
CALL HTD1(N, Z, D, E)
CALL QLS1(N, D, E, SHIFTMODE)
IF (SORTFLAG) CALL SORT1(N, D)
END IF
IF (JOBZ .EQ. 'V') THEN
CALL HTD2(N, Z, D, E)
CALL QLS2(N, D, E, Z, SHIFTMODE)
IF (SORTFLAG) CALL SORT2(N, D, Z)
END IF
RETURN
END SUBROUTINE HTDQLS
```

A.1.2. Tridiagonalization. The subroutine TD1 tridiagonalizes a symmetric matrix A and does not store the rotation matrices, while TD2 tridiagonalizes A and does store the rotation matrix. The diagonal elements are stored in D, while the first sub diagonal is stored in E. Of course the first super diagonal is the same as the first sub diagonal. The process is described in section 4.2

```
SUBROUTINE HTD1(N, A, D, E)
IMPLICIT NONE
INTEGER N
COMPLEX*32 A(N, N), D(N), E(0:N)
INTEGER I, J, K, L
COMPLEX*32 P, Q
LOGICAL REQTDS
DO I=N, 3, -1
D(I)=A(I, I)
P=0.0_16
REQTDS = .FALSE.
DO J=1, I-1
IF ((A(I,J) .NE. 0.0_16) .AND. (J .LE. I-2)) REQTDS = .TRUE.
P=P+A(I, J)**2
END DO
IF (REQTDS) THEN
Q=SQRT(P)
IF (ABS(A(I,I-1)+Q) .GE. ABS(A(I,I-1)-Q)) THEN
E(I)=-Q
P=P+A(I, I-1)*Q
A(I, I-1)=A(I, I-1)+Q
ELSE
E(I)=Q
P=P-A(I,I-1)*Q
A(I,I-1)=A(I,I-1)-Q
END IF
Q=0.0_16
DO J=1, I-1
E(J)=0.0_16
DO K=1, J
E(J)=E(J)+A(J, K)*A(I, K)
END DO
DO K=J+1, I-1
E(J)=E(J)+A(K, J)*A(I, K)
END DO
E(J)=E(J)/P
Q=Q+A(I, J)*E(J)
END DO
Q=Q/(2.0_16*P)
DO J=1, I-1
E(J)=E(J)-Q*A(I, J)
```

```
DO K=1, J
A(J, K)=A(J, K)-E(J)*A(I, K)-A(I, J)*E(K)
END DO
END DO
ELSE
E(I)=A(I,I-1)
END IF
END DO
D(2)=A(2, 2)
E(2)=A(2, 1)
D(1)=A(1, 1)
DO I=1, N-1
E(I)=E(I+1)
END DO
E(0)=0.0_16
E(N)=0.0_16
RETURN
END SUBROUTINE HTD1


SUBROUTINE HTD2(N, A, D, E)
IMPLICIT NONE
INTEGER N
COMPLEX*32 A(N, N), D(N), E(0:N)
INTEGER I, J, K, L
COMPLEX*32 P, Q
LOGICAL REQTDS
DO I=N, 3, -1
D(I)=A(I, I)
P=0.0_16
REQTDS = .FALSE.
DO J=1, I-1
IF ((A(I,J) .NE. 0.0_16) .AND. (J .LE. I-2)) REQTDS = .TRUE.
P=P+A(I, J)**2
END DO
IF (REQTDS) THEN
Q=SQRT(P)
IF (ABS(A(I,I-1)+Q) .GE. ABS(A(I,I-1)-Q)) THEN
E(I)=-Q
P=P+A(I, I-1)*Q
A(I, I-1)=A(I, I-1)+Q
ELSE
E(I)=Q
P=P-A(I,I-1)*Q
```

```
A(I,I-1)=A(I,I-1)-Q
END IF
Q=0.0_16
DO J=1, I-1
A(J, I)=A(I, J)/P
E(J)=0.0_16
DO K=1, J
E(J)=E(J)+A(J, K)*A(I, K)
END DO
DO K=J+1, I-1
E(J)=E(J)+A(K, J)*A(I, K)
END DO
E(J)=E(J)/P
Q=Q+A(I, J)*E(J)
END DO
Q=Q/(2.0_16*P)
DO J=1, I-1
E(J)=E(J)-Q*A(I, J)
DO K=1, J
A(J, K)=A(J, K)-E(J)*A(I, K)-A(I, J)*E(K)
END DO
END DO
ELSE
E(I)=A(I,I-1)
DO J=1, I-1
A(I,J) = 0.0_16
A(J,I) = 0.0_16
END DO
END IF
END DO
D(2)=A(2, 2)
E(2)=A(2, 1)
D(1)=A(1, 1)
DO I=1, N-1
E(I)=E(I+1)
END DO
E(0)=0.0_16
E(N)=0.0_16
A(1, 1)=1.0_16
A(2, 1)=0.0_16
DO I=2, N
DO J=1, I-1
P=0.0_16
DO K=1, I-1
```

```
      P=P+A(I, K)*A(K, J)
      END DO
      DO K=1, I-1
      A(K, J)=A(K, J)-A(K, I)*P
      END DO
      END DO
      A(I, I)=1.0_16
      DO J=1, I-1
      A(I, J)=0.0_16
      A(J, I)=0.0_16
      END DO
      END DO
      RETURN
      END SUBROUTINE HTD2
```

A.1.3. Diagonalization. The subroutine QLS1 diagonalizes the tridiagonal matrix stored in D and E, and does not store the rotation matrix while the subroutine QLS2 diagonalizes the tridiagonal matrix stored in D and E, and does store the rotation matrix. These routines detect premature zeroes, and perform the deflation steps when necessary. The procedure is described in chapter 4.3

```
      SUBROUTINE QLS1(N, D, E, SHIFTMODE)
      IMPLICIT NONE
      INTEGER N
      COMPLEX*32 D(N), E(0:N)
      INTEGER SHIFTMODE
      INTEGER I, J, K, L, M
      COMPLEX*32 C, P, Q, R, S, T, U
      DO I=1, N-1
 10   IF (D(I)+E(I) .NE. D(I)) THEN
      DO M=I+1, N-1
      IF (D(M)+E(M) .EQ. D(M)) GOTO 20
      END DO
 20   CALL SHIFT(N, I, M, D, E, S, SHIFTMODE)
      P=D(M)-S
      Q=E(M-1)
      T=D(M)
      U=E(M-1)
      DO K=M-1, I, -1
      R=SQRT(Q**2+P**2)
      E(K+1)=R
      C=P/R
```

```
      S=Q/R
      D(K+1)=C**2*T+2.0_16*C*S*U+S**2*D(K)
      P=(C**2-S**2)*U+C*S*(D(K)-T)
      T=C**2*D(K)-2.0_16*C*S*U+S**2*T
      U=C*E(K-1)
      Q=S*E(K-1)
      END DO
      D(I)=T
      E(I)=P
      E(M)=0.0_16
      GOTO 10
      END IF
      END DO
      RETURN
      END SUBROUTINE QLS1


      SUBROUTINE QLS2(N, D, E, Z, SHIFTMODE)
      IMPLICIT NONE
      INTEGER N
      COMPLEX*32 D(N), E(0:N), Z(N, N)
      INTEGER SHIFTMODE
      INTEGER I, J, K, L, M
      COMPLEX*32 C, P, Q, R, S, T, U
      DO I=1, N-1
10    IF (D(I)+E(I) .NE. D(I)) THEN
      DO M=I+1, N-1
      IF (D(M)+E(M) .EQ. D(M)) GOTO 20
      END DO
20    CALL SHIFT(N, I, M, D, E, S, SHIFTMODE)
      P=D(M)-S
      Q=E(M-1)
      T=D(M)
      U=E(M-1)
      DO K=M-1, I, -1
      R=SQRT(Q**2+P**2)
      E(K+1)=R
      C=P/R
      S=Q/R
      D(K+1)=C**2*T+2.0_16*C*S*U+S**2*D(K)
      P=(C**2-S**2)*U+C*S*(D(K)-T)
      T=C**2*D(K)-2.0_16*C*S*U+S**2*T
      U=C*E(K-1)
      Q=S*E(K-1)
```

```
DO L=1, N
R=Z(L, K+1)
Z(L, K+1)=S*Z(L, K)+C*R
Z(L, K)=C*Z(L, K)-S*R
END DO
END DO
D(I)=T
E(I)=P
E(M)=0.0_16
GOTO 10
END IF
END DO
RETURN
END SUBROUTINE QLS2
```

A.1.4. Shift. The subroutine SHIFT is directed by SHIFTMODE. The parameter SHIFTMODE can be 0, 1, 2 or 3, which will then direct the subroutine to implement no shift, a linear shift, a quadratic shift or a cubic shift respectively. For the quadratic and cubic shifts, the subroutine additionally chooses which of the two or three shifts should be used based on which is closest to the element which the routine is working to converge. The different shifts are discussed in chapter 4.3.1

```
SUBROUTINE SHIFT(N, K, V, D, E, S, SHIFTMODE)
IMPLICIT NONE
INTEGER N, K, V
COMPLEX*32 D(N), E(0:N), S
REAL*16, PARAMETER::C=2.0_16**(1.0_16/3.0_16)
COMPLEX*32, PARAMETER::II=(0.0_16,1.0_16)
COMPLEX*32 X, Y, P, Q, R
COMPLEX*32 S1, S2, S3
REAL*16 D1, D2, D3
INTEGER SHIFTMODE
IF ((SHIFTMODE .LT. 0) .OR. (SHIFTMODE .GT. 3)) THEN
PRINT*, 'INVALID SHIFTMODE: OPERATION TERMINATED'
STOP
END IF
IF (SHIFTMODE .EQ. 0) THEN
S = 0.0_16
ELSE IF (SHIFTMODE .EQ. 1) THEN
S = D(K)
ELSE IF ((V-K .GT. 3) .AND. (SHIFTMODE .EQ. 3)) THEN
```

```
P=2.0_16*D(K)-D(K+1)-D(K+2)
Q=D(K)-2.0_16*D(K+1)+D(K+2)
R=P*Q+9.0_16*E(K)**2
P=D(K)+D(K+1)-2.0_16*D(K+2)
Q=2.0_16*D(K)-D(K+1)
X=-P*R+9.0_16*Q*E(K+1)**2
P=D(K)+D(K+1)+D(K+2)
Q=P**2
P=D(K+1)*D(K+2)+D(K)*(D(K+1)+D(K+2))
R=P-E(K)**2-E(K+1)**2
Y=-Q+3.0_16*R
P=-9.0_16*D(K+2)*(E(K+1)**2)+X
Q=P**2+4.0_16*(Y**3)
R=SQRT(Q)
Q=P+R
P=EXP(LOG(Q)/3.0_16)
Q=Y/(3.0_16*P)
R=P/(3.0_16*C)
P=(D(K)+D(K+1)+D(K+2))/3.0_16
X=(1+II*SQRT(3.0_16))
Y=(1-II*SQRT(3.0_16))
S1=P+C*Q-R
S2=P-X*Q/(C**2)+Y*R/(2.0_16)
S3=P-Y*Q/(C**2)+X*R/(2.0_16)
S=S1
D1=ABS(D(K)-S1)
D2=ABS(D(K)-S2)
D3=ABS(D(K)-S3)
IF (D2 .LT. D1) THEN
S=S2
D1=D2
END IF
IF (D3 .LT. D1) THEN
S=S3
END IF
ELSE
P=(D(K+1)-D(K))/(2.0_16*E(K))
Q=SQRT(P**2+1.0_16)
X=-E(K)*(P-Q)
Y=-E(K)*(P+Q)
IF (ABS(X) .GT. ABS(Y)) THEN
S=D(K)+E(K)*(P+Q)
ELSE
S=D(K)+E(K)*(P-Q)
```

```
      END IF
      END IF
      RETURN
      END SUBROUTINE SHIFT
```

A.1.5. Sort. The subroutine SORT1 sorts only the eigenvalues stored in D to have ascending real parts while SORT2 sorts both the eigenvalues and eigenvectors, such that the eigenvalues have ascending real parts, and the eigenvectors are in the corresponding column of A.

```
      SUBROUTINE SORT1(N, D)
      IMPLICIT NONE
      INTEGER N
      COMPLEX*32 D(N)
      INTEGER I, J
      COMPLEX*32 P
      DO I=2, N
      J=I
 10   IF (DBLE(D(J)) .LT. DBLE(D(J-1))) THEN
      P=D(J-1)
      D(J-1)=D(J)
      D(J)=P
      IF (J .GT. 2) THEN
      J=J-1
      GOTO 10
      END IF
      END IF
      END DO
      RETURN
      END SUBROUTINE SORT1


      SUBROUTINE SORT2(N, D, A)
      IMPLICIT NONE
      INTEGER N
      COMPLEX*32 D(N), A(N, N)
      INTEGER I, J, K
      COMPLEX*32 P
      DO I=2, N
      J=I
 10   IF (DBLE(D(J)) .LT. DBLE(D(J-1))) THEN
      P=D(J-1)
```

```
D(J-1)=D(J)
D(J)=P
DO K=1, N
P=A(K, J-1)
A(K, J-1)=A(K, J)
A(K, J)=P
END DO
IF (J .GT. 2) THEN
J=J-1
GOTO 10
END IF
END IF
END DO
RETURN
END SUBROUTINE SORT2
```

A.2. IMPLEMENTATION OF HTDQRS

Here we present an explicit implementation of HTDQRS, the complementary algorithm to HTDQLS. As the two algorithms are very similar in how they are implemented, we refer you to chapter A.1 for details on how each subroutine functions.

```
SUBROUTINE HTDQRS(JOBZ, N, A, D, Z, SORTFLAG, SHIFTMODE)
IMPLICIT NONE
CHARACTER JOBZ
INTEGER N
COMPLEX*32 A(N, N), Z(N, N), D(N)
LOGICAL SORTFLAG
INTEGER SHIFTMODE, I
COMPLEX*32 E(0:N)
CALL COPYM(N, A, Z)
DO I = 0,N
E(I) = (0.0_16, 0.0_16)
END DO
IF (JOBZ .EQ. 'N') THEN
CALL HTD1(N, Z, D, E)
CALL QRS1(N, D, E, SHIFTMODE)
IF (SORTFLAG) CALL SORT1(N, D)
END IF
IF (JOBZ .EQ. 'V') THEN
CALL HTD2(N, Z, D, E)
CALL QRS2(N, D, E, Z, SHIFTMODE)
```

```
IF (SORTFLAG) CALL SORT2(N, D, Z)
END IF
RETURN
END SUBROUTINE HTDQRS


SUBROUTINE COPYM(N, A, Z)
IMPLICIT NONE
INTEGER N
complex*32 A(N, N), Z(N, N)
INTEGER I, J
DO I=1, N
DO J=1, N
Z(I,J)=A(I,J)
END DO
END DO
RETURN
END SUBROUTINE COPYM


SUBROUTINE HTD1(N, A, D, E)
IMPLICIT NONE
INTEGER N
COMPLEX*32 A(N, N), D(N), E(0:N)
INTEGER I, J, K, L
COMPLEX*32 P, Q
LOGICAL REQTDS
DO I=1, N-2
D(I)=A(I, I)
P=0.0_16
REQTDS = .FALSE.
DO J=I+1, N
IF ((A(I,J) .NE. 0.0_16) .AND. (J .GE. I+2)) REQTDS = .TRUE.
P=P+A(I, J)**2
END DO
IF (REQTDS) THEN
Q=SQRT(P)
IF (ABS(A(I,I-1)+Q) .GE. ABS(A(I,I-1)-Q)) THEN
E(I)=-Q
P=P+A(I, I+1)*Q
A(I, I+1)=A(I, I+1)+Q
ELSE
E(I)=Q
P=P-A(I,I+1)*Q
```

```
A(I,I+1)=A(I,I+1)-Q
END IF
Q=0.0_16
DO J=N, I+1,-1
E(J)=0.0_16
DO K=N, J+1, -1
E(J)=E(J)+A(J, K)*A(I, K)
END DO
DO K=J, I+1, -1
E(J)=E(J)+A(K, J)*A(I, K)
END DO
E(J)=E(J)/P
Q=Q+A(I, J)*E(J)
END DO
Q=Q/(2.0_16*P)
DO J=N, I+1, -1
E(J)=E(J)-Q*A(I, J)
DO K=N, J, -1
A(J, K)=A(J, K)-E(J)*A(I, K)-A(I, J)*E(K)
END DO
END DO
ELSE
E(I)=A(I,I+1)
END IF
END DO
D(N-1)=A(N-1, N-1)
E(N-1)=A(N-1, N)
D(N)=A(N, N)
E(0) = 0.0_16
E(N) = 0.0_16
RETURN
END SUBROUTINE HTD1


SUBROUTINE HTD2(N, A, D, E)
IMPLICIT NONE
INTEGER N
COMPLEX*32 A(N, N), D(N), E(0:N)
LOGICAL REQTDS
INTEGER I, J, K, L
COMPLEX*32 P, Q
DO I=1, N-2
D(I)=A(I, I)
P=0.0_16
```

```
REQTDS = .FALSE.
DO J=I+1, N
IF ((A(I,J) .NE. 0.0_16) .AND. (J .GE. I+2)) REQTDS = .TRUE.
P=P+A(I, J)**2
END DO
IF (REQTDS) THEN
Q=SQRT(P)
IF (ABS(A(I,I-1)+Q) .GE. ABS(A(I,I-1)-Q)) THEN
E(I)=-Q
P=P+A(I, I+1)*Q
A(I, I+1)=A(I, I+1)+Q
ELSE
E(I)=Q
P=P-A(I,I+1)*Q
A(I,I+1)=A(I,I+1)-Q
END IF
Q=0.0_16
DO J=N, I+1,-1
A(J, I)=A(I, J)/P
E(J)=0.0_16
DO K=N, J+1, -1
E(J)=E(J)+A(J, K)*A(I, K)
END DO
DO K=J, I+1, -1
E(J)=E(J)+A(K, J)*A(I, K)
END DO
E(J)=E(J)/P
Q=Q+A(I, J)*E(J)
END DO
Q=Q/(2.0_16*P)
DO J=N, I+1, -1
E(J)=E(J)-Q*A(I, J)
DO K=N, J, -1
A(J, K)=A(J, K)-E(J)*A(I, K)-A(I, J)*E(K)
END DO
END DO
ELSE
E(I)=A(I,I+1)
DO J=I+1, N
A(I, J) = 0.0_16
A(J, I) = 0.0_16
END DO
END IF
END DO
```

```
      D(N-1)=A(N-1, N-1)
      E(N-1)=A(N-1, N)
      D(N)=A(N, N)
      A(N, N)=1.0_16
      A(N-1, N)=0.0_16
      E(0) = 0.0_16
      E(N) = 0.0_16
      DO I=N-1, 1, -1
      DO J=N, I+1, -1
      P=0.0_16
      DO K=N, I+1, -1
      P=P+A(I, K)*A(K, J)
      END DO
      DO K=N, I+1, -1
      A(K, J)=A(K, J)-A(K, I)*P
      END DO
      END DO
      A(I, I)=1.0_16
      DO J=N, I+1, -1
      A(I, J)=0.0_16
      A(J, I)=0.0_16
      END DO
      END DO
      RETURN
      END SUBROUTINE HTD2


      SUBROUTINE QRS1(N, D, E, SHIFTMODE)
      IMPLICIT NONE
      INTEGER N
      COMPLEX*32 D(N), E(0:N)
      INTEGER SHIFTMODE
      INTEGER I, J, K, L, M
      COMPLEX*32 C, P, Q, R, S, T, U
      DO I=N-1, 1, -1
10    IF (D(I+1)+E(I) .NE. D(I+1)) THEN
      DO M=I-1, 1, -1
      IF (D(M+1)+E(M) .EQ. D(M+1)) GOTO 20
      END DO
20    CALL SHIFT(N, I, M, D, E, S, SHIFTMODE)
      P=D(M+1)-S
      Q=E(M+1)
      T=D(M+1)
      U=E(M+1)
```

```
      DO K=M+1, I
      R=SQRT(Q**2+P**2)
      E(K-1)=R
      C=P/R
      S=-Q/R
      D(K)=C**2*T-2.0_16*C*S*U+S**2*D(K+1)
      P=(C**2-S**2)*U+C*S*(T-D(K+1))
      T=C**2*D(K+1)+2.0_16*C*S*U+S**2*T
      U=C*E(K+1)
      Q=-S*E(K+1)
      END DO
      D(I+1)=T
      E(I)=P
      E(M)=0.0_16
      GOTO 10
      END IF
      END DO
      RETURN
      END SUBROUTINE QRS1


      SUBROUTINE QRS2(N, D, E, Z, SHIFTMODE)
      IMPLICIT NONE
      INTEGER N
      COMPLEX*32 D(N), E(0:N), Z(N, N)
      INTEGER SHIFTMODE
      INTEGER I, J, K, L, M
      COMPLEX*32 C, P, Q, R, S, T, U
      DO I=N-1, 1, -1
10    IF (D(I+1)+E(I) .NE. D(I+1)) THEN
      DO M=I-1, 1, -1
      IF (D(M+1)+E(M) .EQ. D(M+1)) GOTO 20
      END DO
20    CALL SHIFT(N, I, M, D, E, S, SHIFTMODE)
      P=D(M+1)-S
      Q=E(M+1)
      T=D(M+1)
      U=E(M+1)
      DO K=M+1, I
      R=SQRT(Q**2+P**2)
      E(K-1)=R
      C=P/R
      S=-Q/R
      D(K)=C**2*T-2.0_16*C*S*U+S**2*D(K+1)
```

```
P=(C**2-S**2)*U+C*S*(T-D(K+1))
T=C**2*D(K+1)+2.0_16*C*S*U+S**2*T
U=C*E(K+1)
Q=-S*E(K+1)
DO L=1, N
R=Z(L, K+1)
Z(L, K+1)=S*Z(L, K)+C*R
Z(L, K)=C*Z(L, K)-S*R
END DO
END DO
D(I+1)=T
E(I)=P
E(M)=0.0_16
GOTO 10
END IF
END DO
RETURN
END SUBROUTINE QRS2


SUBROUTINE SHIFT(N, K, M, D, E, S, SHIFTMODE)
IMPLICIT NONE
INTEGER N, K, M
COMPLEX*32 D(N), E(0:N), S
REAL*16, PARAMETER::C=2.0_16**(1.0_16/3.0_16)
COMPLEX*32, PARAMETER::II=(0.0_16,1.0_16)
COMPLEX*32 X, Y, P, Q, R
COMPLEX*32 S1, S2, S3
REAL*16 D1, D2, D3
INTEGER SHIFTMODE
IF ((SHIFTMODE .LT. 0) .OR. (SHIFTMODE .GT. 3)) THEN
PRINT*, 'INVALID SHIFTMODE: OPERATION TERMINATED'
STOP
END IF
IF (SHIFTMODE .EQ. 0) THEN
S = 0.0_16
ELSE IF (SHIFTMODE .EQ. 1) THEN
S = D(K+1)
ELSE IF ((K-M .GT. 3) .AND. (SHIFTMODE .EQ. 3)) THEN
P=2.0_16*D(K-1)-D(K)-D(K+1)
Q=D(K-1)-2.0_16*D(K)+D(K+1)
R=P*Q+9.0_16*E(K-1)**2
P=D(K-1)+D(K)-2.0_16*D(K+1)
Q=2.0_16*D(K-1)-D(K)
```

```
X=-P*R+9.0_16*Q*E(K)**2
P=D(K-1)+D(K)+D(K+1)
Q=P**2
P=D(K)*D(K+1)+D(K-1)*(D(K)+D(K+1))
R=P-E(K-1)**2-E(K)**2
Y=-Q+3.0_16*R
P=-9.0_16*D(K+1)*(E(K)**2)+X
Q=P**2+4.0_16*(Y**3)
R=SQRT(Q)
Q=P+R
P=EXP(LOG(Q)/3.0_16)
Q=Y/(3.0_16*P)
R=P/(3.0_16*C)
P=(D(K-1)+D(K)+D(K+1))/3.0_16
X=(1+II*SQRT(3.0_16))
Y=(1-II*SQRT(3.0_16))
S1=P+C*Q-R
S2=P-X*Q/(C**2)+Y*R/(2.0_16)
S3=P-Y*Q/(C**2)+X*R/(2.0_16)
S=S1
D1=ABS(D(K+1)-S1)
D2=ABS(D(K+1)-S2)
D3=ABS(D(K+1)-S3)
IF (D2 .LT. D1) THEN
S=S2
D1=D2
END IF
IF (D3 .LT. D1) THEN
S=S3
END IF
ELSE
P=(D(K)-D(K+1))/(2.0_16*E(K))
Q=SQRT(P**2+1.0_16)
D1=ABS(-E(K)*(P-Q))
D2=ABS(-E(K)*(P+Q))
IF (D1 .LT.D2) THEN
S=D(K+1)+E(K)*(P-Q)
ELSE
S=D(K+1)+E(K)*(P+Q)
END IF
END IF
RETURN
END SUBROUTINE SHIFT
```

```
      SUBROUTINE SORT1(N, D)
      IMPLICIT NONE
      INTEGER N
      COMPLEX*32 D(N)
      INTEGER I, J
      COMPLEX*32 P
      DO I=2, N
      J=I
10    IF (DBLE(D(J)) .LT. DBLE(D(J-1))) THEN
      P=D(J-1)
      D(J-1)=D(J)
      D(J)=P
      IF (J .GT. 2) THEN
      J=J-1
      GOTO 10
      END IF
      END IF
      END DO
      RETURN
      END SUBROUTINE SORT1


      SUBROUTINE SORT2(N, D, A)
      IMPLICIT NONE
      INTEGER N
      COMPLEX*32 D(N), A(N, N)
      INTEGER I, J, K
      COMPLEX*32 P
      DO I=2, N
      J=I
10    IF (DBLE(D(J)) .LT. DBLE(D(J-1))) THEN
      P=D(J-1)
      D(J-1)=D(J)
      D(J)=P
      DO K=1, N
      P=A(K, J-1)
      A(K, J-1)=A(K, J)
      A(K, J)=P
      END DO
      IF (J .GT. 2) THEN
      J=J-1
      GOTO 10
      END IF
```

```
END IF
END DO
RETURN
END SUBROUTINE SORT2
```

APPENDIX B

PLAIN QL AND QR ALGORITHMS

B.1. OVERVIEW

As discussed in chapter 4.2.4, it is possible to perform QL and QR decompositions, and subsequently matrix diagonalization using solely Householder reflections. Without loss of generality, here we discuss the plain QL implementation (`PQL`), as the plain QR algorithm works in the much the same way, save for the fact that it performs QR decompositions in place of the QL decompositions.

Unlike the `HTDQLS` and `HTDQRS` routines, the `PQL` and `PQR` algorithms are not presented as a master subroutine with versions of subroutines that carry out the calculations. Instead we present four algorithms which are independent of each other. Two of which are implementations of the `PQL` algorithm, while the other two are implementations of the `PQR` algorithm. The difference between the two `PQL` and the two `PQR` algorithms is whether they are calculating the only the eigenvalues (denoted by a '`1`' following the title) or calculating both the eigenvalues and eigenvectors (denoted by a '`2`' at the end of the title). In actuality the difference is contained in about 24 lines of code.

As in the case of the `HTDQLS` algorithm, we do not need to explicitly calculate the $\underline{Q}$ matrices, and as we shall see we actually only ever explicitly calculate the first one. We do however need to calculate explicit values for at least a portion of the $\underline{L}_i$ matrix (here $\underline{L}_i$ is the in terms of what was presented in chapter 4.3.1). In the case of the `PQL` algorithm, we will have three arrays to keep track of (2 if just the eigenvalues are to be found, and the following algorithm is modified as prescribed).

B.2. PLAIN QL ALGORITHM

```
SUBROUTINE PQLX1(N, A, D)
IMPLICIT NONE
INTEGER N
COMPLEX*32 A(N,N), D(N)
COMPLEX*32 L(N,N), V(N), W(N)
```

```
COMPLEX*32 P, Q
INTEGER I, J, K, M
M=N
DO
Q=0.0_16
DO I=1, M-1
Q=Q+A(I,M)
END DO
IF(A(M,M)+Q .EQ. A(M,M)) M=M-1
IF(M .EQ. 1) EXIT
DO I=1, M
DO J=1, M
L(I,J)=A(I,J)
END DO
END DO
DO K=M, 2, -1
Q=0.0_16
DO I=1, K
V(I)=L(I,K)
Q=Q+V(I)*V(I)
END DO
P=SQRT(Q)
IF (P*L(K,K) .EQ. -Q) P = -P
V(K)=V(K)+P
P=P*L(K,K)+Q
Q=0.0_16
DO I=1, K
W(I)=0.0_16
DO J=1, K
W(I)=W(I)+A(I,J)*V(J)/P
END DO
Q=Q+V(I)*W(I)/(2*P)
END DO
DO I=1, K
W(I)=W(I)-Q*V(I)
END DO
DO I=1, K
DO J=1, K
A(I,J)=A(I,J)-V(I)*W(J)-W(I)*V(J)
END DO
END DO
DO I=K+1, N
W(I)=0.0_16
DO J=1, K
```

```
W(I)=W(I)+V(J)*A(J,I)/P
END DO
DO J=1, K
A(J,I)=A(J,I)-V(J)*W(I)
A(I,J)=A(J,I)
END DO
END DO
DO I=1, k-1
W(I)=0.0_16
DO J=1, K
W(I)=W(I)+V(J)*L(J,I)/P
END DO
DO J=1, k-1
L(J,I)=L(J,I)-V(J)*W(I)
END DO
END DO
END DO
END DO
DO I=1, N
D(I)=A(I,I)
END DO
RETURN
END SUBROUTINE PQLX1


SUBROUTINE PQLX2(N, A, Z, D)
IMPLICIT NONE
INTEGER N
COMPLEX*32 A(N,N), Z(N,N), D(N)
COMPLEX*32 L(N,N), V(N), W(N)
COMPLEX*32 P, Q
INTEGER I, J, K, M
DO I=1, N
DO J=1, N
IF(I .EQ. J) Z(I,J)=1.0_16
IF(I .NE. J) Z(I,J)=0.0_16
END DO
END DO
M=N
DO
Q=0.0_16
DO I=1, M-1
Q=Q+A(I,M)
END DO
```

```
IF(A(M,M)+Q .EQ. A(M,M)) M=M-1
IF(M .EQ. 1) EXIT
DO I=1, M
DO J=1, M
L(I,J)=A(I,J)
END DO
END DO
DO K=M, 2, -1
Q=0.0_16
DO I=1, K
V(I)=L(I,K)
Q=Q+V(I)*V(I)
END DO
P=SQRT(Q)
IF (P*L(K,K) .EQ. -Q) P = -P
V(K)=V(K)+P
P=P*L(K,K)+Q
Q=0.0_16
DO I=1, K
W(I)=0.0_16
DO J=1, K
W(I)=W(I)+A(I,J)*V(J)/P
END DO
Q=Q+V(I)*W(I)/(2*P)
END DO
DO I=1, K
W(I)=W(I)-Q*V(I)
END DO
DO I=1, K
DO J=1, K
A(I,J)=A(I,J)-V(I)*W(J)-W(I)*V(J)
END DO
END DO
DO I=K+1, N
W(I)=0.0_16
DO J=1, K
W(I)=W(I)+V(J)*A(J,I)/P
END DO
DO J=1, K
A(J,I)=A(J,I)-V(J)*W(I)
A(I,J)=A(J,I)
END DO
END DO
DO I=1, k-1
```

```
W(I)=0.0_16
DO J=1, K
W(I)=W(I)+V(J)*L(J,I)/P
END DO
DO J=1, k-1
L(J,I)=L(J,I)-V(J)*W(I)
END DO
END DO
DO I=1, N
W(I)=0.0_16
DO J=1, K
W(I)=W(I)+Z(I,J)*V(J)/P
END DO
DO J=1, K
Z(I,J)=Z(I,J)-W(I)*V(J)
END DO
END DO
END DO
END DO
DO I=1, N
D(I)=A(I,I)
END DO
RETURN
END SUBROUTINE PQLX2
```

B.3. PLAIN QR ALGORITHM

```
SUBROUTINE PQRX1(N, A, D)
IMPLICIT NONE
INTEGER N
COMPLEX*32 A(N,N), D(N)
COMPLEX*32 R(N,N), V(N), W(N)
COMPLEX*32 P, Q
INTEGER I, J, K, M
M=1
DO
Q=0.0_16
DO I=M+1, N
Q=Q+A(I,M)
END DO
IF(A(M,M)+Q .EQ. A(M,M)) M=M+1
IF(M .EQ. N-1) EXIT
DO I=M, N
```

```
DO J=M, N
R(I,J)=A(I,J)
END DO
END DO
DO K=M, N-1
Q=0.0_16
DO I=K, N
V(I)=R(I,K)
Q=Q+V(I)*V(I)
END DO
P=SQRT(Q)
IF (P*R(K,K) .EQ. -Q) P = -P
V(K)=V(K)+P
P=P*R(K,K)+Q
Q=0.0_16
DO I=K, N
W(I)=0.0_16
DO J=K, N
W(I)=W(I)+A(I,J)*V(J)/P
END DO
Q=Q+V(I)*W(I)/(2*P)
END DO
DO I=K, N
W(I)=W(I)-Q*V(I)
END DO
DO I=K, N
DO J=K, N
A(I,J)=A(I,J)-V(I)*W(J)-W(I)*V(J)
END DO
END DO
DO I=1, K-1
W(I)=0.0_16
DO J=K, N
W(I)=W(I)+V(J)*A(I,J)/P
END DO
DO J=K, N
A(I,J)=A(I,J)-V(J)*W(I)
A(J,I)=A(I,J)
END DO
END DO
DO I=K+1, N
W(I)=0.0_16
DO J=K, N
W(I)=W(I)+V(J)*R(J,I)/P
```

```
END DO
DO J=K+1, N
R(J,I)=R(J,I)-V(J)*W(I)
END DO
END DO
END DO
END DO
DO I=1, N
D(I)=A(I,I)
END DO
RETURN
END SUBROUTINE PQRX1


SUBROUTINE PQRX2(N, A, Z, D)
IMPLICIT NONE
INTEGER N
COMPLEX*32 A(N,N), Z(N,N), D(N)
COMPLEX*32 R(N,N), V(N), W(N)
COMPLEX*32 P, Q
INTEGER I, J, K, M
DO I=1, N
DO J=1, N
IF(I .EQ. J) Z(I,J)=1.0_16
IF(I .NE. J) Z(I,J)=0.0_16
END DO
END DO
M=1
DO
Q=0.0_16
DO I=M+1, N
Q=Q+A(I,M)
END DO
IF(A(M,M)+Q .EQ. A(M,M)) M=M+1
IF(M .EQ. N-1) EXIT
DO I=M, N
DO J=M, N
R(I,J)=A(I,J)
END DO
END DO
DO K=M, N-1
Q=0.0_16
DO I=K, N
V(I)=R(I,K)
```

```
Q=Q+V(I)*V(I)
END DO
P=SQRT(Q)
IF (P*R(K,K) .EQ. -Q) P = -P
V(K)=V(K)+P
P=P*R(K,K)+Q
Q=0.0_16
DO I=K, N
W(I)=0.0_16
DO J=K, N
W(I)=W(I)+A(I,J)*V(J)/P
END DO
Q=Q+V(I)*W(I)/(2*P)
END DO
DO I=K, N
W(I)=W(I)-Q*V(I)
END DO
DO I=K, N
DO J=K, N
A(I,J)=A(I,J)-V(I)*W(J)-W(I)*V(J)
END DO
END DO
DO I=1, K-1
W(I)=0.0_16
DO J=K, N
W(I)=W(I)+V(J)*A(I,J)/P
END DO
DO J=K, N
A(I,J)=A(I,J)-V(J)*W(I)
A(J,I)=A(I,J)
END DO
END DO
DO I=K+1, N
W(I)=0.0_16
DO J=K, N
W(I)=W(I)+V(J)*R(J,I)/P
END DO
DO J=K+1, N
R(J,I)=R(J,I)-V(J)*W(I)
END DO
END DO
DO I=1, N
W(I)=0.0_16
DO J=K, N
```

```
W(I)=W(I)+Z(I,J)*V(J)/P
END DO
DO J=K, N
Z(I,J)=Z(I,J)-W(I)*V(J)
END DO
END DO
END DO
END DO
DO I=1, N
D(I)=A(I,I)
END DO
RETURN
END SUBROUTINE PQRX2
```

APPENDIX C

RELATIVISTIC QUANTUM MECHANICS: THE BASICS

C.1. SCHRÖDINGER EQUATION

We obtain the Schrödinger equation by identifying that when moving from (nonrelativistic) classical mechanics to (again nonrelativistic) quantum mechanics

$$E \to \mathrm{i}\partial_t \,, \quad \vec{p} \to -\mathrm{i}\vec{\nabla} \,. \tag{C.1}$$

When applied to the classical equation for a free particle,

$$E = \frac{\vec{p}^{\,2}}{2m} \,, \tag{C.2}$$

we obtain the free Schrödinger equation (describing a free quantum particle)

$$\mathrm{i}\partial_t \, \phi(t, \vec{r}) = -\frac{\vec{\nabla}^{\,2}}{2m} \phi(t, \vec{r}) \,, \tag{C.3}$$

which we can rewrite as

$$\left(\mathrm{i}\partial_t + \frac{1}{2m} \vec{\nabla}^{\,2} \right) \phi(t, \vec{r}) = 0 \,. \tag{C.4}$$

Taking the complex conjugate we find

$$\left(-\mathrm{i}\partial_t + \frac{1}{2m} \vec{\nabla}^{\,2} \right) \phi^*(t, \vec{r}) = 0 \,. \tag{C.5}$$

We now multiply (on the left) (C.4) and (C.5) by $\phi^*(t, \vec{r})$ and $\phi(t, \vec{r})$ respectively, yielding

$$\phi^*(t, \vec{r}) \left(\mathrm{i}\partial_t + \frac{1}{2m} \vec{\nabla}^{\,2} \right) \phi(t, \vec{r}) = \phi(t, \vec{r}) \left(-\mathrm{i}\partial_t + \frac{1}{2m} \vec{\nabla}^{\,2} \right) \phi^*(t, \vec{r}) = 0 \,. \tag{C.6}$$

By moving all the terms to the l.h.s. we find

$$\phi^*(t,\vec{r})\,\mathrm{i}\partial_t\,\phi(t,\vec{r})\ +\phi(t,\vec{r})\,\mathrm{i}\partial_t\,\phi^*(t,\vec{r}) \\ +\frac{1}{2m}\left(\phi^*(t,\vec{r})\,\vec{\nabla}^{\,2}\phi(t,\vec{r})-\phi(t,\vec{r})\,\vec{\nabla}^{\,2}\phi^*(t,\vec{r})\right)=0\,, \tag{C.7}$$

we now multiply both sides by $-\mathrm{i}$ and add $(\mathrm{i}/(2m)[(\vec{\nabla}\phi^*)(\vec{\nabla}\phi)-(\vec{\nabla}\phi)(\vec{\nabla}\phi^*)]=0$ to the equation, i.e.

$$\partial_t\,|\phi(t,\vec{r})|^2-\frac{\mathrm{i}}{2m}\left(\phi^*(t,\vec{r})\vec{\nabla}^{\,2}\phi(t,\vec{r})+\left(\vec{\nabla}\phi^*(t,\vec{r})\right)\left(\vec{\nabla}\phi(t,\vec{r})\right)\right. \\ \left.-\phi(t,\vec{r})\vec{\nabla}^{\,2}\phi^*(t,\vec{r})-\left(\vec{\nabla}\phi(t,\vec{r})\right)\left(\vec{\nabla}\phi^*(t,\vec{r})\right)\right)=0\,, \tag{C.8}$$

which simplifies to

$$\partial_t\,|\phi(t,\vec{r})|^2-\frac{\mathrm{i}}{2m}\vec{\nabla}\left(\phi^*(t,\vec{r})\vec{\nabla}\phi(t,\vec{r})-\phi(t,\vec{r})\vec{\nabla}\phi^*(t,\vec{r})\right) \\ =\partial_t\,|\phi(t,\vec{r})|^2+\vec{\nabla}\cdot\left(-\frac{\mathrm{i}}{2m}\phi^*(t,\vec{r})\,\overleftrightarrow{\vec{\nabla}}\,\phi(t,\vec{r})\right)=0\,, \tag{C.9}$$

where $\overleftrightarrow{\vec{\nabla}}$ is the antisymmetric differential operator, and acts as

$$f(\vec{r})\,\overleftrightarrow{\vec{\nabla}}\,g(\vec{r})=f(\vec{r})\vec{\nabla}g(\vec{r})-g(\vec{r})\vec{\nabla}f(\vec{r})\,. \tag{C.10}$$

By defining $\rho(t,\vec{r})$ and $\vec{j}(t,\vec{r})$ as

$$\rho(t,\vec{r})=|\phi(t,\vec{r})|^2\,,\quad \vec{j}(t,\vec{r})=-\frac{\mathrm{i}}{2m}\phi^*(t,\vec{r})\,\overleftrightarrow{\vec{\nabla}}\,\phi(t,\vec{r})\,, \tag{C.11}$$

(C.9) becomes the continuity equation

$$\partial_t\rho(t,\vec{r})+\vec{\nabla}\cdot\vec{j}(t,\vec{r})=0\,, \tag{C.12}$$

where $\rho(t,\vec{r})$ is positive definite, and is interpreted as the probability density.

C.2. KLEIN–GORDON EQUATION

We would like to generalize this to relativistic quantum mechanics. To do so we we use the relativistic dispersion relation $E^2 = \vec{p}^{\,2} + m^2$ to find

$$\partial_t^2 \phi(t, \vec{r}) = \left(\vec{\nabla}^{\,2} - m^2 \right) \phi(t, \vec{r}) \,, \tag{C.13}$$

i.e.

$$\left(\partial_t^2 - \vec{\nabla}^{\,2} + m^2 \right) \phi(t, \vec{r}) = 0 \,. \tag{C.14}$$

Taking the complex conjugate we find

$$\left(\partial_t^2 - \vec{\nabla}^{\,2} + m^2 \right) \phi^*(t, \vec{r}) = 0 \,. \tag{C.15}$$

We now multiply the l.h.s. of (C.14) and (C.15) by $\phi^*(t, \vec{r})$ and $\phi(t, \vec{r})$ respectively, to obtain

$$\begin{aligned}
&\phi^*(t, \vec{r}) \left(\partial_t^2 - \vec{\nabla}^{\,2} + m^2 \right) \phi(t, \vec{r}) = \phi(t, \vec{r}) \left(\partial_t^2 - \vec{\nabla}^{\,2} + m^2 \right) \phi^*(t, \vec{r}) = 0 \\
\Rightarrow\, &\phi^*(t, \vec{r}) \partial_t^2 \phi(t, \vec{r}) - \phi(t, \vec{r}) \partial_t^2 \phi^*(t, \vec{r}) - \phi^*(t, \vec{r}) \vec{\nabla} \phi(t, \vec{r}) + \phi(t, \vec{r}) \vec{\nabla} \phi^*(t, \vec{r}) = 0 \\
\Rightarrow\, &\partial_t \left(\phi^*(t, \vec{r}) \overleftrightarrow{\partial_t} \phi(t, \vec{r}) \right) + \vec{\nabla} \cdot \left(-\phi^*(t, \vec{r}) \overleftrightarrow{\vec{\nabla}} \phi(t, \vec{r}) \right) = 0 \,.
\end{aligned} \tag{C.16}$$

For the last step we used the fact that

$$f(x) \partial_x^2 g(x) - g(x) \partial_x^2 f(x) = \partial_x \left(g(x) \partial_x f(x) - f(x) \partial_x g(x) \right) = \partial_x \left(g(x) \overleftrightarrow{\partial_x} f(x) \right) \,. \tag{C.17}$$

We then have to redefine what ρ is for the KG equation, but we can keep the same definition for $\vec{j}$, with this in mind we define

$$\rho(t,\vec{r}) = \frac{\mathrm{i}}{2m}\phi^*(t,\vec{r})\overleftrightarrow{\partial_t}\phi(t,\vec{r})\,, \tag{C.18}$$

$$\vec{j}(t,\vec{r}) = -\frac{\mathrm{i}}{2m}\phi^*(t,\vec{r})\overleftrightarrow{\vec{\nabla}}\phi(t,\vec{r})\,. \tag{C.19}$$

Then (C.16) tells us

$$\partial_t\rho(t,\vec{r}) + \vec{\nabla}\cdot\vec{j}(t,\vec{r}) = 0\,. \tag{C.20}$$

It is then natural to again interpret ρ as the probability density. However there are two solutions to the KG equation, a positive and a negative energy solution, i.e.

$$\phi_+(t,\vec{r}) = N\,\exp\left[-\mathrm{i}\left(E\,t - \vec{k}\cdot\vec{r}\right)\right]\,, \tag{C.21}$$

$$\phi_-(t,\vec{r}) = N\,\exp\left[\mathrm{i}\left(E\,t - \vec{k}\cdot\vec{r}\right)\right]\,, \tag{C.22}$$

where $E = \sqrt{\vec{k}^{\,2} + m^2} \geq 0$. If we choose $\phi_-(t,\vec{r})$, then

$$\begin{aligned}\rho(t,\vec{r}) &= \frac{\mathrm{i}}{2m}\left(\phi_-^*(t,\vec{r})\partial_t\phi_-(t,\vec{r}) - \phi_-(t,\vec{r})\partial_t\phi_-^*(t,\vec{r})\right)\\ &= \frac{\mathrm{i}}{2m}N^2\left(\mathrm{i}E - (-\mathrm{i}E)\right) = -N^2\frac{E}{m} \leq 0\,.\end{aligned} \tag{C.23}$$

Thus for the Klein–Gordon (KG) equation, ρ is *not* positive definite, and *cannot* be interpreted as the probability density. Instead it should be interpreted as a charge density. However, this difference initially was extremely worrisome to physicists, and lead to the KG equation being rejected by physicists, and the discovery of the Dirac equation.

C.3. DIRAC EQUATION

The Dirac equation was born out of a desire to combine the advantages of the Schrödinger equation (positive–definite probability density) and the KG equation (relativistic invariance). Such an equation would have to be linear in terms of the time derivative, and consequently momentum (i.e., the spatial derivatives). Additionally, the square of such an equation would necessarily recover the KG equation. Let us rewrite (C.13) as

$$-\partial_t^2 \phi(t,\vec{r}) = \left(-\vec{\nabla}^{\,2} + m^2\right) \phi(t,\vec{r})\,, \tag{C.24}$$

which is simply another way of expressing the KG equation. We now assume that the r.h.s. of (C.24) can be factored in such a way that (suppressing the wave function)

$$\begin{aligned}
\left(-\vec{\nabla}^{\,2} + m^2\right) &= \left(-\mathrm{i}\vec{\alpha}\cdot\vec{\nabla} + \beta m\right)^2 = \left(-\mathrm{i}\alpha^i\partial_i + \beta m\right)\left(-\mathrm{i}\alpha^j\partial_j + \beta m\right) \\
&= -\,\alpha^i\alpha^j\,\partial_i\partial_j - \mathrm{i}\left(\alpha^i\beta\partial_i + \beta\alpha^j\partial_j\right) m + \beta^2 m^2 \\
&= -\,\frac{1}{2}\left(\alpha^i\alpha^j + \alpha^j\alpha^i\right)\partial_i\partial_j - \mathrm{i}\left(\alpha^i\beta + \beta\alpha^i\right) m\,\partial_i + \beta^2 m^2 \\
&= -\,\frac{1}{2}\left\{\alpha^i,\alpha^j\right\}\partial_i\partial_j - \mathrm{i}\left\{\alpha^i,\beta\right\} m\,\partial_i + \beta^2 m^2\,.
\end{aligned} \tag{C.25}$$

Comparing the beginning and end of the equation, we can make a number of deductions (bear in mind that $\vec{\nabla}^{\,2} = \partial_i\partial_i$), namely

$$\left\{\alpha^i,\alpha^j\right\} = 2\delta^{ij}\,, \quad \left\{\alpha^i,\beta\right\} = 0\,, \quad \beta^2 =, 1\,. \tag{C.26}$$

We know that scalars commute, thus these conditions cannot be satisfied if α^i and β are scalars. Instead we use matrices. First let us define the Dirac γ matrices

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} , \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} , \tag{C.27}$$

where σ^i are the Pauli spin matrices, and are defined as

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma^2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} , \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{C.28}$$

The Dirac γ matrices have the property

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} , \tag{C.29}$$

where in flat space, the metric is

$$[g^{\mu\nu}] = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} . \tag{C.30}$$

Thus

$$\left(\gamma^0\right)^2 = 1 , \quad \left(\gamma^i\right)^2 = -1 , \quad \{\gamma^\mu, \gamma^\nu\} = 0 \quad (\mu \neq \nu) . \tag{C.31}$$

We find that the conditions set fourth by (C.26) are satisfied if we define

$$\alpha^i = \gamma^0\gamma^i , \quad \beta = \gamma^0 . \tag{C.32}$$

Thus (C.24) becomes

$$(\mathrm{i}\partial_t)^2\,\phi(t,\vec{r}) = \left(-\mathrm{i}\vec{\alpha}\cdot\vec{\nabla} + \beta m\right)^2 \phi(t,\vec{r})\,, \tag{C.33}$$

which is equivalent to

$$\mathrm{i}\partial_t\phi(t,\vec{r}) = \left(-\mathrm{i}\vec{\alpha}\cdot\vec{\nabla} + \beta m\right)\phi(t,\vec{r}) = \left(\vec{\alpha}\cdot\vec{p} + \beta m\right)\phi(t,\vec{r}) = H\phi(t,\vec{r})\,, \tag{C.34}$$

where $\vec{p} = -\mathrm{i}\vec{\nabla}$ is the momentum operator. By multiplying this equation with γ^0 on the left, and collecting all the terms together we find an alternative way of writing the Dirac equation, i.e.

$$(\mathrm{i}\gamma^\mu\partial_\mu - m)\,\phi(x) = 0\,, \quad x = (t,\vec{r})\,. \tag{C.35}$$

We now take the adjoint (transpose and complex conjugate) of (C.35), multiply by γ^0 on the right, and use the fact that $(\gamma^0)^2 = 0$ to find

$$\begin{aligned}\phi^+(x)\gamma^0\gamma^0\left(-\mathrm{i}\,(\gamma^\mu)^+\,\overleftarrow{\partial}_\mu - m\right)\gamma^0 &= \overline{\phi}(x)\left(-\mathrm{i}\gamma^0\,(\gamma^\mu)^+\,\gamma^0\overleftarrow{\partial}_\mu - \left(\gamma^0\right)^2 m\right)\\ &= \overline{\phi}(x)\left(-\mathrm{i}\gamma^\mu\overleftarrow{\partial}_\mu - m\right) = 0\,,\end{aligned} \tag{C.36}$$

where we used the identity $\gamma^0(\gamma^\mu)^+\gamma^0 = \gamma^\mu$, and defined $\overline{\phi}(x) = \phi^+(x)\gamma^0$. We now multiply (C.35) by $\overline{\phi}(x)$ on the left and (C.36) by $\phi(x)$ on the right, and equate the two to find

$$\overline{\phi}(x)\,(\mathrm{i}\gamma^\mu\partial_\mu - m)\,\phi(x) = \overline{\phi}(x)\left(-\mathrm{i}\gamma^\mu\overleftarrow{\partial}_\mu - m\right)\phi(x)\,, \tag{C.37}$$

which simplifies to (when multiplied by −i)

$$\overline{\phi}(x)\gamma^{\mu}\partial_{\mu}\phi(x) + \partial_{\mu}\overline{\phi}(x)\gamma^{\mu}\phi(x) = \partial_{\mu}\left(\overline{\phi}(x)\gamma^{\mu}\phi(x)\right) = 0\,. \tag{C.38}$$

We now define the probability current as

$$j^{\mu} = \overline{\phi}(x)\gamma^{\mu}\phi(x)\,, \tag{C.39}$$

and by (C.38) we have

$$\partial_{\mu}j^{\mu} = 0\,. \tag{C.40}$$

Recall that the issue many physicists had with the KG equation was that $\rho = j^0$ was not positive definite. Well, here we have

$$\rho = j^0 = \overline{\phi}(x)\gamma^0\phi(x) = \phi^{+}(x)\gamma^0\gamma^0\phi(x) = |\phi(x)|^2\,, \tag{C.41}$$

which is positive definite.

C.4. LORENTZ INVARIANCE IN FLAT SPACE

Here we want to show that the Dirac equation is invariant under Lorentz transformations, i.e., “the form of the Dirac Equation is identical in equivalent frames of reference”– [131]. A Lorentz transform tells us that the coordinates transform according to

$$x'^{\nu} = \Lambda^{\nu}{}_{\mu}x^{\mu}\,. \tag{C.42}$$

The associated differential operator will also be transformed. To understand how the differential operator is transformed, we note that the Lorentz transforms leave the quantity $s^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$ invariant, i.e.

$$\begin{aligned} & x^\mu g_{\mu\nu} x^\nu = x'^\mu g_{\mu\nu} x'^\nu \\ \Rightarrow\, & x^\mu g_{\mu\nu} x^\nu = \left(\Lambda^\mu{}_\sigma x^\sigma\right) g_{\mu\nu} \left(\lambda^\mu{}_\rho x^\rho\right) \\ \Rightarrow\, & g_{\sigma\rho} x^\sigma x^\rho = \Lambda^\mu{}_\sigma g_{\mu\nu} \Lambda^\nu{}_\rho x^\sigma x^\rho \\ \Rightarrow\, & \Lambda^\mu{}_\sigma g_{\mu\nu} \Lambda^\nu{}_\rho = g_{\sigma\rho}\,. \end{aligned} \tag{C.43}$$

We now multiply (C.42) by $\Lambda^\rho{}_\sigma g_{\rho\nu}$,

$$\Lambda^\rho{}_\sigma g_{\rho\nu} x'^\nu = \Lambda^\rho{}_\sigma g_{\rho\nu} \Lambda^\nu{}_\mu x^\mu = g_{\sigma\mu} x^\mu\,. \tag{C.44}$$

We now use the identity $g^{\mu\sigma} g_{\sigma\mu} = \delta^\mu{}_\mu$ to find

$$x^\mu = g^{\mu\sigma} \Lambda^\rho{}_\sigma g_{\rho\nu} x'^\nu\,. \tag{C.45}$$

Then by the chain rule we find

$$\partial'_\nu = \frac{\partial x^\mu}{\partial x'^\nu} \frac{\partial}{\partial x^\mu} = g^{\mu\sigma} \Lambda^\rho{}_\sigma g_{\rho\nu} \partial_\mu = \Lambda_\nu{}^\mu \partial_\mu\,, \tag{C.46}$$

where $\Lambda_\mu{}^\nu = (\Lambda^\nu{}_\mu)^{-1}$. We now multiply this equation by $\Lambda^\nu{}_\rho$, yielding

$$\partial_\rho = \Lambda^\nu{}_\rho \partial'_\nu\,. \tag{C.47}$$

Alternatively, we could have used the chain rule to find

$$\partial_\mu = \frac{\partial x'^\nu}{\partial x^\mu} \frac{\partial}{\partial x'^\nu} = \Lambda^\nu{}_\mu \partial'_\nu\,, \tag{C.48}$$

where we used (C.42) to perform the differentiation, and managed to avoid quite a bit of algebra to get to the same point.

First we consider a frame F with an observer O, then O describes the particle using the wave function $\psi(x^\mu)$, which obeys the equation

$$(\mathrm{i}\gamma^\mu \partial_\mu - m)\,\psi(x^\mu) = 0\,. \tag{C.49}$$

For Lorentz invariance, we then want an observer O' in frame F' to describe the same particle using the wave function $\psi'(x'^\nu)$, which fulfills the equation

$$(\mathrm{i}\gamma'^\nu \partial'_\nu - m)\,\psi'(x'^\nu) = 0\,. \tag{C.50}$$

since we are looking for form invariance, the γ'^ν matrices must fulfill the same properties as the γ^μ matrices, which are uniquely defined up to to a similarity transform, thus we are looking for

$$(\mathrm{i}\gamma^\nu \partial'_\nu - m)\,\psi'(x'^\nu) = 0\,, \tag{C.51}$$

which is of the same form as (C.49). The transformation which takes $\psi(x^\mu) \to \psi'(x'^\nu)$ is assumed to be

$$\psi'(x'^\nu) = S(\Lambda)\psi(x^\mu)\,. \tag{C.52}$$

In flat space, we can safely assume that the transformation matrix $S(\Lambda)$ is independent of the coordinates, i.e., it commutes with the differential operator ∂_μ. Applying the transform to (C.49) we then find

$$\big(\mathrm{i}S(\Lambda)\gamma^\mu S^{-1}(\Lambda)\partial_\mu - m\big)\,\psi'(x'^\nu) = \big(\mathrm{i}S(\Lambda)\gamma^\mu S^{-1}(\Lambda)\Lambda^\nu{}_\mu \partial'_\nu - m\big)\,\psi'(x'^\nu) = 0\,. \tag{C.53}$$

By comparing this result to (C.51) we deduce that

$$S(\Lambda)\gamma^{\mu}S^{-1}(\Lambda)\Lambda^{\nu}{}_{\mu} = \gamma^{\nu}\,. \tag{C.54}$$

We will be using this identity to construct the operator S, to do so we first rework this equation into a more useable form. We begin by multiplying both sides of the equation by $\Lambda^{\sigma}{}_{\rho}g_{\sigma\nu}$ on the left, yielding

$$\begin{aligned} S(\Lambda)g_{\rho\mu}\gamma^{\mu}S^{-1}(\Lambda) &= \Lambda^{\sigma}{}_{\rho}g_{\sigma\nu}\gamma^{\nu} \\ \Rightarrow S(\Lambda)\gamma_{\rho}S^{-1}(\Lambda) &= \Lambda^{\sigma}{}_{\rho}\gamma_{\sigma}\,, \end{aligned} \tag{C.55}$$

where we used (C.43) and $g_{\mu\nu}\gamma^{\nu} = \gamma_{\mu}$. Finally we multiply both sides by $g^{\rho\mu}$, yielding

$$S(\Lambda)\gamma^{\mu}S^{-1}(\Lambda) = \Lambda^{\nu\mu}\gamma_{\nu}\,, \tag{C.56}$$

where $\Lambda^{\nu\mu} = g^{\mu\rho}\Lambda^{\nu}{}_{\rho}$.

We are now ready to begin constructing S. We begin by looking at the infinitesimal Lorentz transformation, which is of the form

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \epsilon^{\mu}{}_{\nu}\,. \tag{C.57}$$

Now, (C.43) can be rewritten as

$$\Lambda^{\sigma}{}_{\mu}g^{\mu\nu}\Lambda^{\rho}{}_{\nu} = g^{\sigma\rho}\,. \tag{C.58}$$

Plugging (C.57) into (C.58) we find

$$\left(1+\epsilon^{\sigma}{}_{\mu}\right)g^{\mu\nu}\left(1+\epsilon^{\rho}{}_{\nu}\right) = \left(g^{\mu\nu}+\epsilon^{\sigma\nu}\right)\left(1+\epsilon^{\rho}{}_{\nu}\right) = g^{\mu\nu}+\epsilon^{\sigma\nu}+\epsilon^{\rho\mu} = g^{\sigma\rho}\,, \tag{C.59}$$

to the first order in ϵ, where $\epsilon^{\sigma}\nu = \epsilon^{\sigma}{}_{\rho} g^{\rho\nu}$. Notice that there is no summation in this equation, so we can set $\sigma = \mu$ and $\rho = \nu$, giving us

$$\epsilon^{\mu\nu} + \epsilon^{\nu\mu} = 0\,, \tag{C.60}$$

i.e., $\epsilon^{mu\nu}$ is antisymmetric. We now use the appropriate elements of the metric g to raise the indices of (C.57) to find

$$\Lambda^{\mu\nu} = g^{\mu\nu} + \epsilon^{\mu\nu}\,. \tag{C.61}$$

The corresponding infinitesimal transformation $S(\epsilon^{\mu\nu})$ can be written as

$$S(\Lambda) = S(\epsilon^{\mu\nu}) = 1 - \frac{\mathrm{i}}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu}\,, \quad S^{-1}(\epsilon^{\mu\nu}) = 1 + \frac{\mathrm{i}}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu}\,. \tag{C.62}$$

It is fairly trivial to prove the inverse, we use the fact that $(a+b)(a-b) = a^2 + b^2$ and the fact that we are only keeping terms to the first order in $\epsilon^{\mu\nu}$ to find $SS^{-1} = 1$. We now plug (C.62) into (C.56), yielding

$$\begin{aligned}
&\left(1 - \frac{\mathrm{i}}{4}\sigma_{\alpha\beta}\epsilon^{\alpha\beta}\right)\gamma^{\mu}\left(1 + \frac{\mathrm{i}}{4}\sigma_{\alpha\beta}\epsilon^{\alpha\beta}\right) = \left(g^{\mu\nu} + \epsilon^{\nu\mu}\right)\gamma_{\nu} \\
\Rightarrow &\left(\gamma^{\mu} - \frac{\mathrm{i}}{4}\sigma_{\alpha\beta}\gamma^{\mu}\epsilon^{\alpha\beta}\right)\left(1 + \frac{\mathrm{i}}{4}\sigma_{\alpha\beta}\epsilon^{\alpha\beta}\right) = g^{\mu\nu}\gamma_{\nu} + \epsilon^{\nu\mu}\gamma_{\nu} \\
\Rightarrow &\,\gamma^{\mu} - \frac{\mathrm{i}}{4}\sigma_{\alpha\beta}\gamma^{\mu}\epsilon^{\alpha\beta} + \frac{\mathrm{i}}{4}\sigma_{\alpha\beta}\gamma^{\mu}\epsilon^{\alpha\beta} = \gamma^{\mu} + \epsilon^{\nu\mu}\gamma_{\nu} \\
\Rightarrow &-\frac{\mathrm{i}}{4}\epsilon^{\alpha\beta}\left(\sigma_{\alpha\beta}\gamma^{\mu} - \gamma^{\mu}\sigma_{\alpha\beta}\right) = \epsilon^{\nu\mu}\gamma_{\nu}\,.
\end{aligned} \tag{C.63}$$

Now, we can rewrite the r.h.s. as follows

$$\epsilon^{\nu\mu}\gamma_{\nu} = \frac{1}{2}\epsilon^{\alpha\mu}\gamma_{\alpha} + \frac{1}{2}\epsilon^{\beta\mu}\gamma_{\beta} = \frac{1}{2}\epsilon^{\alpha\beta}\delta^{\mu}{}_{\beta}\gamma_{\alpha} + \frac{1}{2}\epsilon^{\beta\alpha}\delta^{\mu}{}_{\alpha}\gamma_{\beta} = \frac{1}{2}\epsilon^{\alpha\beta}\left(\delta^{\mu}{}_{\beta}\gamma_{\alpha} - \delta^{\mu}{}_{\alpha}\gamma_{\beta}\right)\,, \tag{C.64}$$

where we used the fact that $\epsilon^{\mu\nu} = -\epsilon^{\nu\mu}$, as shown in (C.60). Combining this result with (C.63), and simplifying we find

$$[\sigma_{\alpha\beta}, \gamma^\mu] = 2\mathrm{i}\left(\delta^\mu{}_\beta\gamma_\alpha - \delta^\mu{}_\alpha\gamma_\beta\right) . \tag{C.65}$$

A solution to (C.65) is

$$\sigma_{\alpha\beta} = \frac{\mathrm{i}}{2}\left[\gamma_\alpha, \gamma_\beta\right] . \tag{C.66}$$

We will now show that (C.66) satisfies (C.65),

$$\begin{aligned}
[\sigma_{\alpha\beta}, \gamma^\mu] &= \frac{\mathrm{i}}{2}\left(\gamma_\alpha\gamma_\beta\gamma^\mu - \gamma_\beta\gamma_\alpha\gamma^\mu - \gamma^\mu\gamma_\alpha\gamma_\beta + \gamma^\mu\gamma_\beta\gamma_\alpha\right) \\
&= \frac{\mathrm{i}}{2}\left(\gamma_\alpha\{\gamma_\beta, \gamma^\mu\} - \gamma_\alpha\gamma^\mu\gamma_\beta - \gamma_\beta\{\gamma_\alpha, \gamma^\mu\} + \gamma_\beta\gamma^\mu\gamma_\alpha\right. \\
&\quad \left.-\{\gamma^\mu, \gamma_\alpha\}\gamma_\beta + \gamma_\alpha\gamma^\mu\gamma_\beta + \{\gamma^\mu, \gamma_\beta\}\gamma_\alpha - \gamma_\beta\gamma^\mu\gamma_\alpha\right) \\
&= 2\mathrm{i}\left(\delta^\mu{}_\beta\gamma_\alpha - \delta^\mu{}_\alpha\gamma_\beta\right) ,
\end{aligned} \tag{C.67}$$

confirming (C.65). To get the result (C.67), we used $\{\gamma^\mu, \gamma_\nu\} = \{\gamma_\nu, \gamma^\mu\} = 2\delta^\mu{}_\nu$. This property comes from the definition of $g^{\mu\nu} = \frac{1}{2}\{\gamma^\mu, \gamma^\nu\}$ and the fact that $g^{\mu\nu}g_{\nu\rho} = \delta^\mu{}_\rho$, explicitly

$$\{\gamma^m u, \gamma_\nu\} = g_{\nu\rho}\{\gamma^\mu, \gamma^\rho\} = 2g_{\nu\rho}g^{\mu\rho} = 2g^{\mu\rho}g_{\rho\nu} = 2\delta^\mu{}_\nu\,. \tag{C.68}$$

We have thus confirmed (C.56), and therefore (C.54), with our definition of $\sigma_{\mu\nu}$ (C.66) and the infinitesimal $S(\Lambda)$ (C.62). So the Dirac equation is invariant under an infinitesimal Lorentz transform, and as such must be invariant under a finite transform, which can be interpreted as a series of infinitesimal transforms, given by

$$\Lambda^{\mu}{}_{\nu} = \exp\left(\frac{1}{2}\Omega^{\alpha\beta}(M_{\alpha\beta})^{\mu}{}_{\nu}\right) , \quad (M_{\alpha\beta})^{\mu}{}_{\nu} = g^{\mu}{}_{\alpha}g_{\nu\beta} - g^{\mu}{}_{\beta}g_{\nu\alpha} \,, \tag{C.69}$$

$$S(\Lambda) = \exp\left(-\frac{\mathrm{i}}{4}\Omega^{\alpha\beta}\sigma_{\alpha\beta}\right) . \tag{C.70}$$

Then for the infinitesimal transform we left $\Omega^{\alpha\beta} \to \epsilon^{\alpha\beta} \ll 1$.

Notice that

$$\Omega^{\alpha\beta}(M_{\alpha\beta})^{\mu}{}_{\nu} = \Omega^{\alpha\beta}\left(g^{\mu}{}_{\alpha}g_{\nu\beta} - g^{\mu}{}_{\beta}g_{\nu\alpha}\right) = \Omega^{\mu\beta}g_{\beta\nu} - \left(-\Omega^{\mu\alpha}\right)g_{\alpha\nu} = 2\Omega^{\mu}{}_{\nu} \,, \tag{C.71}$$

where we have used the fact that $\Omega^{\alpha\beta} = -\Omega^{\beta\alpha}$, i.e., it antisymmetric.

APPENDIX D

GENERAL RELATIVITY: A CRASH COURSE

D.1. GENERAL OUTLINE

In this chapter (inspired by the notes of Rainer Dick [128]) we will be going over some of the basics of General Relativity, which should give some context to the material in chapter 8 to those less familiar with the subject. The main body of this thesis deals with general relativity in the context of relativistic quantum mechanics, here we will look at general relativity in the classical sense, without the added intricacies of quantum dynamics. As in the main body of this work we will be using lowercase Greek characters for the holonomic spacetime ($\mu, \nu, ... = 0, 1, 2, 3$), lower case Latin characters starting at i for holonomic space ($i, j, k, ... = 1, 2, 3$), capital Latin characters for anholonomic spacetime, i.e., the anholonomic basis ($A, B, C, ... = 0, 1, 2, 3$), and capital Latin characters starting at I for the anholonomic space ($I, J, K... = 1, 2, 3$). Additionally we will be using η for the Minkowski metric, $[\eta_{AB}] = \text{diag}[1, -1, -1, -1]$, and any other metric will be denoted using g.

Let us begin by considering an object moving in two dimensional space. The object's position can be described by the vector $\vec{r}$, which in turn is defined using a linear combination of linearly independent vectors. For example, in 2D Cartesian coordinates we would have

$$d\vec{r} = dx\, \hat{e}_x + dy\, \hat{e}_y\,, \tag{D.1}$$

which is of course exceptionally useful since we choose our linearly independent to be orthonormal. However, there is nothing that forces us to use orthogonal, nor normal, basis vectors. Thus a more general description of $\vec{r}$ would be

$$d\vec{r} = dx^1 \hat{\mathbf{e}}_1 + dx^2 \hat{\mathbf{e}}_2 = dx^I \hat{\mathbf{e}}_I\,. \tag{D.2}$$

We can then define the corresponding metric tensor using the basis vectors, i.e.

$$g_{IJ} = \hat{\mathbf{e}}_I \cdot \hat{\mathbf{e}}_J \,, \quad [g_{IJ}] = \begin{pmatrix} \hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_2 \\ \hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_2 \end{pmatrix} . \tag{D.3}$$

Then the square of any vector in thus basis is

$$dr^2 = d\vec{r} \cdot d\vec{r} = dx^I \hat{\mathbf{e}}_I \cdot dx^J \hat{\mathbf{e}}_J = g_{IJ} dx^I dx^J \,. \tag{D.4}$$

Note that this is true in all bases. We can also define the inverse of the metric tensor as $[g_{IJ}]^{-1} = [g^{IJ}]$, in which case $[g_{IJ}] \cdot [g^{IJ}] = \mathbb{1}$, or

$$g^{IJ} g_{JK} = \delta^I_K \,. \tag{D.5}$$

Using the inverse metric tensor, we can define a new set of basis vectors, denoted as $\hat{\mathbf{e}}^I$. Furthermore we require that

$$\hat{\mathbf{e}}^I \cdot \hat{\mathbf{e}}_K = \delta^I_K \,. \tag{D.6}$$

By equating (D.5) and (D.6), and using the definition of g_{IJ} (D.3), we find

$$\hat{\mathbf{e}}^I \cdot \hat{\mathbf{e}}_K = g^{IJ} g_{JK} = g^{IJ} \hat{\mathbf{e}}_J \cdot \hat{\mathbf{e}}_K \,, \tag{D.7}$$

which leaves us to conclude that

$$\hat{\mathbf{e}}^I = g^{IJ} \hat{\mathbf{e}}_J \,. \tag{D.8}$$

The inversion of which is

$$\hat{\mathbf{e}}_J = g_{IJ}\hat{\mathbf{e}}^I \, . \tag{D.9}$$

We can then calculate the inverse of the metric tensor

$$g^{IJ} = g^{IK}\delta_K^J = g^{IK}g_{KL}g^{LJ} = g^{IK}g^{LJ}\hat{\mathbf{e}}_K \cdot \hat{\mathbf{e}}_L = \left(g^{IK}\hat{\mathbf{e}}_K\right) \cdot \left(g^{LJ}\hat{\mathbf{e}}_L\right) = \hat{\mathbf{e}}^I \cdot \hat{\mathbf{e}}^J \, . \tag{D.10}$$

There are a variety of different sets of basis vectors, and the basis set that one works in can have a significant impact on the complexity of the problem. As such it is useful to know how to transform from one basis set to another. Suppose we want to transform an equation from the basis set $\hat{\mathbf{e}}_I$ to the basis set $\hat{\mathbf{e}}_i$ (note that the different basis sets are differentiated using upper and lower case Latin characters). Then there must be a linear combination of the original basis set which results in the new basis set, i.e.

$$\hat{\mathbf{e}}_i = e_i^J \hat{\mathbf{e}}_J \, . \tag{D.11}$$

We can write the vector $\vec{r}$ in terms of both the original basis and the new basis, i.e.

$$\vec{r} = x^i \hat{\mathbf{e}}_i = x^i e_i^J \hat{\mathbf{e}}_J = x^J \hat{\mathbf{e}}_J \, , \tag{D.12}$$

giving us

$$x^i e_i^J = x^J \, . \tag{D.13}$$

We define the inverse of e_i^J as e_J^i, or

$$e_i^J e_K^i = \delta_K^J \, , \quad e_j^I e_I^k = \delta_j^k \, . \tag{D.14}$$

Thus

$$x^i = e^i_J x^J \,. \tag{D.15}$$

We can also find the elements of the tensors in the new basis set,

$$g_{ij} = \hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = e^K_i \hat{\mathbf{e}}_K \cdot e^L_j \hat{\mathbf{e}}_L = e^K_i e^L_j g_{KL} \,, \quad g^{ij} = e^i_K e^j_L g^{KL} \,. \tag{D.16}$$

On the other hand, we can look at what happens when we transform the dual basis vectors, for which $x^J = x_I g^{IJ}$ and $x_I = g_{IJ} x^J$. Then

$$dx_i = g_{ij} dx^j = e^K_i e^L_j g_{KL} dx^j = e^K_i g_{KL} dx^L = e^K_i dx_K \,. \tag{D.17}$$

So while the components of the basis sets transform contravariantly, the components of the dual basis set transform covariantly. Armed with these results we can show that the scalar product is invariant under coordinate transforms. To show this we begin with two vectors in the original basis set,

$$\vec{u} = u^I \hat{\mathbf{e}}_I \,, \quad \vec{v} = v^I \hat{\mathbf{e}}_I \,, \quad \vec{u} \cdot \vec{v} = \hat{\mathbf{e}}_I \cdot \hat{\mathbf{e}}_J u^I v^J = g_{IJ} u^I v^J = u^I v_I = u_J v^J \,, \tag{D.18}$$

this result can then be compared to the result which we obtain from the transformed vectors $\vec{u}$ and $\vec{v}$,

$$\vec{v} \cdot \vec{u} = g_{ij} u^i v^j = u^i v_i = u^J e^i_J v_i = u^J v_J \,. \tag{D.19}$$

Thus, as stated, and as should be expected, the scalar product of two vectors is invariant under coordinate transforms.

Due of the complex nature of many of the bases, movement which is simply described in one basis, i.e., movement along the x-axis in the Cartesian coordinate system, is more complicated to describe in other systems. As such, we must formulate a system for describing such movements within the basis transform. Let us start with the vector $\vec{r}(x^j) = x^I(x^j)\hat{e}_I$ in the Cartesian basis. Then the vector $d\vec{r}(x^j)$ which connects $\vec{r}(x^j)$ and $\vec{r}(x^j + dx^j)$ is

$$d\vec{r}(x^j) = \vec{r}(x^j + dx^j) - \vec{r}(x^j) = dx^k \partial_k \vec{r}(x^j) = dx^k \hat{\mathbf{e}}_k \,, \tag{D.20}$$

which tells us that

$$\hat{\mathbf{e}}_k = \partial_k \vec{r} = \partial_k x^I \hat{e}_I \,, \tag{D.21}$$

We can then conclude that

$$e_k^I = \partial_k x^I \,. \tag{D.22}$$

We can also solve for the distance squared between the two points,

$$ds^2 = d\vec{r}^2 = \big(dx^i \hat{\mathbf{e}}_i(x)\big) \cdot \big(dx^j \hat{\mathbf{e}}_j(x)\big) = dx^i dx^j \hat{\mathbf{e}}_i(x) \cdot \hat{\mathbf{e}}_j(x) = g_{ij} dx^i dx^j \,. \tag{D.23}$$

While most of the work done in this section was done in two dimensions, all of it can be generalized two four dimensions (spacetime), and will apply to our 4-vectors. We would simply have $I, J, K... \to A, B, C...$ and $i, j, k... \to \mu, \nu, \rho....$

D.2. HOLONOMIC COVARIANT DERIVATIVE

For a vector $A = A^\mu \hat{\mathbf{e}}_\mu$ in the holonomic basis, we take the partial derivative an find

$$\begin{aligned} \partial_\nu A &= \partial_\nu \left(A^\mu \hat{\mathbf{e}}_\mu\right) = \left(\partial_\nu A^\mu\right) \hat{\mathbf{e}}_\mu + A^\mu \partial_\nu \hat{\mathbf{e}}_\mu = \left(\partial_\nu A^\mu\right) \hat{\mathbf{e}}_\mu + A^\mu \Gamma^\rho_{\mu\nu} \hat{\mathbf{e}}_\rho \\ &= \big(\partial_\nu A^\mu + A^\rho \Gamma^\mu_{\nu\rho}\big) \hat{\mathbf{e}}_\mu = \left(\nabla_\nu A^\mu\right) \hat{\mathbf{e}}_\mu \,, \end{aligned} \tag{D.24}$$

where the Christoffel symbols (of the second kind) are defined as

$$\partial_\nu \hat{\mathbf{e}}_\mu = \Gamma^\rho_{\mu\nu} \hat{\mathbf{e}}_\rho \,. \tag{D.25}$$

Using the fact that $\hat{\mathbf{e}}^\rho \cdot \hat{\mathbf{e}}_\rho = 1$, we find that the Christoffel symbols can be expressed as

$$\Gamma^\rho_{\mu\nu} = \hat{\mathbf{e}}^\rho \cdot \partial_\nu \hat{\mathbf{e}}_\mu \,. \tag{D.26}$$

We can also rewrite the Christoffel symbols in terms of the metric g. The key observation is to recall that by definition $\hat{\mathbf{e}}_\mu = \partial_\mu \vec{r}$ (spacetime generalization of (D.21)). Thus [129]

$$\partial_\nu \hat{\mathbf{e}}_\mu = \partial_\nu \partial_\mu \vec{r} = \partial_\mu \hat{\mathbf{e}}_\nu = \frac{1}{2}\left(\partial_\nu \hat{\mathbf{e}}_\mu + \partial_\mu \hat{\mathbf{e}}_\nu\right) \,. \tag{D.27}$$

Applying this to (D.26), we find

$$\begin{aligned}
\Gamma^\rho_{\mu\nu} &= \hat{\mathbf{e}}^\rho \cdot \frac{1}{2}\left(\partial_\nu \hat{\mathbf{e}}_\mu + \partial_\mu \hat{\mathbf{e}}_\nu\right) = \frac{1}{2} g^{\rho\sigma}\left(\hat{\mathbf{e}}_\sigma \cdot \partial_\nu \hat{\mathbf{e}}_\mu + \hat{\mathbf{e}}_\sigma \cdot \partial_\mu \hat{\mathbf{e}}_\nu\right) \\
&= \frac{1}{2} g^{\rho\sigma}\left[\left(\hat{\mathbf{e}}_\sigma \cdot \partial_\nu \hat{\mathbf{e}}_\mu + \hat{\mathbf{e}}_\mu \cdot \partial_\nu \hat{\mathbf{e}}_\sigma\right) + \left(\hat{\mathbf{e}}_\sigma \cdot \partial_\mu \hat{\mathbf{e}}_\nu + \hat{\mathbf{e}}_\nu \cdot \partial_\mu \hat{\mathbf{e}}_\sigma\right) - \left(\hat{\mathbf{e}}_\mu \cdot \partial_\nu \hat{\mathbf{e}}_\sigma + \hat{\mathbf{e}}_\nu \cdot \partial_\mu \hat{\mathbf{e}}_\sigma\right)\right] \\
&= \frac{1}{2} g^{\rho\sigma}\left[\left(\hat{\mathbf{e}}_\sigma \cdot \partial_\nu \hat{\mathbf{e}}_\mu + \hat{\mathbf{e}}_\mu \cdot \partial_\nu \hat{\mathbf{e}}_\sigma\right) + \left(\hat{\mathbf{e}}_\sigma \cdot \partial_\mu \hat{\mathbf{e}}_\nu + \hat{\mathbf{e}}_\nu \cdot \partial_\mu \hat{\mathbf{e}}_\sigma\right) - \left(\hat{\mathbf{e}}_\mu \cdot \partial_\sigma \hat{\mathbf{e}}_\nu + \hat{\mathbf{e}}_\nu \cdot \partial_\sigma \hat{\mathbf{e}}_\mu\right)\right] \\
&= \frac{1}{2} g^{\rho\sigma}\left[\partial_\nu\left(\hat{\mathbf{e}}_\sigma \cdot \hat{\mathbf{e}}_\mu\right) + \partial_\mu\left(\hat{\mathbf{e}}_\sigma \cdot \hat{\mathbf{e}}_\nu\right) - \partial_\sigma\left(\hat{\mathbf{e}}_\mu \cdot \hat{\mathbf{e}}_\nu\right)\right] = \frac{1}{2} g^{\rho\sigma}\left(\partial_\nu g_{\sigma\mu} + \partial_\mu g_{\sigma\nu} - \partial_\sigma g_{\mu\nu}\right)
\end{aligned} \tag{D.28}$$

Let us go through the derivation in detail. In doing from the first to the second line we added and subtracted identical terms. Going from the second to the third line we used the identity $\partial_\mu \hat{\mathbf{e}}_\nu = \partial_\nu \hat{\mathbf{e}}_\mu$ on the last term in the square brackets. Finally we apply the product rule (in reverse), and use the fact that $\hat{\mathbf{e}}_\mu \cdot \hat{\mathbf{e}}_\nu = g_{\mu\nu}$. We can now define the covariant derivative operating in the holonomic basis as

$$\nabla_\nu A^\mu \equiv \partial_\nu A^\mu + \Gamma^\mu_{\nu\rho} A^\rho \,. \tag{D.29}$$

D.3. ANHOLONOMIC DERIVATIVE

We have now seen the emergence of the Christoffel symbols in the covariant derivative by looking at how a vector in the holonomic space is differentiated. Now let us look at how a vector in the anholonomic space is effected by the covariant derivative. We start with a vector $V = V^A \hat{\mathbf{e}}_A$ in the anholonomic basis. Then

$$\begin{aligned} \partial_\mu V =& \partial_\mu V^A \hat{\mathbf{e}}_A = \left(\partial_\mu V^A\right) \hat{\mathbf{e}}_A + V^A \partial_\mu \hat{\mathbf{e}}_A = \left(\partial_\mu V^A\right) \hat{\mathbf{e}}_A + V^A \partial_\mu \left[e^\rho_A \hat{\mathbf{e}}_\rho\right] \\ &= \left(\partial_\mu V^A\right) \hat{\mathbf{e}}_A + V^A \left[\left(\partial_\mu e^\rho_A\right) \hat{\mathbf{e}}_\rho + e^\rho_A \partial_\mu \hat{\mathbf{e}}_\rho\right] \,, \end{aligned} \tag{D.30}$$

we then apply the definition of the Christoffel symbols (D.25), yielding

$$\partial_\mu V = \left(\partial_\mu V^A\right) \hat{\mathbf{e}}_A + V^A \left[\left(\partial_\mu e^\rho_A\right) \hat{\mathbf{e}}_\rho + e^\rho_A \Gamma^\lambda_{\rho\mu} \hat{\mathbf{e}}_\lambda\right] \,, \tag{D.31}$$

we now apply the identity $\hat{\mathbf{e}}_\rho = e^A_\rho \hat{\mathbf{e}}_A$,

$$\begin{aligned} \partial_\mu V &= \left(\partial_\mu V^A\right) \hat{\mathbf{e}}_A + V^A \left[\left(\partial_\mu e^\rho_A\right) \left(e^B_\rho \hat{\mathbf{e}}_B\right) + e^\rho_A \Gamma^\lambda_{\rho\mu} \left(e^B_\lambda \hat{\mathbf{e}}_B\right)\right] \\ &= \left(\partial_\mu V^B\right) \hat{\mathbf{e}}_B + V^A \left[e^B_\rho \partial_\mu e^\rho_A + e^B_\rho \Gamma^\rho_{\mu\lambda} e^\lambda_A\right] \hat{\mathbf{e}}_B \\ &= \left(\partial_\mu V^B + V^A\, e^B_\rho \left[\partial_\mu e^\rho_A + \Gamma^\rho_{\mu\lambda} e^\lambda_A\right]\right) \hat{\mathbf{e}}_B \,, \end{aligned} \tag{D.32}$$

we now use the definition of the holonomic covariant derivative (D.29) to rewrite this as

$$\partial_\mu V = \left(\partial_\mu V^B + e^B_\rho \nabla_\mu e^\rho_A\right) \hat{\mathbf{e}}_B = \left(\partial_\mu V^B + \omega^B_{\mu A}\right) \hat{\mathbf{e}}_B \tag{D.33}$$

where we define the Ricci rotation coefficient as

$$\omega^B_{\mu A} \equiv e^B_\rho \nabla_\mu e^\rho_A \,. \tag{D.34}$$

We then find that the anholonomic covariant derivative is

$$\nabla_\mu V^A \equiv \partial_\mu V^A + \omega^A_{\mu B}\,. \tag{D.35}$$

Using the fact that $e^A_\rho e^\rho_C = \delta^A_C$, and $\partial_\mu \delta^A_C = 0$ we find

$$\omega^{AB}_\mu = -\omega^{BA}_\mu\,. \tag{D.36}$$

D.4. MOTION OF A PARTICLE IN SPACETIME

We now switch back to general relativity. Earlier we solved for the distance squared between two points in 2D. In our 4-vector space time, it is generalized to be

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu\,. \tag{D.37}$$

In flat–spacetime this would be

$$ds^2 = \eta_{AB} dx^A dx^B = c^2 dt^2 - d\vec{x}^2\,. \tag{D.38}$$

Then by taking the square root and integrating we find the action to be

$$S = mc \int dt \sqrt{c^2 - \dot{\vec{x}}^{\,2}} = mc \int \sqrt{g_{\mu\nu} dx^\mu dx^\nu} = mc \int d\xi \sqrt{g_{\mu\nu} \frac{dx^\mu}{d\xi} \frac{dx^\nu}{d\xi}}\,. \tag{D.39}$$

From here, we can extract the Lagrangian

$$\mathcal{L} = \sqrt{g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu}\,, \tag{D.40}$$

where $\dot{x}^\mu = dx^\mu/d\xi$. Recall that the proper time is given by

$$d\tau^2 = dt^2 - \frac{1}{c^2}d\vec{x}^{\,2} = \frac{1}{c^2}\left(c^2dt^2 - d\vec{x}^{\,2}\right) = \frac{1}{c^2}g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu d\xi^2\,, \tag{D.41}$$

then

$$cd\tau = \sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}d\xi\,. \tag{D.42}$$

We can then find the equations of motion using the fact that

$$\left(\frac{\partial}{\partial x^\mu} - \frac{d}{d\xi}\frac{\partial}{\partial \dot{x}^\mu}\right)\mathcal{L} = 0\,. \tag{D.43}$$

Looking at these differentials separately we find

$$\frac{\partial}{\partial x^\mu}\mathcal{L} = \partial_\mu\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta} = \frac{1}{2}\frac{1}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}}\partial_\mu g_{\rho\nu}\dot{x}^\rho\dot{x}^\nu\,, \tag{D.44}$$

and

$$\begin{aligned}\frac{d}{d\xi}\frac{\partial}{\partial\dot{x}^\mu}\mathcal{L} &= \frac{d}{d\xi}\frac{\partial}{\partial\dot{x}^\mu}\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta} = \frac{d}{d\xi}\left(\frac{1}{2}\frac{1}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}}\frac{\partial}{\partial\dot{x}^\mu}g_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma\right)\\ &= \frac{d}{d\xi}\left(\frac{1}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}}g_{\mu\nu}\dot{x}^\mu\right)\\ &= g_{\mu\nu}\dot{x}^\mu\frac{d}{d\xi}\frac{1}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}} + g_{\mu\nu}\ddot{x}^\mu\frac{1}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{c}^\beta}} + \dot{x}^\mu\dot{x}^\rho\partial_\rho g_{\mu\nu}\frac{1}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{c}^\beta}}\,,\end{aligned} \tag{D.45}$$

where we used $\frac{d}{d\xi} = \frac{dx^\sigma}{d\xi}\frac{d}{dx^\sigma} = \dot{x}^\sigma\partial_\sigma$. Notice that

$$f(x)\frac{d}{dx}\ln\left(C\,f(x)\right) = f(x)\frac{f'(x)}{f(x)} = f'(x)\,, \tag{D.46}$$

where C is a constant. Then

$$\frac{d}{d\xi}\frac{1}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}} = \frac{1}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}}\frac{d}{d\xi}\ln\left(\frac{c}{\sqrt{g_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma}}\right)$$
$$= -\frac{1}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}}\frac{d}{d\xi}\ln\left(\frac{1}{c}\sqrt{g_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma}\right), \tag{D.47}$$

and (D.45) becomes

$$\frac{d}{d\xi}\frac{\partial}{\partial\dot{x}^\mu}\mathcal{L} = \frac{1}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}}\left(-g_{\mu\nu}\dot{x}^\mu\frac{d}{d\xi}\ln\left(\frac{1}{c}\sqrt{g_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma}\right) + g_{\mu\nu}\ddot{x}^\mu + \dot{x}^\mu\dot{x}^\rho\partial_\rho g_{\mu\nu}\right). \tag{D.48}$$

We can now plug (D.44) and (D.48) into (D.43), and multiply by $-g^{\mu\rho}$ to find

$$\frac{1}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}}\left(\ddot{x}^\mu + g^{\mu\rho}\left(\partial_\sigma g_{\rho\nu} - \frac{1}{2}\partial_\rho g_{\sigma\nu}\right)\dot{x}^\sigma\dot{x}^\nu - g_{\mu\nu}\dot{x}^\mu\frac{d}{d\xi}\ln\left(\frac{1}{c}\sqrt{g_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma}\right)\right) = 0, \tag{D.49}$$

i.e.

$$\ddot{x}^\mu + \Gamma^\mu_{\sigma\nu}\dot{x}^\sigma\dot{x}^\nu = g_{\mu\nu}\dot{x}^\mu\frac{d}{d\xi}\ln\left(\frac{1}{c}\sqrt{g_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma}\right), \tag{D.50}$$

where we used the identities

$$\Gamma^\mu_{\sigma\nu}\dot{x}^\sigma\dot{x}^\nu = g^{\mu\rho}\left(\partial_\sigma g_{\rho\nu} - \frac{1}{2}\partial_\rho g_{\sigma\nu}\right)\dot{x}^\sigma\dot{x}^\nu. \tag{D.51}$$

If we choose $\xi = \tau$, then (D.42) becomes

$$\sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu} = c, \tag{D.52}$$

and due to the fact that $\ln(1) = 0$, (D.50) becomes

$$\ddot{x}^{\mu} + \Gamma^{\mu}_{\nu\rho}\dot{x}^{\nu}\dot{x}^{\rho} = 0\,. \tag{D.53}$$

We have derived the equations of motion in spacetime.

D.5. SCHWARZSCHILD METRIC

The Schwarzschild metric was derived in 1916 by Karl Schwarzschild [130], and describes the metric for a non–rotating gravitational center, and is given as

$$ds^2 = \left(1 - \frac{r_{\rm s}}{r}\right) c^2 dt^2 - \left(1 - \frac{r_{\rm s}}{r}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\varphi\,, \tag{D.54}$$

where $r_{\rm s}$ is the Schwarzschild radius, and is given by

$$r_{\rm s} = \frac{2GM}{c^2}\,. \tag{D.55}$$

To derive the equations of motion, let us set

$$B = \left(1 - \frac{r_{\rm s}}{r}\right)\,, \quad A = \frac{1}{B} = \left(1 - \frac{r_{\rm s}}{r}\right)^{-1}\,, \tag{D.56}$$

thus we can rewrite the metric as

$$ds^2 = c^2 B^2 dt^2 - A dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta\, d\varphi^2\,. \tag{D.57}$$

We quickly deduce that the only non-vanishing terms of the metric tensor are

$$\begin{aligned} g_{tt} &= -\,c^2 B\,, \quad g_{rr} = A\,, \quad g_{\theta\theta} = r^2\,, \quad g_{\varphi\varphi} = r^2\sin^2\theta\,, \\ g^{tt} &= -\,\frac{1}{c^2 B}\,, \quad g^{rr} = \frac{1}{A}\,, \quad g^{\theta\theta} = \frac{1}{r^2}\,, \quad g^{\varphi\varphi} = \frac{1}{r^2\sin^2\theta}\,. \end{aligned} \tag{D.58}$$

Then using

$$\Gamma^{\rho}_{\nu\mu} = \frac{1}{2} g^{\rho\sigma} (\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\nu\mu}) \,, \tag{D.59}$$

We can solve for all the Christoffel symbols. To make things simpler we notice that all the elements of the metric tensor (co- and contravariant) are nonzero along the diagonal (i.e., $g_{\mu\nu} \neq 0$ iff $\mu = \nu$), as such we will only include these elements when writing out the solutions for the Christoffel symbols. Furthermore, for a Christoffel symbol to be non-vanishing, at least two of it's indices must match (in this case), so we can narrow our scope. After all the calculations are complete, we find that the non–vanishing Christoffel symbols are

$$\begin{gathered}
\Gamma^r{}_{rr} = \frac{A'}{2A} \,, \quad \Gamma^r{}_{\theta\theta} = -\frac{r}{A} \,, \quad \Gamma^r{}_{\varphi\varphi} = -\frac{r \sin^2 \theta}{A} \,, \quad \Gamma^\theta{}_{\theta r} = \Gamma^\varphi{}_{\varphi r} = \frac{1}{r} \,, \\
\Gamma^\theta{}_{\varphi\varphi} = -\sin\theta\cos\theta \,, \quad \Gamma^\varphi{}_{\varphi\theta} = \frac{\cos\theta}{\sin\theta} \,, \quad \Gamma^r{}_{tt} = \frac{c^2 B'}{2A} \,, \quad \Gamma^t{}_{tr} = \frac{B'}{2B} \,.
\end{gathered} \tag{D.60}$$

Then the equations of motion given by

$$\ddot{x}^\mu + \Gamma^\mu{}_{\nu\rho} \dot{x}^\nu \dot{x}^\rho = 0 \,, \tag{D.61}$$

become

$$\frac{d^2 r}{d\tau^2} + \frac{A'}{2A} \left(\frac{dr}{d\tau} \right)^2 - \frac{r}{A} \left(\frac{d\theta}{d\tau} \right)^2 - \frac{r \sin^2\theta}{A} \left(\frac{d\varphi}{d\tau} \right)^2 + \frac{c^2 B'}{2A} \left(\frac{dt}{d\tau} \right)^2 = 0 \,, \tag{D.62}$$

$$\frac{d^2\theta}{d\tau^2} + \frac{2}{r} \frac{d\theta}{d\tau} \frac{dr}{d\tau} - \sin\theta\cos\theta \left(\frac{d\varphi}{d\tau} \right)^2 = 0 \,, \tag{D.63}$$

$$\frac{d^2\varphi}{d\tau^2} + \frac{2}{r} \frac{d\varphi}{d\tau} \frac{dr}{d\tau} + 2 \frac{\cos\theta}{\sin\theta} \frac{d\varphi}{d\tau} \frac{d\theta}{d\tau} = 0 \,, \tag{D.64}$$

$$\frac{d^2 t}{d\tau^2} + \frac{B'}{B} \frac{dr}{d\tau} \frac{dt}{d\tau} = 0 \,. \tag{D.65}$$

Recall

$$\begin{aligned}\hat{e}_r &= \sin\theta\cos\varphi\hat{e}_x + \sin\theta\sin\varphi\hat{e}_y + \cos\theta\hat{e}_z\,,\\ \hat{e}_\theta &= \cos\theta\cos\varphi\hat{e}_x + \cos\theta\sin\varphi\hat{e}_y - \sin\theta\hat{e}_z\,,\\ \hat{e}_\varphi &= -\sin\varphi\hat{e}_x + \cos\varphi\hat{e}_y\,,\end{aligned} \tag{D.66}$$

then

$$\frac{d}{d\tau}\hat{e}_r = \dot{\theta}\hat{e}_\theta + \dot{\varphi}\sin\theta\hat{e}_\varphi\,, \quad \frac{d}{d\tau}\hat{e}_\theta = -\dot{\theta}\hat{e}_r + \dot{\varphi}\cos\theta\hat{e}_\varphi\,, \quad \frac{d}{d\tau}\hat{e}_\varphi = -\dot{\varphi}(\sin\theta\hat{e}_r + \cos\theta\hat{e}_\theta)\,, \tag{D.67}$$

and

$$\dot{\vec{r}} = \dot{r}\hat{e}_r + r\dot{\theta}\hat{e}_\theta + r\dot{\varphi}\sin\theta\hat{e}_\varphi\,, \tag{D.68}$$

$$\begin{aligned}\ddot{\vec{r}} =& (\ddot{r} - r\dot{\theta}^2 - r\dot{\varphi}^2\sin^2\theta)\hat{e}_r + (2\dot{r}\dot{\theta} + r\ddot{\theta} - r\dot{\varphi}^2\sin\theta\cos\theta)\hat{e}_\theta\\ &+ (2\dot{r}\dot{\varphi}\sin\theta + r\ddot{\varphi}\sin\theta + 2r\dot{\varphi}\dot{\theta}\cos\theta)\hat{e}_\varphi\,.\end{aligned} \tag{D.69}$$

We can rewrite (D.63) as

$$\ddot{\theta} = \dot{\varphi}^2\sin\theta\cos\theta - \frac{2}{r}\dot{\theta}\dot{r}\,, \tag{D.70}$$

and plug it into the θ component of (D.69), giving us

$$2\dot{r}\dot{\theta} + r\ddot{\theta} - r\dot{\varphi}^2\sin\theta\cos\theta = 2\dot{r}\dot{\theta} + r\left(\dot{\varphi}^2\sin\theta\cos\theta - \frac{2}{r}\dot{\theta}\dot{r}\right) - r\dot{\varphi}^2\sin\theta\cos\theta = 0\,. \tag{D.71}$$

Similarly, we can rewrite (D.64) as

$$\ddot{\varphi} = -\frac{2}{r}\dot{\varphi}\dot{r} - 2\frac{\sin\theta}{\cos\theta}\dot{\varphi}\dot{\theta}\,, \tag{D.72}$$

and plug it into the φ component of (D.69), yielding

$$2\dot{r}\dot{\varphi}\sin\theta + r\ddot{\varphi}\sin\theta + 2r\dot{\varphi}\dot{\theta}\cos\theta$$
$$= 2\dot{r}\dot{\varphi}\sin\theta + r\left(-\frac{2}{r}\dot{\varphi}\dot{r} - 2\frac{\sin\theta}{\cos\theta}\dot{\varphi}\dot{\theta}\right)\sin\theta + 2r\dot{\varphi}\dot{\theta}\cos\theta = 0\,. \tag{D.73}$$

Thus both of the angular components of $\ddot{\vec{r}}$ vanish, and we are left with

$$\ddot{\vec{r}} = (\ddot{r} - r\dot{\theta}^2 - r\dot{\varphi}^2\sin^2\theta)\hat{e}_r\,, \tag{D.74}$$

and by rewriting (D.62) as

$$\ddot{r} = -\frac{A'}{2A}\dot{r}^2 + \frac{r}{A}\dot{\theta}^2 + \frac{r\sin^2\theta}{A}\dot{\varphi}^2 - \frac{c^2B'}{2A}\dot{t}^2\,, \tag{D.75}$$

(D.74) becomes

$$\ddot{\vec{r}} = -\left(\frac{A'}{2A}\dot{r}^2 - \frac{r}{A}\dot{\theta}^2 - \frac{r\sin^2\theta}{A}\dot{\varphi}^2 + \frac{c^2B'}{2A}\dot{t}^2 + r\dot{\theta}^2 + r\dot{\varphi}^2\sin^2\theta\right)\hat{e}_r\,, \tag{D.76}$$

telling us that the gravity around a spherically symmetric non-rotating mass still produces a force parallel to $\vec{r}$. This means that absent of any outside influence, the motion will remain in a plane. We can choose this plain to be at $\theta = \pi/2$. The equations of motion then reduce to

$$\frac{d^2r}{d\tau^2} + \frac{A'}{2A}\left(\frac{dr}{d\tau}\right)^2 - \frac{r}{A}\left(\frac{d\varphi}{d\tau}\right)^2 + \frac{c^2B'}{2A}\left(\frac{dt}{d\tau}\right)^2 = 0\,, \tag{D.77}$$
$$\frac{d^2\varphi}{d\tau^2} + \frac{2}{r}\frac{d\varphi}{d\tau}\frac{dr}{d\tau} = \frac{1}{r^2}\frac{d}{d\tau}(r^2\dot{\varphi}) = 0\,, \tag{D.78}$$
$$\frac{d^2t}{d\tau^2} + \frac{B'}{B}\frac{dr}{d\tau}\frac{dt}{d\tau} = \frac{1}{B}\frac{d}{d\tau}(B\dot{t}) = 0\,. \tag{D.79}$$

We now have sufficient information to consider how light will be affected by the curvature of spacetime due to a gravitational center.

D.6. BENDING OF LIGHT

Again, the work in this appendix is primarily inspired by [128]. Dealing with massless particles can be somewhat problematic. The first obstacle that we must overcome is the fact that eigentime of massless particles vanishes,

$$c^2 d\tau^2 = -ds^2 = -g_{\mu\nu} dx^\mu dx^\nu = 0\,.$$

Think of it this way, the special theory of relativity (STR) tells us that massless particles must travel at the speed of light. As such massless particles will travel at the same rate regardless of reference frame. Under such conditions the notion of eigentime does not makes much sense, since it was devised to deal with the differences that observers in different reference frames would see. For a particle traveling at the speed of light, these differences no longer exist. With $d\tau = 0$ it is clear that we do not have an eigenvelocity ($u^\mu = \frac{dx^\mu}{d\tau}$). To circumvent these difficulties we will rewrite the equations for massive particles such that they are independent of mass and the eigentime. We will then assume that these hold in the limit where $m \to 0$, which makes sense since the equations are independent of the *particle's mass*.

Let us begin by recalling the equations for massive particles (in terms of their eigentime):

$$\frac{d^2 r}{d\tau^2} + \frac{A'}{2A}\left(\frac{dr}{d\tau}\right)^2 - \frac{r}{A}\left(\frac{d\varphi}{d\tau}\right)^2 + \frac{c^2 B'}{2A}\left(\frac{dt}{d\tau}\right)^2 = 0\,, \tag{D.80}$$

$$\frac{1}{r^2}\frac{d}{d\tau}\left(r^2\dot{\varphi}\right) = 0\,, \tag{D.81}$$

$$\frac{1}{B}\frac{d}{d\tau}\left(B\dot{t}\right) = 0\,. \tag{D.82}$$

From (D.82) we find

$$B\dot{t} = C \quad \Rightarrow \quad \frac{dt}{d\tau} = \frac{C}{B} \quad \Rightarrow \quad d\tau = \frac{B}{C}dt\,, \tag{D.83}$$

where C is a constant. From which we find

$$\begin{aligned}
\frac{d^2r}{d\tau^2} &= \frac{C}{B}\frac{d}{dt}\frac{C}{B}\frac{d}{dr} = \frac{C}{B}\left(-C\frac{B'}{B^2}\left(\frac{dr}{dt}\right)^2 + \frac{C}{B}\frac{d^2r}{dt^2}\right) \\
&= \frac{C^2}{B^2}\left(\frac{d^2r}{dt} - \frac{B'}{B}\left(\frac{dr}{dt}^2\right)\right)\,, \\
\left(\frac{dr}{d\tau}\right)^2 &= \frac{C^2}{B^2}\left(\frac{dr}{dt}\right)^2\,, \quad \left(\frac{d\varphi}{d\tau}\right)^2 = \frac{C^2}{B^2}\left(\frac{d\varphi}{dt}\right)^2\,, \quad \left(\frac{dt}{d\tau}\right)^2 = \frac{C^2}{B^2}\,, \\
\frac{d}{d\tau}\left(r^2\frac{d\varphi}{d\tau}\right) &= \frac{C^2}{B}\frac{d}{dt}\left(\frac{r^2}{B}\frac{d\varphi}{dt}\right)\,.
\end{aligned}$$

Thus (D.80) and (D.81) become (after we divide out common factors)

$$\frac{d^2r}{dt^2} + \left(\frac{A'}{2A} - \frac{B'}{B}\right)\left(\frac{dr}{dt}\right)^2 - \frac{r}{A}\left(\frac{d\varphi}{dt}\right)^2 + \frac{c^2B'}{2A} = 0\,, \tag{D.84}$$

$$\frac{d}{dt}\left(\frac{r^2}{B}\frac{d\varphi}{dt}\right) = 0\,. \tag{D.85}$$

Which are independent of mass and eigentime as desired. (D.85) gives us

$$\dot{\varphi} = \frac{JB}{r^2}\,, \tag{D.86}$$

where J is a constant. Plugging this into (D.84), and multiplying by $\frac{A}{B}\dot{r}$ we find

$$\begin{aligned}
&\frac{A}{B^2}\dot{r}\ddot{r} + \frac{A'}{2B^2}\dot{r}^3 - \frac{B'}{B^3}\dot{r}^3 - \frac{J^2}{r^3}\dot{r} + \frac{c^2B'}{2B^2}\dot{r} = 0 \\
\Rightarrow\ &\frac{d}{dt}\left(\frac{A}{2B^2}\dot{r}^2 + \frac{J^2}{2r^2} - \frac{c^2}{2B}\right) = 0 \\
\Rightarrow\ &\frac{A}{2B^2}\dot{r}^2 + \frac{J^2}{2r^2} - \frac{c^2}{2B} = K\,.
\end{aligned} \tag{D.87}$$

We know that $ds^2 = 0$, massless particle, and we can rewrite ds^2 (bear in mind that $\theta = \frac{\pi}{2}$) as (we will use the equations that we just found plugged into $\dot{\varphi}^2$ and $\dot{r}$)

$$\begin{aligned} ds^2 = & - c^2 B dt^2 + A dr^2 + r^2 d\varphi^2 = \left(-B + \frac{A}{c^2}\dot{r}^2 + \frac{r^2}{c^2}\dot{\varphi}^2 \right) c^2 dt^2 \\ = & \left(-B + \frac{A}{c^2}\frac{2B^2}{A}\left(K - \frac{J^2}{2r^2} + \frac{c^2}{2B} \right) + \frac{r^2}{c^2}\frac{J^2 B^2}{r^4} \right) c^2 dt^2 \,, \end{aligned}$$

most of the terms cancel out, and we are left with

$$ds^2 = 2B^2 K dt^2 = 0\,. \tag{D.88}$$

Thus, for a massless particle

$$K = 0\,. \tag{D.89}$$

The radial equation is then

$$\frac{1}{2}\dot{r} + \frac{B^2 J^2}{2Ar^2} - \frac{c^2}{B}2A = 0\,. \tag{D.90}$$

We can now use a parameter τ which we define as

$$d\tau = B dt = \frac{r - r_{\mathrm{s}}}{r} dt\,. \tag{D.91}$$

This defines the affine parameters for light–like geodesics. τ can be rescaled by an arbitrary *positive* factor. Applying this to the radial equation, along with $A = \frac{1}{B}$, we find

$$\frac{1}{2}B^2\left(\frac{dr}{d\tau}\right)^2 + B^2\frac{J^2}{2Ar^2} - B^2\frac{c^2}{2AB} = 0\,.$$

We now divide out B^2, yielding

$$\frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 + \frac{BJ^2}{2r^2} - \frac{c^2}{2} = 0\,,$$

and plug in for B

$$\frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 + \frac{1}{2}\frac{r-r_\mathrm{s}}{r}\frac{J^2}{r^2} = \frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 + \left(\frac{1}{2} - \frac{GM}{c^2 r}\right)\frac{J^2}{r^2} = \frac{c^2}{2}\,, \tag{D.92}$$

The effective potential is then

$$\widetilde{V} = \frac{1}{2}\frac{r-r_\mathrm{s}}{r}\frac{J^2}{r^2} = \left(\frac{1}{2} - \frac{GM}{c^2 r}\right)\frac{J^2}{r^2}\,. \tag{D.93}$$

In this case, the effective potential will always produce a centrifugal barrier, provided that $J^2 > 0$:

$$\begin{aligned} \frac{d\widetilde{V}}{dr} &= \frac{J^2}{2}\frac{d}{dr}\left(\frac{1}{r^2} - \frac{r_\mathrm{s}}{r^3}\right) = \frac{J^2}{2}\left(-2\frac{1}{r^3} + 3\frac{r_\mathrm{s}}{r^4}\right) = 0 \\ \Rightarrow r &= \frac{3}{2}r_\mathrm{s}\,, \end{aligned} \tag{D.94}$$

i.e., the maximum (it is maximum) height of the effective potential always occurs at $r = \frac{3}{2}r_\mathrm{s}$, this is then the lower bound of the *periastron*, the point of closest approach, of a photon, assuming that it doesn't get sucked in. Clearly, the other extremums can only occur when $r \to \infty$, i.e., there is only the one maximum, and no minima. As such, there are no stable bound orbits in the case of a massless particle. The height of the barrier is

$$\widetilde{V}\left(\frac{3}{2}r_\mathrm{s}\right) = \frac{2}{27}\frac{J^2}{r_\mathrm{s}^2}\,, \tag{D.95}$$

meaning that the photons do not fall into the event horizon provided (see figure D.1)

$$J \geq \frac{3\sqrt{3}}{2}cr_\mathrm{s} = 3\sqrt{3}\frac{GM}{c}\,. \tag{D.96}$$

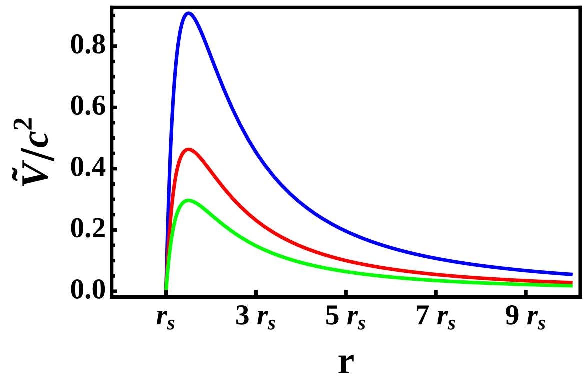


Figure D.1: Here we plot the effective potential $\widetilde{V}(r)/c^2$, as found in (D.93) for $J = 3.5\,c\,r_\mathrm{s}$ (blue), $J = 2.5\,c\,r_\mathrm{s}$ (red) and $J = 2\,c\,r_\mathrm{s}$ (green). We note that for the latter two cases, the effective potential is less than the classical kinetic energy ($\frac{1}{2}\,c^2$), and the photons will fall into the Schwarzschild radius. This agrees with (D.96).

i.e.

$$\widetilde{V}\left(\frac{3}{2}r_\mathrm{s}\right) \geq c^2\,. \tag{D.97}$$

The impact parameter is given as $b = J/c$, thus in terms of the impact parameter, photons do not fall into the Schwarzschild radius unless

$$b \geq \frac{3\sqrt{3}}{2}r_\mathrm{s} = 3\sqrt{3}\frac{GM}{c^2}\,. \tag{D.98}$$

D.6.1. Deflection of Light in a Gravitational Field. Here we are considering photons which escape to infinity, i.e. $b > \frac{3\sqrt{3}}{2}r_\mathrm{s}$. (D.86) gives us

$$\frac{d\varphi}{dt} = \frac{JB}{r^2} \quad \Rightarrow \quad \frac{d\varphi}{d\tau} = \frac{J}{r^2} \quad \Rightarrow \quad \frac{d\tau}{d\varphi} = \frac{r^2}{J}\,. \tag{D.99}$$

We can now multiply (D.92) by $\left(\frac{d\tau}{d\varphi}\right)^2$

$$\frac{1}{2}\left(\frac{dr}{d\tau}\right)^2\left(\frac{d\tau}{d\varphi}\right)^2 + \frac{1}{2}\frac{r-r_\mathrm{s}}{r}\frac{J^2}{r^2}\left(\frac{d\tau}{d\varphi}\right)^2 = \frac{c^2}{2}\left(\frac{d\tau}{d\varphi}\right)^2,$$

We now simplify and rearrange to find

$$\left(\frac{dr}{d\varphi}\right)^2 = -\frac{r-r_\mathrm{s}}{r}\frac{J^2}{r^2}\frac{r^4}{J^2} + c^2\frac{r^4}{J^2} = -r^2 + r_\mathrm{s}r + \frac{c^2}{J^2}r^4 \,.$$

Taking the square root, and rearranging further we find

$$d\varphi = \pm dr \frac{J}{\sqrt{c^2r^4 + J^2 r_\mathrm{s} r - J^2 r^2}} \,. \tag{D.100}$$

Integration leads to

$$\varphi - \varphi_0 = \pm \int_{r_0}^{r} dr \frac{J}{\sqrt{r(c^2r^3 - J^2 r + J^2 r_\mathrm{s})}} = \pm \int_{r_0}^{r} dr \frac{b}{\sqrt{r(r^3 - b^2 r + b^2 r_\mathrm{s})}} \tag{D.101}$$

We now set the origin at the center of the mass M. Since $\theta = \frac{\pi}{2}$ we know the photon is traveling in the xy-plane. We choose the x-axis such that the photon is falling in from $x \to -\infty$ along $y = -b$. In which case $\varphi_0 = -\pi$ and $r_0 \to \infty$. The particle will fall to the minimum value of r, r_1. We then have

$$\varphi(r_1) = -\pi - \int_{\infty}^{r_1} dr \frac{b}{\sqrt{r(r^3 - b^2 r + b^2 r_\mathrm{s})}} \,. \tag{D.102}$$

The deflection angle of the photon is given by (r goes back out to ∞)

$$\begin{aligned}\Delta\varphi &= \varphi(r \to \infty) = \varphi(r_1) + \int_{r_1}^{\infty} dr \frac{b}{\sqrt{r(r^3 - b^2 r + b^2 r_\mathrm{s})}} \\ &= -\pi - 2\int_{\infty}^{r_1} dr \frac{b}{\sqrt{r(r^3 - b^2 r + b^2 r_\mathrm{s}}} \,.\end{aligned} \tag{D.103}$$

We now need to know what r_1 is. Since, by definition, it is the minimum value of r, $\frac{dr}{d\varphi}|_{r=r_1} = 0$, i.e.

$$\sqrt{r_1(r_1^3 - b^2 r_1 + b^2 r_\mathrm{s})} = 0 \quad \Rightarrow \quad (r_1 - r_\mathrm{s})b^2 = r_1^3 \,. \tag{D.104}$$

Clearly if $r_\mathrm{s} = 0$ then $r_1 = b$, because there is no gravitational attraction. As such we can propose that $r = b + \epsilon$ in the case that $r_\mathrm{s} \ll b$, then to the first order in epsilon

$$\begin{aligned}(b+\epsilon-r_\mathrm{s})b^2 = (b+\epsilon)^3 &\Rightarrow (b+\epsilon-r_\mathrm{s})b^2 = b^3 + 3b^2\epsilon + \cancel{3b\epsilon^2} + \cancel{\epsilon^3} \\ &\Rightarrow b+\epsilon-r_\mathrm{s} = b + 3\epsilon \\ &\Rightarrow \epsilon = -\frac{r_\mathrm{s}}{2}\end{aligned} \tag{D.105}$$

We now expand the $\Delta\varphi$ to the first order in r_2, yielding

$$\Delta\varphi \approx -\pi - 2\int_\infty^{b-\frac{r_\mathrm{s}}{2}} dr \frac{b}{r\sqrt{r^2-b^2}} + \int_\infty^{b-\frac{r_\mathrm{s}}{2}} dr \frac{b^3 r_\mathrm{s}}{r^2\sqrt{r^2-b^2}^3}\,. \tag{D.106}$$

Using the identity

$$\int^{f(x+\epsilon)} d\xi I(\xi) \approx \int^{f(x)} d\xi I(\xi) + \epsilon f'(x) I(f(x)) \tag{D.107}$$

we find

$$\begin{aligned}\Delta\varphi \approx & -\pi - 2\int_\infty^{b} dr \frac{b}{r\sqrt{r^2-b^2}} + \lim_{r\to b}\frac{r_\mathrm{s}}{\sqrt{r^2-b^2}} \\ & + \int_\infty^{b} dr \frac{b^3 r_\mathrm{s}}{r^2\sqrt{r^2-b^2}^3} - \cancel{\lim_{r\to b}\frac{r_\mathrm{s}}{2}\frac{br_\mathrm{s}}{\sqrt{r^2-b^2}^3}} \\ = & -\pi - 2\sin^{-1}\left(-\frac{b}{r}\right)\Bigg|_\infty^b + \lim_{r\to b}\frac{r_\mathrm{s}}{\sqrt{r^2-b^2}} - r_\mathrm{s}\frac{2r^2-b^2}{br\sqrt{r^2-b^2}}\Bigg|_\infty^b \\ = & -\pi + \pi + \lim_{r\to b}\frac{r_\mathrm{s}}{\sqrt{r^2-b^2}} - \lim_{r\to b}\frac{r_\mathrm{s}}{\sqrt{r^2-b^2}} + \frac{2r_\mathrm{s}}{b} = \frac{2r_\mathrm{s}}{b}\,.\end{aligned} \tag{D.108}$$

Thus the deflection angle of a photon with impact parameter b in a gravitational field is

$$\Delta\varphi \approx \frac{2r_\mathrm{s}}{b} = \frac{4GM}{c^2 b}\,. \tag{D.109}$$

D.7. ISOTROPIC SCHWARZSCHILD METRIC

The Schwarzschild metric is derived as (D.54)

$$ds^2 = -\frac{r - r_\mathrm{s}}{r} dt^2 + \frac{r}{r - r_\mathrm{s}} dr^2 + r^2 d\Omega^2 \,, \quad d\Omega^2 \equiv d\theta^2 + \sin^2\theta d\varphi^2 \tag{D.110}$$

where r_s is the Schwarzschild radius, and as such is positive. Clearly this metric runs into difficulties at the event horizon, i.e., $r = r_\mathrm{s}$. Inside a black hole the metric becomes

$$ds^2 = \frac{r_\mathrm{s} - r}{r} dt^2 - \frac{r}{r_\mathrm{s} - r} dr^2 + r^2 d\Omega^2 \,, \tag{D.111}$$

and our time coordinate t becomes spacelike while the space coordinate r becomes timelike. This singularity and change in behavior of the t and r coordinates provide challenges when crossing the event horizon and is dealt with using Kruskal-Szekeres coordinates (see Chapter 6.4 of [126], and page 97–102 of [128]). Another example of a coordinate transformation is the transformation into isotropic coordinates which is performed by setting

$$r = r_1 \left(1 + \frac{r_\mathrm{s}}{4r_1}\right)^2 = r_1 f^2 \,, \tag{D.112}$$

as illustrated in figure D.1, from which we find

$$\begin{aligned} dr &= dr_1 \left(1 + \frac{r_\mathrm{s}}{4r_1}\right)^2 + 2\cancel{r_1} \left(1 + \frac{r_\mathrm{s}}{4r_1}\right) \left(-\frac{r_\mathrm{s}}{4r_1^{\cancel{2}}}\right) dr_1 \\ &= \left(1 + \frac{r_\mathrm{s}}{4r_1} - \frac{r_\mathrm{s}}{2r_1}\right) \left(1 + \frac{r_\mathrm{s}}{4r_1}\right) dr_1 \\ &= \left(1 - \frac{r_\mathrm{s}}{4r_1}\right) \left(1 + \frac{r_\mathrm{s}}{4r_1}\right) dr_1 = g f dr_1 \,, \end{aligned} \tag{D.113}$$

$$\begin{aligned} r - r_\mathrm{s} &= r_1 \left(1 + \frac{r_\mathrm{s}}{4r_1}\right)^2 - r_2 = r_1 + \frac{r_\mathrm{s}}{2} + \frac{r_\mathrm{s}^2}{16 r_1} - r_\mathrm{s} = r_1 \left(1 - \frac{r_\mathrm{s}}{2} + \left(\frac{r_\mathrm{s}}{4r_1}\right)^2\right) \\ &= r_1 \left(1 - \frac{r_\mathrm{s}}{4r_1}\right)^2 = r_1 g^2 \,, \end{aligned} \tag{D.114}$$

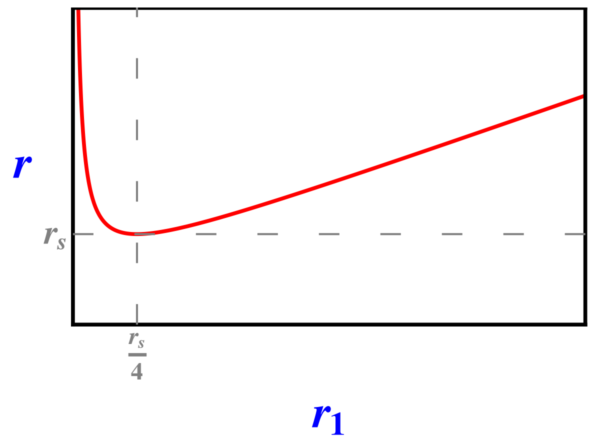


Figure D.2: Here we plot r as a function of r_1, from which it is evident that r covers all the possible values outside the black hole, while the values for r inside the black hole are absent.

where

$$f = f(r_1) = 1 + \frac{r_\mathrm{s}}{4r_1}\,, \quad g = g(r_1) = 1 - \frac{r_\mathrm{s}}{4r_1}\,. \tag{D.115}$$

We can then plug these into the Schwarzschild metric

$$\begin{aligned} ds^2 &= -\frac{r - r_\mathrm{s}}{r}c^2dt^2 + \frac{r}{r - r_\mathrm{s}}dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2) \\ &= -\frac{\cancel{r_1} g^2}{\cancel{r_1} f^2}c^2dt^2 + \frac{\cancel{r_1} f^2}{\cancel{r_1} \cancel{g^2}}\cancel{g^2} f^2 dr_1^2 + f^4 r_1^2(d\theta^2 + \sin^2\theta d\varphi^2) \\ &= -\frac{g^2}{f^2}c^2dt^2 + f^4(dr_1^2 + r_1^2 d\theta^2 + r_1^2\sin^2\theta d\varphi^2) = -\frac{g^2}{f^2}c^2dt^2 + f^4(dx^2 + dy^2 + dz^2) \end{aligned}$$

i.e.,

$$ds^2 = -\frac{\left(1 - \frac{r_\mathrm{s}}{4r_1}\right)^2}{\left(1 + \frac{r_\mathrm{s}}{4r_1}\right)^2}c^2dt^2 + \left(1 + \frac{r_\mathrm{s}}{4r_1}\right)^4 (dr_1^2 + r_1^2 d\Omega^2)\,, \tag{D.116}$$

where

$$r = r_1\left(1 + \frac{r_\mathrm{s}}{4r_1}\right)^2\,, \quad r_1 = \frac{r}{2} - \frac{r_\mathrm{s}}{4} + \frac{1}{2}\sqrt{r(r - r_\mathrm{s})}\,. \tag{D.117}$$

This is a commonly used transform, and is generally used without a hint of the difficulties that arise from using such a transform. Not only is the spatial coordinate transformed, but the time coordinate t is implicitly transformed as well, albeit as

$t = t_1$. Let us consider the possible values of r implied by this transformation. As r_1 goes from $\frac{r_s}{4}$ to ∞, r goes from the horizon to ∞, as we might expect. However when r_1 goes from 0 to $\frac{r_s}{4}$, r goes from ∞ to r_s. Thus the possible values for r *outside* the black hole are covered twice, while the values for r *inside* the black hole are absent, see figure D.2. Furthermore, when solving for r_1 we find that there is a minimum value for r, below which r_1 becomes complex and no longer physical.

BIBLIOGRAPHY

[1] W. Pauli, *On Dirac's New Method of Field Quantization*, Rev. Mod. Phys. **15**, 175 (1943).

[2] C. M. Bender and S. Boettcher, *Real Spectra in Non–Hermitian Hamiltonians Having PT Symmetry*, Phys. Rev. Lett. **80**, 5243 (1998).

[3] A. Mostafazadeh, *Pseudo-Hermiticity versus PT-Symmetry: The necessary condition for the reality of the spectrum of a non-Hermitian Hamiltonian*, J. Math. Phys. **43**, 205 (2002).

[4] A. Mostafazadeh, *Pseudo–Hermiticity and Generalized PT- and CPT-Symmetries*, J. Math. Phys. **44**, 974 (2003).

[5] U. D. Jentschura, A. Surzhykov, and J. Zinn-Justin, *Multi-Instantons and Exact Results III: Unification of Even and Odd Anharmonic Oscillators*, Ann. Phys. (N.Y.) **325**, 1135 (2010).

[6] C. M. Bender, M. V. Berry, and A. Mandilara, *Generalized PT Symmetry and Real Spectra*, J. Phys. A **35**, L467 (2002).

[7] C. M. Bender, K. A. Milton, and V. M. Savage, *Solution of Schwinger-Dyson Equations for PT-Symmetric Quantum Field Theory*, Phys. Rev. D **62**, 085001 (2000).

[8] C. M. Bender, S. Boettcher, and P. N. Meisinger, *PT-symmetric quantum mechanics*, J. Math. Phys. **40**, 2201 (1999).

[9] C. M. Bender and G. V. Dunne, *Large-Order Perturbation Theory for a Non-Hermitian PT-Symmetric Hamiltonian*, J. Math. Phys. **40**, 4616 (1999).

[10] C. M. Bender and E. J. Weniger, *Numerical Evidence that the Perturbation Expansion for a Non-Hermitian PT-symmetric Hamiltonian is Stieltjes*, J. Math. Phys. **42**, 2167 (2001).

[11] C. M. Bender, D. C. Brody, and H. F. Jones, *Complex Extension of Quantum Mechanics*, Phys. Rev. Lett. **89**, 270401 (2002).

[12] U. D. Jentschura, A. Surzhykov, and J. Zinn-Justin, *Unified Treatment of Even and Odd Anharmonic Oscillators of Arbitrary Degree*, Phys. Rev. Lett. **102**, 011601 (2009).

[13] A. Mostafazadeh, *Pseudo-Hermiticity versus PT-Symmetry II: A complete characterization of non-Hermitian Hamiltonians with a real spectrum*, J. Math. Phys. **43**, 2814 (2002).

[14] A. Mostafazadeh, *Pseudo-Hermiticity versus PT-Symmetry III: Equivalence of pseudo-Hermiticity and the presence of antilinear symmetries*, J. Math. Phys. **43**, 3944 (2002).

[15] C. M. Bender, *Making sense of non–Hermitian Hamiltonians*, Rep. Prog. Phys. **70**, 947 (2007).

[16] P. Dorey, C. Dunning and R. Tateo, *Spectral equivalences, Bethe ansatz equations, and reality properties in PT–symmetric quantum mechanics*, J. Phys. A **34**, 5679 (2001).

[17] P. Dorey, C. Dunning and R. Tateo, *A reality proof in PT–symmetric quantum mechanics*, Czech J. Phys. **54**, 35 (2003).

[18] E. Caliceti and S. Graffi, *On a class on non selfadjoint quantum non–linear oscillators with a real spectrum*, J. Nonliniar Math. Phys. **12** Supp. 1, 138 (2005).

[19] A. Mostafazadeh, *Exact PT–symmetry is equivalent to Hermiticity*, J. Phys. A **36**, 7081 (2003).

[20] Petr Siegl, *The non–equivalence of pseudo–Hermiticity and presence of antilinear symmetry*, Pramana **73**, 279 (2009).

[21] P. Siegl and D. Krejčiŗík, *On the metric operator for the imaginary cubic oscillator*, Phys. Rev. D **86**, 121702 (2012).

[22] L. L. Foldy and S. A. Wouthuysen, *On the Dirac Theory of Spin 1/2 Particles and Its Non–Relativistic Limit*, Phys. Rev., **78**, 29 (1950).

[23] J.D. Bjorken and S. Drell, *Relativistic Quantum Mechanics*, (McGraw–Hill, New York, 1964).

[24] E. Eriksen and M. Kolsrud, *Canonical transformations of Dirac's equation to even forms*, Suppl. Nuovo Cimento **18**, 1 (1958).

[25] A. Chodos, A. I. Hauser, and V. A. Kostelecky, **The neutrino as a tachyon**, Phys. Lett. B **150**, 431 (1985).

[26] O. M. P. Bilaniuk, V. K. Beshpande and E. C. G. Sudarshan, *"Meta" Relativity*, Am. J. Phys. **30**, 718 (1962).

[27] G. Feinberg, *Possibility of Faster–Than–Light Particles*, Phys. Rev. **159**, 1089 (1967).

[28] M. E. Arons and E. C. G. Sudarshan, *Lorentz Invariance, Local Field Theory, and Faster–than–Light Particles* Phys. Rev. **173**, 1622 (1968).

[29] J. Dhar and E. C. G. Sudarshan, *Quantum Field Theory of Interacting Tachyons*, Phys. Rev. **174**, 1808 (1968).

[30] E. Recami and R. Mignani, *Classical Theory of Tachyons (Special Relativity Extended to Superluminal Frames and Objects)*, Rivista Nuovo Cimento **4**, 209 (1974).

[31] C. M. Bender, P. N. Mesinger, and Q. Wang, *Calculation of the Hidden Symmetry Operator in PT-Symmetric Quantum Mechanics*, J. Phys. A **36** 1973 (2003).

[32] C. M. Bender, D. C. Brody, and H. F. Jones, *Extension of PT-Symmetric Quantum Mechanics to Quantum Field Theory with Cubic Interaction*, Phys.Rev. D **70**, 025001 (2004).

[33] H. F. Jones, *Use of equivalent Hermitian Hamiltonian for PT–symmetric sinusoidal optical lattices*, J. Phys. A **44** 345302 (2011).

[34] A. Mostafazadeh, *Pseudo-Hermitian Representation of Quantum Mechanics*, Int. J. Geom. Methods Mod. Phys.**7**, 1191 (2010).

[35] C. M. Bender, S. F. Brandt, J. H. Chen, and Q. Wang, *Ghost Busting: PT-Symmetric Interpretation of the Lee Model*, Phys. Rev. D **71**, 025014 (2008).

[36] C. M. Bender, P. Meisinger and H. Yang, *Calculation of the one-point Greens function for a $-g\varphi^4$ quantum field theory*, Phys. Rev. D **74**, 45001 (2001).

[37] H. F. Jones, J. Mateo and R. J. Rivers, *Path-integral derivation of the anomaly for the Hermitian equivalent of the complex PT-symmetric quartic Hamiltonian*, Phys. Rev. D **74**, 125002 (2006).

[38] A. A. Andrianov, F. Cannata, A. Y. Kamenshchik, and D. Regoli, *Phantom cosmology based on PT symmetry*, Int. J. Mod. Phys. D **10**, 97 (2010).

[39] L. Chen, R. Li, N. Yang, D. Chen, and L. Li, *Optical Modes in PT-Symmetric Double-Channel Waveguides*, P. Romanian Acad. A **13**, 46 (2012).

[40] A. Szameit, M. C. Rechtsman, O. Bahat-Treidel, and M. Segev, *PT–Symmetry in honeycomb photonic lattices*, Phys. Rev. A **84**, 021806 (2011).

[41] H. Ramezani, T. Kottos, V. Kovanis and D. N. Christodoulides, *Exceptional–point dynamics in photonic honeycomb lattieces with PT symmetry*, Phys. Rev. A **85**, 013818 (2012).

[42] A. Regensburger, C. Bersch, M.-A. Miri, G. Onishchukov, D. N. Christodoulides and U. Peschel, *Parity–time synthetic photonic lattices*, Nature **488**, 167 (2012).

[43] U. D. Jentschura and B. J. Wundt, *Pseudo-Hermitian Quantum Dynamics of Tachyonic Spin-1/2 Particles*, J. Phys. A **45**, 444017 (2012).

[44] U. D. Jentschura and B. J. Wundt, *From Generalized Dirac Equations to a Candidate for Dark Energy*, ISRN High Energy Physics 2013, 374612 (2013).

[45] R. Ehrlich, *Tachyonic neutrinos and the neutrino masses*, Astropart. Phys. **41**, 1 (2013).

[46] J. Zinn-Justin and U. D. Jentschura, *Multi-Instantons and Exact Results I: Conjectures, WKB Expansions, and Instanton Interactions*, Ann. Phys. (N.Y.) **313**, 197 (2004).

[47] J. Zinn-Justin and U. D. Jentschura, *Multi-Instantons and Exact Results II: Specific Cases, Higher-Order Effects, and Numerical Calculations* Ann. Phys. (N.Y.) **313**, 269 (2004).

[48] N. Moiseyev, *Quantum theory of resonances: calculating energies, widths and cross-sections by complex scaling*, Phys. Rep. **302**, 211 (1998).

[49] A. Mostafazadeh, *Non-Hermitian Hamiltonians with a Real Spectrum and Their Physical Applications*, Pramana-J. Phys. **73**, 269 (2009).

[50] C. M. Bender, D. C. Brody, and H. F. Jones, *Scalar Quantum Field Theory with Complex Cubic Interaction*, Phys. Rec. Lett. **93**, 251601 (2004).

[51] C. M. Bender and T. T. Wu, *Large-Order Behavior of Perturbation Theory*, Phys. Rev. Lett. **27**, 461 (1971).

[52] C. M. Bender and T. T. Wu, *Anharmonic Oscillator. II. A Study of Perturbation Theory in Large Order*, Phys. Rev. D **7**, 1620 (1973).

[53] C. M. Bender and T. T. Wu, *Anharmonic Oscillator*, Phys. Rev. **184**, 1231 (1969).

[54] J. H. Noble, M. Lubasch, and U. D. Jentschura, *Generalized Householder Transformations for the Complex Symmetric Eigenvalue Problem*, Eur. Phys. J. Plus **128**, 93 (2013).

[55] A. Messiah, *Quantum Mechanics II* (Dover, North-Holland, Amsterdam, 1962).

[56] U. D. Jentschura, A. Surzhykov, M. Lubasch, and J. Zinn-Justin, *Structure, Time Propagation and Dissipative Terms for Resonances*, J. Phys. A **41**, 095302 (2008).

[57] J. Zinn-Justin and U. D. Jentschura, *Imaginary Cubic Perturbation: Numerical and Analytic Study*, J. Phys. A **43**, 425301 (2010).

[58] C. M. Bender, J. Brod, A. Refig, and M. E. Reuter, *The C Operator in PT-Symmetric Quantum Theories*, J. Phys. A **37**, 10139 (2004).

[59] A. Messiah, *Quantum Mechanics I* (Dover, North–Holland, Amsterdam, 1962).

[60] C. Itzykson and J. B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980).

[61] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, D. Sorensen, `LAPACK` Users' Guide, 3rd Edition (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999).

[62] W. E. Arnoldi, *The principle of minimized iterations in the solution of the matrix eigenvalue problem*, Quart. Appl. Math. **9**, 17 (1951).

[63] C. M. Bender and D. J. Weir, *PT phase transitions in multidimensional quantum systems*, J. Phys. A **45**, 425303 (2012).

[64] J. G. F. Francis, *The QR Transformation, I*, Comput. J. **4**, 265 (1961).

[65] J. G. F. Francis, *The QR Transformation, II*, Comput. J. **4**, 332 (1962).

[66] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, (Oxford University Press, Oxford, UK, 1965).

[67] J. H. Wilkinson, *Global Convergence of Tridiagonal QR Algorithm With Origin Shifts*, Lin. Alg. Applic. **1** 409 (1968).

[68] G. Golub, and C. F. van Loan, *Matrix Computations, 3rd Edition*, (The Johns Hopkins University Press, Baltimore, MD, 1996).

[69] B. N. Parlett, *The Symmetric Eigenvalue Problem*, (Prentice Hall, Englewood Cliffs,NJ, 1998).

[70] A. S. Householder, *Unitary Triangularization of a Nonsymmetric Matrix*, J. ACM **5**, 339 (1958).

[71] J. K. Cullum, R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations, Volume I: Theory*, (Birkhäuser, Boston, 1985).

[72] J. K. Cullum, R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations, Volume II: Programs*, (Birkhäuser, Boston, 1985).

[73] H. Schabauer, C. Pacher, A. G. Sunderland, W. N. Gansterer, *Toward a parallel solver for generalized complex symmetric eigenvalue problems*, Procedia Computer Science **1**, 437 (2012).

[74] I. Bar-On, V. Ryaboy, *Non-Hermitian electronic theory and applications to clusters*, SIAM J. Sci. Comput. **18**, 1412 (1997).

[75] J. K. Cullum, R. A. Willoughby, *A QL Procedure for Computing the Eigenvalues of Complex Symmetric Tridiagonal Matrices*, SIAM J. Math. Anal. **17** 83 (1996).

[76] Z. Bai, J. Demmel, *On a Block Implementation of Hessenberg Multishift QR Iteration*, Int. J. High Speed Comput. **1** 97 (1989).

[77] W. Givens, *Computaion of Plane Unitary Rotations Transforming a Generl Matrix to Triangular Form*, J. Soc. Indust. Appl. Math. **6**, 26 (1958).

[78] A. S. Householder, *On Deflating Matrices*, J. Soc. Indust. Appl. Math. **9** 89 (1961).

[79] G. Fasshauer, *477/577 Numerical Linear Algebra/Computational Mathematics I*, Unpublished Manuscript, Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL.

[80] P. Arbenz, *Numerical Methods for Solving Large Scale Eigenvalue Problems*, Unpublished Manuscript, Computer Science Department, Eidgenössische Technische Hochschule Zürich, Zürich, Switzerland.

[81] L. E. Henderson, *PhD Dissertation: Testing eigenvalue software* (University of Arizona, 1991, unpublished).

[82] D. H. Bailey, *A portable high performance multiprecision package*, NASA Ames Tech. Rep. RNR-90-022.

[83] D. H. Bailey, *Multiprecision translation and execution of Fortran programs*, ACM Trans. Math. Soft. **19** 288 (1993).

[84] D. H. Bailey, *A Fortran-90 based multiprecision system*, NASA Ames Tech. Rep. RNR-94-013, published as ACM Trans. on Mathematical Software **21**, 379 (1995).

[85] D. H. Bailey, *A Fortran-90 based multiprecision system*, ACM Trans. Math. Soft. **21** 379 (1995).

[86] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic Press, New York, 1957).

[87] U. D. Jentschura, and J. H. Noble, *Nonrelativistic Limit of the Dirac-Schwarzschild Hamiltonian: Gravitational Zitterbewegung and Gravitational Spin–Orbit Coupling*, Phys. Rev. A **88** 022121 (2013).

[88] U. D. Jentschura and J. H. Noble, *Foldy-Wouthuysen Transformation, Scalar Potentials and Gravity* J. Phys. A **47**, 045402 (2014).

[89] J. H. Noble and U. D. Jentschura, *Dirac equations with confining potentials*, Int. J. Mod. Phys. A **30**, 1550002 (2015).

[90] M. Bojowald, *Canonical Gravity and Applications* (Cambridge University, Cambridge, England, 2011).

[91] D. R. Brill and J. A. Wheeler, *Interaction of Neutrinos and Gravitaional Fields*, Rev. Mod. Phys. **29**, 465 (1957).

[92] U. D. Jentschura, *Gravitationally coupled Dirac equation for antimatter*, Phys. Rev. A **87**, 032101 (2013), Erratum Phys. Rev. A **87**, 069903 (2013).

[93] J. H. Noble, and U. D. Jentschura, *Ultrarelativistic decoupling transformation for generalized Dirac equations*, Phys. Rev. A **92**, 012101 (2015).

[94] Y. N. Obukhov, *Spin, Gravity, and Inertia*, Phys. Rev. Lett., **86**, 192 (2001).

[95] D. G. Boulware, *Spin–$\frac{1}{2}$ quantum field theory is Schwarzschild space*, Phys. Rev. D **12**, 305 (1975).

[96] M. Soffel, B. Müller and W. Greiner, *Particles in a stationary spherically symmetric gravitational field*, J. Phys. A **10**, 551 (1977).

[97] V. M. Red'kov, *Ricci Coefficients in Covariant Dirac Equation, Symmetry Aspects and Newman–Penrose Approach*, Nonlinear Phenom. Complex Syst. **7**, 250 (2011).

[98] J. Yepez, *Einstein's vierbein field theory of curved space*, arXiv:1106.2037 (2011).

[99] Y. N. Obukhov, *On gravitational interaction of fermions*, Fortsch. Phys. **50**, 711 (2002).

[100] A. S. Eddington, *The Mathematical Theory of Relativity* (Cambridge University Press, Cambridge, England, 1924).

[101] A. J. Silenko and O. V. Teryaev, *Semiclassical limit for Dirac particles interacting with a gravitational field*, Phys. Rev. D **71**, 064016 (2005).

[102] F. W. Hehl and W.-T. Ni, *Inertial effects of a Dirac particle*, Phys. Rev. D **42**, 2045 (1990).

[103] B. Mashhoon, *Neutron interferometry in a rotating frame of reference*, Phys. Rev. Lett. **61**, 2639 (1988).

[104] B. Mashhoon, *Electrodynamis in a rotating frame of reference*, Phys. Lett. A **139**, 103 (1989).

[105] K. Pachucki, *Higher–Order Binding Correction to the Lamb Shift*, Ann. Phys. (N.Y.) **226**, 1 (1993).

[106] U. Jentschura and K. Pachucki, *Higher–Order Binding Corrections to the Lamb Shift of 2P States*, Phys. Rev. A **54**, 1853 (1996).

[107] J. Franklin, *A Simple Dirac Wave Function for a Coulomb Potential with Linear Confinement*, Mod. Phys. Lett. A **14**, 2409 (1999).

[108] A. Soares de Castro and J. Franklin, *Exact Solutions of the Dirac Equation for Modified Coulombic Potentials*, Int. J. Mod. Phys. A **15**, 4355 (2000).

[109] U. D. Jentschura, *Master Theis: The Lamb Shift in Hydrogenlike Systems*, (University of Munich, 1996, unpublished (see e–print hep-ph/0305065)).

[110] B. J. W. Wundt, *PhD Dissertation: Quantum Electrodynamics and Fundamental Constants*, (Missouri University of Science and Technology, 2011, unpublished).

[111] U. D. Jentschura and K. Pachucki, Two–loop self–energy corrections to the fine structure, J. Phys. A: Math. Gen. **35**, 1927 (2002).

[112] E. J. Weniger, *Nonlinear sequence transformations for the acceleration of convergence and the summation of divergent series*, Comput. Phys. Rep. **10**, 189 (1989).

[113] U. D. Jentschura, P. J. Mohr and G. Soff, *Calculation of the Electron Self-Energy for Low Nuclear Charge*, Phys. Rev. Lett. 82, 53 (1999).

[114] U. D. Jentschura, P. J. Mohr and G. Soff, *Electron Self Energy for K and L Shells at Low Nuclear Charge*, Phys. Rev. A **63**, 042512 (2001).

[115] A. D. Fokker, *The Geodesic Precession: A Consequence of Einsteins Theory of Gravitation*, Kon. Akad. Weten. Amsterdam. Proc. **23**, 729 (1921).

[116] R. A. Swainson and G. W. F. Drake, *A unified treatment of the non-relativistic and relativistic hydrogen atom I: The wavefunctions*, J. Phys. A **24**, 79 (1991).

[117] R. A. Swainson and G. W. F. Drake, *A unified treatment of the non-relativistic and relativistic hydrogen atom II: the Green functions*, J. Phys. A **24**, 95 (1991).

[118] R. A. Swainson and G. W. F. Drake, *A unified treatment of the non-relativistic and relativistic hydrogen atom: III. The reduced Green functions*, J. Phys. A **24**, 1801 (1991).

[119] C. W. F. Everitt, D. B. DeBra, B. W. Parkinson, J. P. Turneaure, J. W. Conklin, M. I. Heifetz, G. M. Keiser, A. S. Silbergleit, T. Holmes, J. Kolodziejczak, M. Al-Meshari, J. C. Mester, B. Muhlfelder, V. Solomonik, K. Stahl, P. Worden, W. Bencze, S. Buchman, B. Clarke, A. Al-Jadaan, H. Al-Jibreen, J. Li, J. A. Lipa, J. M. Lockhart, B. Al-Suwaidan, M. Taber and S. Wang, *Gravity Probe B: Final Results of a Space Experiment to Test General Relativity*, Phys. Rev. Lett. **106**, 221101 (2011).

[120] D. Lovelock and H. Rund, *Tensors, Differential Forms, and Variational Principles* (John Wiley & Sons, New York, 1975).

[121] V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics, Volume 4 of the Course on Theoretical Physics, 2nd ed.* (Pergamon, Oxford, 1982).

[122] M. Srednicki, *Quantum Field Theory* (Cambridge University, Cambridge, England, 2007).

[123] M. Cinin and B. Touschek, *The relativistic limit of the theory of spin 1/2 particles*, Nuovo Cimento **7**, 422 (1958).

[124] A. Sankaranaranayan, *On the Transformation of Relativistic Single–Particle Hamiltonians to Odd & Even Forms*, Nuovo Cimento **32**, 1715 (1963).

[125] R. Parthasarathy, *Relativistic Quantum Mechanics* (Alpha Science, Oxford, UK 2009).

[126] R. M. Wald, *General Relativity*, (The University of Chicago Press, Chicago, 1984).

[127] O. W. Greenburg, *CPT Violation Implies Violation of Lorentz Invariance*, Phys. Rec. Lett. **89**, 231602 (2002).

[128] R. Dick, *Astronomy 411.3, Gravitation & Cosmology*, 5th ed., Unpublished manuscript, Department of Physics & Engineering Physics, University of Saskatchewan, Saskatoon, Canada (2012).

[129] D. A. Clark, *A Primer on Tensor Calculus*, Unpublished manuscript, Department of Physics, Saint Mary's University, Halifax NS, Canada (2011).

[130] K. Schwarzschild, *Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Teorie*, Sitz. K. Preuss. Akad. Wiss. 7, 189 (1916).
English Translation: Gen. Rel. Grav. **35**, 1951 (2003).

[131] K. Schulten, *Notes on Quantum Mechanics*, Unpublished manuscript, Department of Physics and Beckman Institute, University of Illinois at Urbana–Champaign, Urbana, IL, USA.

[132] D. Borisov, and D. Krejčiřík, *PT–Symmetric Waveguides*, Int. Eqns Op. Th. **62**, 489 (2008).

[133] C. M. Bender, P. N. Meisinger, and Q. Wang *Calculation of the Hidden Symmetry Operator in PT-Symmetric Quantum Mechanics*, arXiv:quant-ph/0211166v1 (2002).

[134] H. Weyl, *Elektron und Gravitation I*, Z. Phys. **56**, 330 (1929).

[135] A. Fock, *Geometrisierung der Diracschen Theorie des Elektrons*, Z. Phys. **57**, 261 (1929).

VITA

Jonathan Howard Noble was born on September 2nd, 1986 in Arlington, VA, USA. Due to his father's work, he grew up in Honduras, Pakistan, Botswana, and the Republic of Georgia. In 2001, he moved to Ijamsville, MD, where he attended Urbana High school, and graduated with highest honors in 2005. Jonathan then attended Saint Mary's College of Maryland, and graduated cum laude with a B.A. in Mathematics and Physics in 2009. He spent the next year working in an experimental research lab, until moving to Rolla, MO, in 2010 to attend the Missouri University of Science and Technology, where he received a M.S. in Physics in 2014, followed by a Ph.D. in Physics in 2015.