NON-HERMITIAN QUANTUM MECHANICS

by

KATHERINE JONES-SMITH

Submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

Thesis Advisor: Dr. Harsh Mathur

Department of Physics

CASE WESTERN RESERVE UNIVERSITY

May, 2010

CASE WESTERN RESERVE UNIVERSITY
SCHOOL OF GRADUATE STUDIES

We hereby approve the thesis/dissertation of
Katherine Jones-Smith,
candidate for the Ph.D degree.

(signed) Harsh Mathur (chair of the committee)
Idit Zehavi
Tanmay Vachaspati
Phil Taylor

3/31/10

We also certify that written approval has been obtained for any proprietary material contained therein.

To Fema
The Far Flung Correspondent

Contents

List of Figures

Non-Hermitian Quantum Mechanics

Abstract

by

KATHERINE JONES-SMITH

The basic structure of quantum mechanics was delineated in the early days of the theory and has not been modified since. One of the fundamental assumptions used in formulating the theory is that operators are represented by Hermitian matrices. In recent years it has been shown that quantum mechanics can be formulated consistently without making this assumption, using instead a combination of the parity (P) and time-reversal (T) operators and a number of other requirements related to P and T. Only the case of even T has been analyzed in the literature; here we generalize the principles to include odd time-reversal. We use this generalization to construct a non-Hermitian version of the Dirac equation, and in doing so discover a new type of particle not allowed within the (Hermitian) Standard Model. Finally we present a potential application of the ideas of non-Hermitian quantum mechanics to the unsolved problems of quantum magnetism and high temperature superconductivity.

Chapter 1

Introduction

Quantum mechanics originated as a set of ad hoc rules that attempted to explain a number of seemingly unrelated experimental results. These ad hoc rules amalgamated into the basic principles of quantum mechanics that were delineated in the 1930's and have not been modified since [1]. Still, it is desirable to ask whether the structure can be altered and generalized. For example Weinberg showed that it is possible to formulate a non-linear generalization of quantum mechanics and to thereby subject the linearity of quantum mechanics to a quantitative test [2]. A fruitful generalization of the canonical principles, was the discovery that particles can have fractional statistics that interpolate between Bose and Fermi, albeit only in two spatial dimensions [3]. More recently the principle that the Hamiltonian and other observables should be represented by Hermitian operators has been re-examined [4].

When most physicists hear the term 'Hermitian', they think of a matrix which is equal to its complex conjugate transpose. But in this context Hermiticity is just short for self-adjoint. An operator A that is self adjoint has the very reasonable property that its effect on the vectors of the Hilbert space in which it is defined is independent of what vector it acted on first. Using the standard Dirac bra and ket notation we can write this as $\langle\phi|A\psi\rangle = \langle A\phi|\psi\rangle$. But the specific matrix properties enforced by

self-adjoincy depend on the definition of an inner product used, and there are infinite ways to define an inner product on a vector space. Vectors in the Hilbert space are written in terms of the basis vectors, and if the basis vectors are orthonormal then the inner product is just the standard one: $\langle\phi|\psi\rangle = \sum_i \phi_i^* \psi_i = \phi^\dagger \psi$. So an operator that is 'Hermitian' is self-adjoint with respect to a given inner product rule, and in the case of the standard Hermitian inner product this means the matrix representation of the operator is equal to its complex conjugate transpose.

About a decade ago, Carl Bender *et al* showed that the assumption that operators, most importantly the Hamiltonian operator, need not be Hermitian in order to construct a consistent theory of quantum mechanics [4]. He showed that all of the virtues of Hermiticity *e.g.* real eigenvalues, unitary time evolution, etc., can be obtained by adopting an alternative set of assumptions. Non-Hermitian quantum mechanics is only "non-Hermitian" with respect to the standard inner product; operators in Bender's theory are self-adjoint with respect to a different inner product. Unlike the assumption of Hermiticity, the assumptions drafted by Bender cannot be summarized into a single statement, and so at first one might wonder why trade one perfectly good axiom for a whole bunch of them– isnt it easier and more importantly, more elegant to just assume operators are Hermitian? Furthermore Hermiticity is selected out by the orthonormal basis vectors as their preferred inner product, so why bother with a different inner product and different set of assumptions?

The reasons for considering inner products besides the standard one and non-Hermitian Hamiltonians are several. First, non-Hermitian quantum mechanics enlarges the set of Hamiltonians we are allowed to consider quantum mechanically, so it increases the number of systems we can analyze and solve. Another reason is that non-Hermitian quantum mechanics puts physical properties and principles at the forefront of the theory; specifically, the parity (P) and time-reversal (T) operators taken on the analogous role to the Hermitian conjugate. In comparison with the P and

T operators, the complex conjugate transpose seems rather arbitrary, (and we wind up imposing P and T symmetry in Hermitian theories later down the line, so why not at the level of axioms.) And then there is the so-called totalitarian principle: "Anything which is not forbidden is compulsory". First introduced in T.H. White's *The Once and Future King* as the governing principle of a colony of ants [5], and later conjectured by Murray Gell-Mann to apply to the physical laws that govern the universe [6], the totalitarian principle in this case says that there needn't be anything *wrong* with Hermitian quantum mechanics– but if we *can* do non-Hermitian quantum mechanics then we must.

That is a splendid concept but it is not going to convince your grant monitor. So is there any *real* reason to pursue non-Hermitian quantum mechanics? In what follows I hope to convince you that there is, and that non-Hermitian quantum mechanics is the idea that one might discover new and physically relevant theories by considering other inner products than the standard one.

We begin by constructing the formal extension of Bender's PT quantum mechanics to include systems that are odd under time-reversal ($T^2 = -1$). This formalism is used to develop a non-Hermitian version of the Dirac equation in Chapter 2. We show that, remarkably, the Dirac equation constructed according to the principles of PT quantum mechanics is identical to the Hermitian Dirac equation, thereby endowing non-Hermitian quantum mechanics with a host of observed phenomena; anything that is accurately described by the standard Dirac equation is also described by the PT Dirac equation. Even more remarkable is that higher dimensional representations[1] of the non-Hermitian Dirac equation describe new particles, with properties forbidden in ordinary quantum mechanics and the Standard Model. Finally we present an example of non-Hermitan quantum mechanics from the condensed matter literature; in 1956 Dyson [7] showed that high precision calculations of interacting spin

[1]The dimensionality here is that of the Hamiltonian and other operators. The spacetime dimensionality we assume is that of the real world, *ie* 3+1.

waves in a ferromagnet were facilitated by employing a non-Hermitian Hamiltonian. We recast this result in the context of PT quantum mechanics and speculate on whether non-Hermitian quantum mechanics may be able to shed light on the most outstanding problem in condensed matter physics, the theory of high temperature superconductivity[8].

Chapter 2

$\mathrm{T}_{\mathrm{odd}}$ PT Quantum Mechanics

2.1 Time-Reversal in Quantum Mechanics

Wigner was the first to derive properties and consequences of time-reversal symmetry in quantum mechanics [11] . He assumed time-reversal symmetry (T) should be antilinear in order to be consistent with the Schrödinger equation, among other things. It follows that there are only two types of time reversal, even ($T^2 = 1$)and odd ($T^2 = -1$). To see this, assume T is anti-linear:

$$T\psi = L\psi^* \tag{2.1}$$

where L is a linear operator (this is the definition of an anti-linear operator– one that can be written as complex conjugation followed by a linear operator). T^2 should leave a state unchanged up to a phase factor: $T^2\psi = e^{i\phi}\psi$. It follows that

$$\begin{aligned} LL^* &= e^{i\phi} \\ \Rightarrow \quad L^* &= e^{i\phi}L^{-1} \end{aligned} \tag{2.2}$$

.

Complex conjugating both sides gives

$$L = e^{-i\phi}\ (L^{-1}\ )^*. \tag{2.3}$$

So $LL^* = e^{-i\phi}$ also, which implies

$$(e^{\imath\phi}\ )^2 = 1 \rightarrow e^{\imath\phi} = \pm 1. \tag{2.4}$$

Wigner assumed that L is unitary; the proof given here does not make that assumption and hence generalizes the proposition to the non-Hermitian case.

Note that there is a subtle difference in the way anti-linear operators transform under change of basis as compared to linear operators: suppose that in some basis

$$T\psi = L\psi^*. \tag{2.5}$$

where ψ is the N component wave function and L is a $N \times N$ matrix. If we perform a change basis $\psi' = V\psi$ for some transformation matrix V, in this new basis T has the form

$$T\psi' = L'\psi'^*. \tag{2.6}$$

To determine the matrix L' we may proceed as follows. The wave function of the state in the old basis is given by $V^{-1}\psi'$. By eq (2.5) the time-reversed state has wave function $LV^{-1*}\psi'^*$ in the old basis. Thus the time-reversed state has wave function $VLV^{-1*}\psi'^*$ in the new basis. Comparing to eq (2.6) we conclude

$$L' = VLV^{-1*}. \tag{2.7}$$

Note that by contrast if T had been a linear operator represented by a matrix L in the old basis, then in the new basis it would have the matrix VLV^{-1}.

In all work on PT quantum mechanics to date it has been implicitly assumed that time reversal is even, $T^2 = 1$. However in quantum theory this is only true of bosonic systems with integer spin. For fermionic systems, with half-integer spin, time-reversal is odd. Quarks and leptons in particle physics, approximately half of all nuclei and atoms, and a plethora of condensed matter problems including magnetic spin models and solid state electronic matter fall into this category. Thus it is clearly important to generalize the construction of PT quantum mechanics to the case that T is odd.

2.2 $\mathbb{T}_{\mathbf{even}}$ PT Quantum Mechanics

Before embarking on our generalization of Bender's PT quantum mechanics to the case where time-reversal is odd, it is useful to recall the key principles of PT quantum mechanics for the case that time reversal is even. For simplicity let us assume that the Hilbert space of states has finite dimension N so that the state of the system may be specified by a wavefunction that is an N component column vector with complex components $\psi(n)$, with $n = 1, 2, 3, \ldots N$. We will work in a basis such that the operation of time reversal consists of simply taking the complex conjugate of the wavefunction, $T\psi = \psi^*$; thus $T^2 = 1$; such a basis can always be found.

Parity is a linear operator so it may be represented by a matrix that we denote S; thus $P\psi = S\psi$. We assume $[P, T] = 0$ and that parity applied twice is the identity transformation; this implies $S = S^*$ and that $S^2 = \mathbb{I}$. Since $S^2 = \mathbb{I}$ then it must have eigenvalues ± 1, and without loss of generality we can find a basis in which S is diagonal and T still consists of simple conjugation.

To see this, recall how T transforms under basis change from eq (2.7). If we make a change of basis $\psi' = U^{-1}\psi$, then $T\psi' = U^{-1}U^*\psi'^*$, but because P is linear and S is real and squares to the identity, $P\psi' = U^{-1}SU\psi'$. If we can choose U such that

$S' = U^{-1}SU$ is diagonal and $U^{-1}U^* = \mathbb{I}$, then we will have achieved the desired basis transformation: in the primed basis parity will be diagonal and time-reversal will still consist of just complex conjugation.

S is real and squares to the identity, so the eigenvectors of S may be chosen to be real (because they are determined by solving $S\psi = \lambda\psi$ and both S and $\lambda = \pm 1$ are real). Let us construct the matrix

$$U = \begin{pmatrix} | & & | \\ \psi_1 & \dots & \psi_N \\ | & & | \end{pmatrix}; \tag{2.8}$$

here $\psi_1, \dots, \psi_N$ are the real eigenvectors of S. This matrix is real and by the customary reasoning in linear algebra it diagonalizes S

$$U^{-1}SU = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}. \tag{2.9}$$

Hence we can always find a basis in which S can be written

$$S = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix} \tag{2.10}$$

where $\mathbb{I}$ denotes the $N/2$ dimensional identity matrix. (We have assumed N is even and that S is traceless just for simplicity.)

In quantum mechanics one conventionally defines the inner product of two states as $(\phi, \psi) = \phi^\dagger\psi = \sum_{n=1}^{N} \phi^*(n)\psi(n)$. However in PT quantum mechanics a different inner product is used. Because the inner product plays an integral role in the formulation of non-Hermitian quantum mechanics, we digress briefly to recall some properties of

the inner product.

2.2.1 Properties of the Inner Product

'Inner product' refers to the rubric one uses to combine vectors in a vector space [9]. The inner product is not unique, in fact infinitely many inner products exist for a given vector space[**?**]. If u and v are vectors in a vector space, we denote their inner product as (u, v) ; in general it is a complex number. Inner products are assumed to satisfy two conditions. First, it is assumed that $(u, v) = (v, u)^*$. Second, it is assumed that the inner product rule is bilinear which means that:

(i) $(u, v + w) = (u, v) + (u, w)$ and $(u + v, w) = (u, w) + (v, w)$

(ii) $(u, cv) = c(u, v)$ where c is a complex number.

(iii) $(cu, v) = c^*(u, v)$.

In physics we also require that the inner product of a vector with itself be positive, and be zero in the case that one of the vectors is the zero vector. Such inner products are called positive definite.

A vector space is of course spanned by a set of basis vectors; the inner product is related to the basis vectors by calculating inner product of all pairs of basis vectors: if $e_1, \ldots, e_N$ form a basis for vector space V,

$$\kappa_{ij} = (e_i, e_j) \tag{2.11}$$

κ_{ij} are called the *kernel* of the inner product. If the kernel is known, the inner product of any pair of vectors can be evaluated. Specifically consider the vectors

$$\begin{aligned} u &= b_1 e_1 + b_2 e_2 + \ldots + b_N e_N \\ v &= c_1 e_1 + c_2 e_2 + \ldots + c_N e_N \end{aligned} \tag{2.12}$$

where the coefficients are complex numbers. Their inner product is

$$(u, v) = \sum_{i=1}^{N} \sum_{j=1}^{N} b_i^* \kappa_{ij} c_j \tag{2.13}$$

Eq (2.13) may be derived by writing

$$(u, v) = (\sum_{i=1}^{N} b_i e_i, \sum_{j=1}^{N} c_j e_j) \tag{2.14}$$

and using the bilinearity of the inner product.

A more compact notation is to write

$$(u, v) = b^\dagger \kappa c \tag{2.15}$$

where $b^\dagger$ denotes the row vector $(b_1^*, \ldots, b_N^*)$, c is the column vector of similar composition, and κ is the matrix comprised of the kernel elements κ_{ij}.

A basis is said to be orthonormal if the kernel matrix is the identity:

$$(e_i, e_j) = \delta_{ij}. \tag{2.16}$$

In an orthonormal basis the formula for inner product eq (2.13) simplifies to

$$(u, v) = b^\dagger c. \tag{2.17}$$

An orthonormal basis can always be found using the process of Gram-Schmitt orthogonalization [10]. In physics we usually choose an orthonormal basis, hence the inner product is the standard Hermitian conjugate. We will refer to this choice as the standard inner product.

2.2.2 PT Inner Product

Given that in PT quantum mechanics, P and T take on a role analogous to the Hermitian conjugate in ordinary quantum mechanics, a natural way to define the PT inner product is $(\phi, \psi)_{PT} = (PT\phi)^T\psi = \phi^\dagger S\psi$. However with this definition there are some states of negative norm. For example, consider an N dimensional column vector ϕ which for convenience we break up into a pair of $N/2$ dimensional segments:

$$\phi \to \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \tag{2.18}$$

The norm of ϕ under the PT inner product is

$$\begin{aligned} (\phi, \phi)_{PT} = \phi^\dagger S\phi &= (\phi_1^*, \phi_2^*) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \\ &= |\phi_1|^2 - |\phi_2|^2 \end{aligned} \tag{2.19}$$

which is negative for $|\phi_1|^2 < |\phi_2|^2$. So the PT inner product is not a viable inner product to use in quantum mechanics.

Positive definiteness is restored by introducing a linear operator C that takes eigenstates of the Hamiltonian that have negative norm under the PT inner product and turns them into positive [1]. Let ψ_i be eigenvectors of H and s_i is the sign of the PT norm of the eigenvector, $(\psi_i, \psi_i)_{PT}$. C is defined by its action on the ψ_i:

$$C\psi_i = s_i\psi_i \tag{2.20}$$

but because it is a linear operator it can be represented by some matrix K in the

[1] It is also possible that there are eigenvectors that are orthogonal to themselves under the PT inner product. In the absence of degeneracies such an orthogonality is a catastrophe in the sense that it is then impossible to formulate PT quantum mechanics for the Hamiltonian under consideration.

standard basis:

$$C\psi = K\psi. \tag{2.21}$$

The operator C commutes with the combination PT (although it may not commute with either separately), this implies $KS = SK^*$. Furthermore $C^2 = 1$. We now define the CPT inner product:

$$(\phi, \psi)_{CPT} \equiv (CPT\phi)^T\psi = \phi^\dagger K^T S\psi \tag{2.22}$$

This is the inner product used in PT quantum mechanics in lieu of the standard inner product. It is evident from the definition given here that the Hamiltonian plays a crucial role in determining the operator C, so we say that the CPT inner product is 'dynamically determined'.

Under the CPT inner product all non-trivial states have positive norm, time evolution is unitary with respect to this inner product. Thus it is possible to consistently formulate quantum mechanics using the CPT inner product, notwithstanding the non-Hermiticity of the Hamiltonian.

2.2.3 Observables

Conventionally one requires the Hamiltonian H (and all other observables) to be Hermitian, $H^\dagger = H$. In this context Hermitian is synonymous with self-adjoint. An operator A is said to be self-adjoint if $(A\phi, \psi) = (\phi, A\psi)$. Clearly the choice of inner product determines the specific matrix properties A must have in order to be self-adjoint, so it is more precise to say that in ordinary quantum mechanics 'Hermitian' means self-adjoint with respect to the standard inner product. Alternatively we can say that 'non-Hermitian' quantum mechanics is Hermitian but with respect to a non-standard inner product.

We define the CPT adjoint $A^\star$ of an operator A by imposing $(\phi, A\psi)_{CPT} =$

$(A^\star\phi,\psi)_{CPT}$ for all ϕ and ψ. Observables are then required to be CPT self-adjoint, $A = A^\star$. This is sufficient to ensure that the eigenvalues of A are real and that the usual principles of quantum measurement and uncertainty relations may be applied even though the observables are no longer Hermitian in the usual sense.

The proof of this (as in the Hermitian case) relies on the Schwarz inequality $(\alpha|\alpha)(\beta|\beta) \geq |\ (\alpha|\beta)^2$ for any vectors α and β in a vector space with a positive definite inner product rule. Since we have such an inner product with the definition given in eq (2.22) we can simply follow the standard proof given in textbooks on Hermitian quantum mechanics [12].

Assume there are CPT self-adjoint operators A, B, and C such that

$$[A, B] = iC, \tag{2.23}$$

and CPT self-adjoint operators O, F and G such that

$$F = O + O^\star \quad \text{and} \quad G = -i\ (O - O^\star\ ). \tag{2.24}$$

Then

$$O = \frac{O + O^\star}{2} + \frac{O - O^\star}{2} = \frac{F}{2} + \frac{iG}{2} \tag{2.25}$$

Let us define

$$\begin{aligned} |\alpha) &= (A - \langle A\rangle\ )|\psi) \\ |\beta) &= (B - \langle B\rangle\ )|\psi) \end{aligned} \tag{2.26}$$

where$\langle A\rangle = (\psi|A|\psi)$ is the expectation value of A (and similarly for B), and as such

are real numbers. Notice that

$$\begin{aligned} (\alpha|\alpha) &= (\psi|\ (A - \langle A\rangle\ )^2|\psi) = \Delta A^2 \\ (\beta|\beta) &= (\psi|\ (B - \langle B\rangle\ )^2|\psi) = \Delta B^2 \end{aligned} \tag{2.27}$$

using the standard definition of uncertainty, *ie* $(\Delta x\ )^2 = \langle x^2\rangle - \langle x\rangle^2$. Plugging the definition of $(\alpha|\beta)$ into the right hand side of the Schwarz inequality, we have

$$(\alpha|\beta\,) = (\psi|\ (A - \langle A\rangle\ )(B - \langle B\rangle\ )|\psi). \tag{2.28}$$

Now if we define $O = (A - \langle A\rangle\ )\ (B - \langle B\rangle\ )$, then

$$O - O^\star = [A, B] = iC \tag{2.29}$$

from which we conclude $G = C$. Thus,

$$\begin{aligned} |(\alpha|\beta\,)|^2 &= |\frac{1}{2}\ (\psi|F|\psi) + \frac{i}{2}\ (\psi|C|\psi)|^2 \\ &= \frac{|(\psi|F|\psi)|^2}{4} + \frac{|(\psi|C|\psi)|^2}{4} \geq \frac{|(\psi|C|\psi)|^2}{4} \end{aligned} \tag{2.30}$$

because the expectation values are real. From the Schwarz inequality, then, we see

$$\Delta A^2 \Delta B^2 \geq \frac{|\langle C\rangle|^2}{4} \ \Rightarrow\ \Delta A \Delta B \geq \frac{|\langle C\rangle|}{2}. \tag{2.31}$$

So the famous uncertainty relations are preserved in PT quantum mechanics. Note that the proof depends only on the positive definiteness of the inner product rule and the fact that the expectation values are real (and that the operators are self-adjoint,

of course). Thus it can be applied quite generally to non-Hermitian extensions of quantum mechanics.

2.2.4 Criteria for T_{even} Hamiltonians

We have seen that with a different definition of inner product, it is still possible formulate a theory with real-valued observables, positive definite norm, and unitary time evolution. We turn now to the conditions that must be met by the Hamiltonian operator in particular in order to be a valid PT quantum mechanical system.

Invariance Under PT

First, the Hamiltonian must be invariant under PT (*i.e.* it must commute with PT). If we write H as

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{2.32}$$

where A, B, C and D are arbitrary matrices and $[H, PT] = 0$, then expanding $HPT\psi = PTH\psi$:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \psi^* = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} A^* & B^* \\ C^* & D^* \end{pmatrix} \psi^*$$

$$\begin{pmatrix} A & -B \\ C & -D \end{pmatrix} = \begin{pmatrix} A^* & B^* \\ -C^* & -D^* \end{pmatrix}$$

hence

$$H = \begin{pmatrix} A & iB \\ iC & D \end{pmatrix} \tag{2.33}$$

where now A, B, C and D are real matrices. It also follows from invariance under PT that the eigenvalues of H come in conjugate pairs: suppose $[H, PT] = 0$ and ϕ is an

eigenfunction of H with eigenvalue λ, and and $HPT\phi = PTH\phi = PT\ (\lambda\phi\ )$. Let $\psi = PT\phi = S\phi^*$, then $H\psi = \lambda^*\psi$. So

$$H\phi = \lambda\phi \Rightarrow PT\phi = \lambda^*\phi. \tag{2.34}$$

Unbroken PT

The second criterion is that PT must be 'unbroken' in the sense that it should be possible to find eigenvectors of the Hamiltonian, ψ_i, that are invariant under PT (*i.e.* $PT\psi_i = \psi_i$). This is crucial as it ensures the eigenvalues of H are real; since this is a subtlety arising from the prominent role of parity and time-reversal in PT quantum mechanics, we pause to prove that if PT is unbroken then the eigenvalues of H must all be real, and the converse, if the eigenvalues H are real then PT is unbroken. Note that a state that is invariant under PT will have the form

$$\psi = \begin{pmatrix} \xi \\ i\eta \end{pmatrix} \tag{2.35}$$

where ξ and η are real column vectors with $N/2$ components each and we are assuming that S has the form given in eq (2.10). Before embarking on the proof, we note that it is an elementary proposition of linear algebra that if H and A commute and are linear operators then there exists a set of simultaneous eigenvectors. However since PT is not a linear operator we have no reason to believe that H and PT should have simultaneous eigenvectors. We also note that proving that PT is unbroken (or equivalently that the eigenvalues are real) is frequently the most difficult step in PT quantum mechanics [2]. There is at least one instance involving a Schrödinger equation with an imaginary potential where the proof is fifty pages and involves the use of Bethe ansatz!

[2]Hermiticity is a sufficient condition for the eigenvalues to be real but not necessary. Nonetheless, a generic non-Hermitian operator has complex eigenvalues; see for example ref[13].

We first show that if PT is unbroken then the eigenvalues of H are real. Let ψ_i be a set of eigenvectors of H with eigenvalue λ_i that are invariant under PT. Now by the reasoning embodied in eq (2.34) if ψ_i is an eigenvector of H with eigenvalue λ_i then $PT\psi_i$ is an eigenvector of H with eigenvalue λ_i^*. But since ψ_i is invariant under PT, it follows ψ_i is an eigenvector of H with eigenvalue λ_i^*. ψ_i can have eigenvalue λ_i^* and λ_i only if $\lambda_i^* = \lambda_i$. Thus the eigenvalues are real as claimed.

Next let us prove the converse, that if the eigenvalues of H are real, then PT must be unbroken. For simplicity we assume that there are N non-degenerate eigenvalues. Our objective is to find a set of eigenvectors of H that are also invariant under PT given that the eigenvalues are real. We start from eq (2.34) which asserts that if ψ_i is an eigenvector of H with eigenvalue λ_i, then $PT\psi_i$ is an eigenvector with eigenvalue λ_i^*. Since the eigenvalues are assumed real, ψ_i and $PT\psi_i$ both have the same eigenvalue λ_i. Since the spectrum is assumed non-degenerate it follows that the two states ψ_i and $PT\psi_i$ must be linearly dependent on each other; namely $PT\psi_i = \mu\psi_i$. We now invoke a lemma (proved below) that μ must be a pure phase, $\mu = \exp(i\gamma)$. It is then easy to verify that

$$PT\psi_i = \exp(i\gamma)\psi_i \;\Rightarrow PT\tilde{\psi}_i = \tilde{\psi}_i \tag{2.36}$$

where $\tilde{\psi}_i = \exp(-i\gamma/2)\psi_i$. Thus $\tilde{\psi}_i$ constitute a set of eigenvectors of H that are also invariant under PT.

Finally, the lemma that if $PT\psi = \mu\psi$, then μ must be pure phase. Note that

$$\begin{aligned} PT\psi &= \mu\psi \\ \Rightarrow S\psi^* &= \mu\psi \\ \Rightarrow \psi^T S^\dagger &= \mu^*\psi^\dagger \\ \Rightarrow \psi^T S^\dagger S\psi^* &= |\mu|^2\psi^\dagger\psi. \end{aligned} \tag{2.37}$$

Finally note that in the diagonal basis $S^\dagger = S$ and therefore $S^\dagger S = S^2 = \mathbb{I}$. Furthermore $\psi^T \psi^* = \psi^\dagger \psi$ (as can easily be verified by writing both sides in terms of the components of the wave function). Thus eq (2.37) implies that $|\mu|^2 = 1$ and thus μ is a pure phase as claimed.

Self-duality Under PT Inner Product

First it is useful to recall that the PT inner product is given by

$$(\phi, \psi)_{\mathrm{PT}} = (PT\phi)^T \psi = \phi^\dagger S \psi. \tag{2.38}$$

We assume we are working in a basis wherein time-reversal is given by conjugation. S is assumed to be real (because P and T commute), symmetric and to satisfy $S^2 = 1$ (because $P^2 = 1$). It is not necessary to assume that in our basis S is diagonal.

We define A_D, the PT dual of the operator A, by the condition that

$$(A_D \phi, \psi)_{\mathrm{PT}} = (\phi, A\psi)_{\mathrm{PT}}\,. \tag{2.39}$$

This condition should be met for all states ϕ and ψ. Using eq (2.38) one can derive the explicit formula

$$A_D = S A^\dagger S \tag{2.40}$$

The formula eq (2.40) can be simplified further if we assume that A commutes with PT. In that case we find

$$A_D = A^T. \tag{2.41}$$

The proof is as follows: That A commutes with PT implies $AS = SA^*$. Left multiplication by S leads to $SAS = A^*$. Transposing both sides and bearing in mind that S is symmetric leads to $A^\dagger = SA^T S$. Substituting this expression for $A^\dagger$ in eq (2.40) and using $S^2 = 1$ leads to eq (2.41).

Now to the issue of orthogonality of eigenvectors . Let us suppose that H is self-dual. Let us also assume that H commutes with PT and that PT is unbroken. Let ψ_i with $i = 1, \ldots, N$ denote the eigenvectors of H and λ_i the corresponding eigenvalues. It follows from the assumptions we have made about H (chief among them, self-duality) that if two eigenvectors have distinct eigenvalues they must be orthogonal under the PT inner product.

The proof is as follows: By virtue of self-duality, $(\psi_i, H\psi_j)_{\rm PT} = (H\psi_i, \psi_j)_{\rm PT}$. Since the ψ_i's are eigenvectors of H, it follows $(\psi_i, \lambda_j\psi_j)_{\rm PT} = (\lambda_i\psi_i, \psi_j)_{\rm PT}$. By the anti-linear property of inner products this equation may be written as $\lambda_j\,(\psi_i, \psi_j)_{\rm PT} = \lambda_i^*\,(\psi_i, \psi_j)_{\rm PT}$. Since we have assumed that PT is unbroken, the eigenvalues of H are real; hence we may write

$$(\lambda_j - \lambda_i)\,(\psi_i, \psi_j)_{\rm PT} = 0. \tag{2.42}$$

It follows that if $\lambda_i \neq \lambda_j$ then inevitably $(\psi_i, \psi_j)_{\rm PT} = 0$. This establishes the claimed orthogonality.

Philosophically, we want the eigenvectors of any observable to be orthogonal to each other under the dynamical inner product since this is a feature of conventional quantum mechanics. Requiring the Hamiltonian to be self-dual under the PT inner product is a stepping stone to that goal. It ensures that the eigenvectors of H are appropriately orthogonal to each other under the PT inner product.

This concludes our rèsumè of the principles of PT quantum mechanics for the case of even time reversal symmetry. We now construct the extension of these principles to the case that time reversal symmetry is odd.

2.3 $\mathrm{T}_{\mathbf{odd}}$ PT Quantum Mechanics

Having reviewed the case of even time-reversal is useful as now we can go more quickly through the corresponding formalism for the T_{odd} case. Lets begin with the definition

of inner product. We work in a basis where the action of time-reversal is given by

$$T\psi = Z\psi^* \tag{2.43}$$

where Z is a matrix that yields $T^2\psi = -\psi$ and will be specified later. Once again we assume $P^2 = 1$ or equivalently $S^2 = \mathbb{I}$ and that $PT = TP$ which implies $SZ = ZS^*$. Thus parity can be written

$$S = \begin{pmatrix} \mathbb{I}_N & 0 \\ 0 & -\mathbb{I}_N \end{pmatrix}. \tag{2.44}$$

We will prove later that such a basis can always be found. In the even case, the PT inner product was given by $(\phi, \psi)_{PT} = (PT\phi)^T\psi = \phi^\dagger S\psi$. Now that the action of T has changed we expect the PT inner product to reflect this feature. We define the PT inner product for T_{odd} systems as $(\phi, \psi)_{PT} = (PT\phi)^T Z\psi = \phi^\dagger S\psi$; note the crucial insertion of Z. As in the even case, the PT inner product on its own is not a viable inner product for quantum mechanics because it is not positive definite, so again we introduce the C operator that acts on eigenstates of the Hamiltonian ψ_i. C is a linear operator with a corresponding matrix K which has the defining property that $C\psi_i = s_i\psi_i$ where s_i is the sign of the PT norm of that eigenvector, (ψ_i, ψ_i). C commutes with PT so $KSZ = SZK^*$. As in the even case $C^2 = 1$. Observables are CPT self-adjoint, $A = A^\star$, although because the inner product is slightly different from the even case the matrix properties of self-adjoint operators are slightly different as well. And finally, we impose the same three criteria on Hamiltonian operators as in the even case:

(i) H is invariant under PT : $[H, PT] = 0$

(ii) PT symmetry is unbroken.

(iii) H is self-dual under the PT inner product.

The criteria are constructed to be the same but they have quite different implications here in the T_{odd} case, including a PT analogue of Kramers degeneracy from ordinary T_{odd} quantum mechanics. For the interested reader we now present the construction of these criteria in detail; the uninterested reader may skip to the next section.

2.3.1 Construction and Implications of Criteria

Quaternions

First it is useful to review properties and notation relevant to quaternions. In the remainder of this chapter we will refer to 2×2 matrices as quaternions. Any quaternion q can be written

$$\begin{aligned} q &= q_0\sigma_0 + iq_1\sigma_1 + iq_2\sigma_2 + iq_3\sigma_3 \\ &= q_0 + i\mathbf{q}\cdot\boldsymbol{\sigma} \\ &= \begin{pmatrix} q_0+iq_3 & iq_1+q_2 \\ iq_1-q_2 & q_0-iq_3 \end{pmatrix}. \end{aligned} \tag{2.45}$$

The coefficients q_0, q_1, q_2, q_3 are complex numbers in general, but in the case that they are real we will say the quaternion is real. It is important to realize that this does not mean the corresponding 2×2 matrix is real, though. By suitably partitioning a $2N\times 2N$ matrix into 2×2 blocks one can view it as an $N\times N$ matrix of quaternions.

Consider the matrix

$$Z = \begin{pmatrix} e_2 & & \\ & \ddots & \\ & & e_2 \end{pmatrix}, \quad \text{where } e_2 = i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.46}$$

Note that Z is an $N\times N$ quaternion matrix. If another quaternion matrix A

satisfies

$$AZ = ZA^* \tag{2.47}$$

then A is quaternion real (*i.e.*, it is composed of real quaternions), and vice versa.

If we assume its eigenvalues to be real then A has some interesting properties. First, that its eigenvalues come in degenerate pairs, one belonging to the eigenvector ψ and the other belonging to the eigenvector $-Z\psi^*$. To see this, let ψ be an eigenvector of A with the (real) eigenvalue λ. Then

$$\begin{aligned} A\psi &= \lambda\psi \\ \Rightarrow A^*\psi^* &= \lambda\psi^* \\ \Rightarrow -ZA^*\psi^* &= -\lambda Z\psi^* \\ \Rightarrow A(-Z\psi^*) &= \lambda(-Z\psi^*). \end{aligned} \tag{2.48}$$

So ψ and $-Z\psi^*$ are both eigenvectors of A with the same eigenvalue λ.

The next interesting property is that A is made diagonal by a matrix U which is also quaternion real. To see this, let ψ be a $2N$ component eigenvector of A:

$$\psi = \begin{pmatrix} a(1) \\ b(1) \\ \vdots \\ a(N) \\ b(N) \end{pmatrix}. \tag{2.49}$$

Recall that $-Z\psi^*$ is another eigenvector of A with the same eigenvalue as ψ, and

from eq (2.46) and (2.49) it follows that

$$-Z\psi^* = \begin{pmatrix} -b(1)^* \\ a(1)^* \\ \vdots \\ -b(N)^* \\ a(N)^* \end{pmatrix}. \tag{2.50}$$

If we assemble ψ and $-Z\psi^*$ into a pair of columns

$$\begin{pmatrix} | & | \\ \psi & -Z\psi^* \\ | & | \end{pmatrix} = \begin{pmatrix} a(1) & -b(1)^* \\ b(1) & a(1)^* \\ \vdots & \vdots \\ a(N) & -b(N)^* \\ b(N) & a(N)^* \end{pmatrix} \tag{2.51}$$

we may regard this eigenvector-doublet as a $2N \times 2$ complex matrix or an N component quaternion column. If we write the complex numbers $a(1) = q_0(1) + iq_3(1)$ and $b(1) = -q_2(1) + iq_1(1)$ where $q_0(1), q_1(1), q_2(1)$ and $q_3(1)$ are all real, then substitute these definitions into eq (2.51) we find

$$\begin{pmatrix} | & | \\ \psi & -Z\psi^* \\ | & | \end{pmatrix} = \begin{pmatrix} q_0(1) + iq_3(1) & q_2(1) + iq_1(1) \\ -q_2(1) + iq_1(1) & q_0(1) - iq_3(1) \\ \dots & \dots \\ q_0(N) + iq_3(N) & q_2(N) + iq_1(N) \\ -q_2(N) + iq_1(N) & q_0(N) - iq_3(N) \end{pmatrix}. \tag{2.52}$$

Comparing to eq (2.46) we see therefore that the eigenvector-doublet is a column of real quaternions.

Finally we define U, the 'eigenmatrix' of A, by assembling these eigenvector-doublets into columns,

$$U = \begin{pmatrix} | & | & & | & | \\ \psi_1 & -Z\psi_1^* & \dots & \psi_N & -Z\psi_N^* \\ | & | & & | & | \end{pmatrix}. \tag{2.53}$$

Because it is comprised of N eigenvector-doublets that are quaternion real, U itself and U^{-1} are also quaternion real, and these matrices serve to diagonalize A:

$$U^{-1}AU = \begin{pmatrix} \lambda_1\sigma_0 & & \\ & \ddots & \\ & & \lambda_N\sigma_0 \end{pmatrix}. \tag{2.54}$$

So, when regarded as an $N \times N$ quaternion matrix, A is clearly quaternion real in light of the analysis of the eigenvector-doublet above.

Representation of Parity and Time Operators

We work in a basis where $T\psi = Z\psi^*$ and Z is given by eq(2.46) ; note that $T^2\psi = -\psi$ as needed. Of course other matrices could satisfy this property but eq(2.46) is best suited to our purposes. We would like it if , as in the even case, the same basis in which T has the canonical form eq(2.46) allows us to write parity as eq(2.44); we now prove such a basis exists.

We make the same assumptions about parity here as we did in section 2.2: $P^2 = 1$ or equivalently $S^2 = \mathbb{I}$, and $PT = TP$ which implies that $SZ = ZS^*$.

The first assumption tells us that the eigenvalues of S are ± 1, and using the second assumption and the results of the previous section we know S is a quaternion real matrix. As such, the eigenvalues of S come in degenerate pairs and the eigenmatrix U that diagonalizes S may be chosen to be quaternion real.

Recall from section 2.2 that if we now change basis $\psi' = V^{-1}\psi$ then the matrix representing parity changes from S to $V^{-1}SV$ and time reversal in the new basis consists of conjugation followed by multiplication by $V^{-1}ZV^*$.

If we cleverly choose the transformation matrix $V = U$ (the matrix that diagonalizes S), then we transform into a basis where parity is diagonal. Time reversal in this basis consists of conjugation followed by multiplication by $U^{-1}ZU^*$. Since U^{-1} is quaternion real it obeys $U^{-1}Z = ZU^{-1*} \Rightarrow U^{-1}ZU^* = Z$. Thus in the new basis time reversal still consists of conjugation followed by multiplication by Z. This proves that we can always find a basis in which simultaneously S is diagonal and time-reversal has the canonical form.

Unbroken PT and Kramers Degeneracy

A key role is played in even PT quantum mechanics by states that are invariant under PT. In odd PT quantum mechanics however there are no states that are invariant under PT; the nearest analogue is the concept of the PT doublet which we now introduce.

Consider the pair of states ϕ and $-PT\phi$. Together these states constitute a 'PT doublet'. We write the doublet as a $2N \times 2$ complex matrix

$$\begin{pmatrix} | & | \\ \phi & -PT\phi \\ | & | \end{pmatrix}. \tag{2.55}$$

With this definition it is easy to see that if we apply PT to the doublet we get

$$\begin{pmatrix} | & | \\ PT\phi & \phi \\ | & | \end{pmatrix} \tag{2.56}$$

because $(PT)^2 = P^2T^2 = -1$. Exactly the same outcome would result if we post-multiplied the doublet, eq (2.55), by $e_2 = i\sigma_2$. Thus the doublet is not left invariant by PT but it is merely multiplied by a constant matrix. Thus it is a close analogue of the PT invariant state in T_{even} PT quantum mechanics.

It is instructive to write out all the components of the doublet. First we work out the components of ϕ and $PT\phi$:

$$\phi = \begin{pmatrix} a(1) \\ b(1) \\ \cdots \\ a(N/2) \\ b(N/2) \\ a(N/2+1) \\ b(N/2+1) \\ \cdots \\ a(N) \\ b(N) \end{pmatrix} \Rightarrow PT\phi = \begin{pmatrix} b(1)^* \\ -a(1)^* \\ \cdots \\ b(N/2)^* \\ -a(N/2)^* \\ -b(N/2+1)^* \\ a(N/2+1)^* \\ \cdots \\ -b(N)^* \\ a(N)^* \end{pmatrix}. \tag{2.57}$$

For simplicity we have assumed that N is even. Note the subtle difference in the lower half components of $PT\phi$ compared to the upper half. Before we stack ϕ and $-PT\phi$ side by side let us write out the complex components of ϕ in terms of real parameters. We define $a(i) = q_0(i) + iq_3(i)$ and $b(i) = -q_2(i) + iq_1(i)$ for $i = 1, \ldots, N/2$. For the lower components we define $a(i) = iq_0(i) - q_3(i)$ and $b(i) = -iq_2(i) - q_1(i)$. Here the q parameters are all real. Re-writing ϕ and $PT\phi$ in terms of the q parameters and

stacking them side by side we find that the PT doublet may be written as

$$\begin{pmatrix} q_0(1)\sigma_0 + i\mathbf{q}(1)\cdot\boldsymbol{\sigma} \\ \ldots\, q_0(N/2)\sigma_0 + i\mathbf{q}(N/2)\cdot\boldsymbol{\sigma} \\ i[q_0(N/2+1)\sigma_0 + i\mathbf{q}(N/2+1)\cdot\boldsymbol{\sigma}] \\ \ldots \\ i[q_0(1)\sigma_0 + i\mathbf{q}(1)\cdot\boldsymbol{\sigma}] \end{pmatrix}. \tag{2.58}$$

In other words, regarded as an N component column of quaternions, the upper half of the PT doublet is composed of real quaternions and the lower half of pure imaginary quaternions (real quaternions multiplied by i).

Finally we note that the two vectors of a PT doublet are orthogonal to each other under the PT inner product. To prove this, suppose that $\phi = PT\psi$. We want to show that $(\phi,\psi)_{\rm PT} = 0$. To this end note $(\phi,\psi)_{\rm PT} = (PT\phi)^T Z\psi = [(PT)^2\psi]^T Z\psi = -\psi^T Z\psi$. Now

$$\psi = \begin{pmatrix} a_1 \\ b_1 \\ \ldots \\ a_{2N} \\ b_{2N} \end{pmatrix}, \quad -Z\psi = \begin{pmatrix} -b_1 \\ a_1 \\ \ldots \\ -b_{2N} \\ a_{2N} \end{pmatrix}. \tag{2.59}$$

Hence $-\psi^T Z\psi = (-a_1b_1 + a_1b_1) + \ldots + (-a_{2N}b_{2N} + a_{2N}b_{2N}) = 0$.

Additionally, each vector of the doublet has the same inner product with itself: let $\phi = PT\psi \Rightarrow PT\phi = -\psi$. We want to show that $(\phi,\phi)_{\rm PT} = (\psi,\psi)_{\rm PT}$. Note that

$$(\phi,\phi)_{PT} = (PT\phi)^T Z\phi = -\psi^T Z\phi. \tag{2.60}$$

On the other hand

$$(\psi,\psi)_{PT} = (PT\psi)^T Z\psi = -\phi^T Z\psi. \tag{2.61}$$

Since this expression is just a number it will not change if we transpose it. This leads to $-\psi^T Z^T \phi = \psi^T Z \phi$ (because $Z^T = -Z$ in the canonical basis). Thus $(\psi, \psi)_{\rm PT} = (\phi, \phi)_{\rm PT}$.

Having explained PT doublets, we now return to the discussion of unbroken PT. Imposing the first criterion, that $[H, PT] = 0$, has the immediate consequence that the eigenvalues come in conjugate pairs. If ϕ is an eigenvector with eigenvalue λ, then $PT\phi$ is an eigenvector with eigenvalue λ^*:

$$\begin{aligned} H\phi &= \lambda\phi \\ \Rightarrow H^*\phi^* &= \lambda^*\phi^* \\ \Rightarrow SZH^*\phi^* &= \lambda^* SZ\phi^* \\ \Rightarrow HSZ\phi^* &= \lambda^* SZ\phi^* \\ \Rightarrow H(PT\phi) &= \lambda^*(PT\phi). \end{aligned} \tag{2.62}$$

Recall that the condition of unbroken PT when T is even is that we should be able to find eigenvectors of H that are invariant under PT, and that the purpose of this condition is to ensure that the eigenvalues of H be real. In the case of T_{odd}, it is impossible for a state to be invariant under PT, so we generalize the concept of unbroken PT as follows: we say that PT is unbroken if for every eigenvector ϕ we find that the pair ϕ and $PT\phi$ are degenerate.

Clearly this condition ensures the eigenvalues must be real. Note that we have already demonstrated that the eigenvalues of ϕ and $PT\phi$ are a conjugate pair, λ and λ^*. If the states are degenerate, the eigenvalues must be real, $\lambda = \lambda^*$. Conversely if the eigenvalues are all real then clearly ϕ and $PT\phi$ are degenerate and therefore PT is unbroken.

So in odd PT quantum mechanics the condition of unbroken PT not only ensures that the eigenvalues of H are real, it also ensures they come in degenerate pairs.

This is reminiscent of Kramers theorem in ordinary quantum mechanics. Kramer's theorem asserts that if time-reversal is odd and H commutes with time-reversal then the eigenvalues of H will come in degenerate pairs.

Self-duality

The third condition that must be imposed on the Hamiltonian in PT quantum mechanics is that H is self-dual under the PT inner product. The purpose of this condition is to ensure the desirable feature that eigenvectors of H that have distinct eigenvalues of are orthogonal to each other under the PT inner product.

Recall that the PT inner product is given by

$$(\phi, \psi)_{\mathrm{PT}} = (PT\phi)^T Z\psi = \phi^\dagger S^\dagger \psi. \tag{2.63}$$

We assume that we are in a basis where T has the canonical form $T\psi = Z\psi^*$, and in this basis S is quaternion real, $S^2 = \mathbb{I}$, and $S = S^\dagger$. This last assumption would certainly be true in the basis in which S is diagonal and one option is to continue the discussion in such a basis. However it is true even in a basis in which S is not diagonal since eq (2.63) reveals that S is the kernel of the PT inner product (see section 2.2.1). One can show that the kernel of an inner product is a matrix that must be equal to its conjugate transpose because of the requirement that an inner product should be bi-linear. Thus we may also write the PT inner product in the form

$$(\phi, \psi)_{\mathrm{PT}} = \phi^\dagger S\psi. \tag{2.64}$$

A_D, the PT dual of an operator A, is defined in exactly the same way in odd PT quantum mechanics as in even, by the condition that

$$(A_D\phi, \psi)_{\mathrm{PT}} = (\phi, A\psi)_{\mathrm{PT}}. \tag{2.65}$$

This condition should be met for all states ϕ and ψ. Using eq (2.65) one can derive the explicit formula

$$A_D = SA^\dagger S; \tag{2.66}$$

the derivation is exactly the same as in the even case.

Now let us turn to the issue of orthogonality of the eigenvectors of H. Suppose that H satisfies all three conditions of odd PT quantum mechanics, *i.e.*, it commutes with PT, PT is unbroken and that H is self-dual. Let ψ_i denote an eigenfunction of H with the eigenvalue λ_i. Exactly as in the even case it follows from self-duality

$$\begin{aligned} (\psi_i, H\psi_j)_{\rm PT} &= (H\psi_i, \psi_j)_{\rm PT} \\ \Rightarrow (\lambda_j - \lambda_i)(\psi_i, \psi_j)_{\rm PT} &= 0. \end{aligned} \tag{2.67}$$

Thus if $\lambda_i - \lambda_j \neq 0$ then $(\psi_i, \psi_j)_{\rm PT} = 0$; in other words if two eigenvectors of H have distinct eigenvalues, they are orthogonal under the PT inner product.

2.4 New Hamiltonians

Now that we have specified the criteria that must be met in order for a Hamiltonian to be valid for either T_{even} or T_{odd} PT quantum mechanics, let us illustrate these principles by constructing the simplest non-trivial examples. For the even case the simplest example has $N = 2$ for the even case and $N = 4$ for the odd case; the two-level model for the even case has been discussed before in ref [14].

2.4.1 T_{even} Hamiltonians

For the even case the most general 2×2 Hamiltonian matrix that meets all the conditions of PT quantum mechanics is

$$H = \begin{pmatrix} a & ib \\ ib^* & -a \end{pmatrix} \tag{2.68}$$

Here a and b are real numbers and we have imposed the additional condition that H is traceless for simplicity [3]. Note that for $b \neq 0$ this matrix is explicitly non-Hermitian. It is instructive to compare eq (2.68) to the most general two-level Hermitian Hamiltonian that is invariant under even time reversal:

$$H = \begin{pmatrix} a & b \\ b^* & -a \end{pmatrix} \tag{2.69}$$

Clearly, PT quantum mechanics opens up a new class of Hamiltonians physicists can analyze.

The eigenvalues of the H in eq(2.68) are $\pm\sqrt{a^2 - b^2}$. Thus PT is unbroken only for $a^2 > b^2$. If this condition is satisfied the Hamiltonian H may be parametrized as $a = \rho\cosh(\chi)$ and $b = \rho\sinh\chi$ where $\rho > 0$ and $-\infty < \chi < \infty$. This parametrization applies for $a > 0$ which we will assume hereafter. The case $a < 0$ can be parametrized and analyzed in exactly the same way. The eigenmatrix is

$$U = \begin{pmatrix} \cosh\chi/2 & \sinh\chi/2 \\ i\sinh\chi/2 & i\cosh\chi/2 \end{pmatrix} \tag{2.70}$$

Here the first column corresponds to the eigenvector with positive eigenvalue ρ and

[3] If H has a trace it can always be written as the trace times the identity plus a traceless part. Note that the trace term does not affect the eigenvectors and shifts all the eigenvalues by a constant value. Thus the effects of the trace can be trivially incorporated.

the second to the negative eigenvalue $-\rho$; note that the eigenvectors have the PT invariant form in eq (2.35). It is easy to verify that the positive eigenvector also has positive PT norm; the negative has negative norm. Thus the operator C is simply the normalized Hamiltonian (*i.e.* H divided by the magnitude of the eigenvalues $\sqrt{a^2 - b^2}$),

$$C = \begin{pmatrix} \cosh\chi & i\sinh\chi \\ i\sinh\chi & -\cosh\chi \end{pmatrix} \tag{2.71}$$

Finally the most general operator A_{even} that corresponds to an observable by virtue of being CPT self-adjoint when T is even is given by

$$A = \begin{pmatrix} A_0 + A_3 - iA_1\tanh\chi & A_1 - iA_2 + iA_3\tanh\chi \\ A_1 + iA_2 + iA_3\tanh\chi & A_0 - A_3 + iA_1\tanh\chi \end{pmatrix} \tag{2.72}$$

Note that in the limit $\chi \to 0$, the most general observable is simply a Hermitian matrix; in the same limit the Hamiltonian H becomes Hermitian as well.

2.4.2 T_{odd} Hamiltonians

Finally let us consider the simplest non-trivial example of PT quantum mechanics for the case of odd time reversal symmetry with $N = 4$. The most general traceless Hamiltonian matrix that meets the criteria for T_{odd} is given by

$$H = \begin{pmatrix} a & ib \\ ib^\dagger & -a \end{pmatrix} \tag{2.73}$$

which is identical to the form of 2.68 except now b is a real quaternion, $b_0\sigma_0 + ib_1\sigma_1 + ib_2\sigma_2 + ib_3\sigma_3$, and a is a real quaternion proportional to the identity, $a = a_0\sigma_0$. It is instructive to compare this Hamiltonian to the most general four-level Hermitian

Hamiltonian that is invariant under odd time-reversal:

$$H = \begin{pmatrix} a & b \\ b^\dagger & -a \end{pmatrix} \tag{2.74}$$

i.e. eq(2.73) with the pure imaginary off-diagonal quaternions replaced by pure real quaternions, $ib \to b$. The eigenvalues of the PT invariant Hamiltonian are $\pm\sqrt{a^2 - b^2}$ where $a^2 = a_0^2$ and $b^2 = b_0^2 + b_1^2 + b_2^2 + b_3^2$ denote the magnitudes of the quaternions a and b. Thus PT is unbroken only for $a^2 > b^2$. So long as this condition is met (and $a_0 > 0$; the case $a_0 < 0$ can be analyzed similarly) we can parametrize the PT Hamiltonian by writing $a_0 = \cosh\chi$ and adopting polar coordinates $(\sinh\chi, \varphi, \theta, \phi)$ in the four dimensional space of the components of b so that $b_0 = \sinh\chi\cos\varphi, b_3 = \sinh\chi\sin\varphi\cos\theta, b_1 = \sinh\chi\sin\varphi\sin\theta\cos\phi$ and $b_2 = \sinh\chi\sin\varphi\sin\theta\sin\phi$. In terms of this parametrization the eigenmatrix has the form

$$U = \begin{pmatrix} q\cosh\chi/2 & q\sinh\chi/2 \\ iqp\sinh\chi/2 & iqp\cosh\chi/2 \end{pmatrix} \tag{2.75}$$

Here q is the real quaternion corresponding to a rotation about the $n_x = \sin\phi, n_y = -\cos\phi, n_z = 0$ axis by an angle of θ; and $p = \exp(-i\varphi\sigma_z)$, a rotation about the z -axis by an angle 2φ. The first two columns correspond to the positive energy PT doublet; the second two to the negative energy doublet. It is easy to verify that the positive doublet also has positive PT norm; the negative has negative norm. Thus the operator C coincides with the normalized Hamiltonian (*i.e.* H divided by $\sqrt{a^2 - b^2}$). Finally, the most general operator A_{odd} that corresponds to an observable by virtue of being CPT self-adjoint is

$$A_{odd} = \begin{pmatrix} q & 0 \\ 0 & qp \end{pmatrix} A \begin{pmatrix} q^\dagger & 0 \\ 0 & p^\dagger q^\dagger \end{pmatrix} \tag{2.76}$$

where A is still given by eq (2.72) but with A_0, A_1, A_2 and A_3 now interpreted as arbitrary 2×2 Hermitian matrices.

It is worth recalling that in conventional quantum mechanics a variety of complicated quantum mechanical problems can be truncated to a two level model [21]. Thus the two and four level models presented here should be regarded not merely as toy models but as effective Hamiltonians that can be used as the basis for further investigation of the quantum dynamics of PT quantum systems.

Chapter 3

Relativistic Non-Hermitian Quantum Mechanics and the Non-Hermitian Dirac Equation

In his seminal work on relativistic quantum mechanics Dirac set out to discover a wave equation that was first order in space and time derivatives and consistent with special relativity [16]. A key assumption made by Dirac was that the corresponding Hamiltonian would be Hermitian. In this way Dirac was led to his celebrated equation which predicted antimatter and describes both electrons and quarks. Should it turn out to describe neutrinos as well, the Dirac theory would govern all known fermionic matter in nature.

Remarkably, as we will show in this chapter, the Dirac equation in its fundamental representation is not unique to Hermitian quantum mechanics/quantum field theory. By relaxing the assumption of Hermiticity and adopting instead the principles of PT quantum mechanics outlined in the previous chapter, we do not modify the Dirac equation at all. The fundamental representation of the Dirac equation emerges completely in tact, identical in every aspect to the Dirac equation as derived from Hermi-

tian theory. This is a most intriguing result as it suggests Hermiticity is less crucial of an assumption than one might think (given that it is one of the fundamental tenets of quantum mechanics). It also puts non-Hermitian quantum mechanics 'on the map' in the sense that the theory can lay claim to the entire body of physical phenomena that are successfully described by the Hermitian Dirac equation. An even more intriguing feature of non-Hermitian quantum mechanics is that higher dimensional representations of the Dirac equation, which ordinarily decouple into independent fermions in Hermitian Dirac theory, here describe new types of particles, with exciting properties forbidden within the Standard Model.

3.1 Dirac Hamiltonian

The Schrödinger equation states that the time evolution of a wave function is dictated by the Hamiltonian that governs the system:

$$i\frac{\partial\psi}{\partial t} = H\psi. \tag{3.1}$$

(Here and in the rest of the document we work in units where $\hbar = 1$.) A free particle of mass m has $H = \mathbf{p}^2/2m$; suppose we take ψ to be a plane wave eigenstate of H and $\mathbf{p} = -i\nabla$, then the energy E where $H\psi = E\psi$ is given by $E = p^2/2m$. For each energy eigenvalue there are two degenerate states, one with momentum $\mathbf{p}$ and one with momentum -$\mathbf{p}$, which of course correspond to a right or left moving particle.

The principles of special relativity give a different energy for a massive relativistic particle, namely $E = \sqrt{p^2 + m^2}$. Dirac sought to construct a Hamiltonian that would give the relativistically correct energy eigenvalue when inserted in the Schrödinger equation; in 1927 he postulated the following Hamiltonian

$$H_D = -i\boldsymbol{\alpha}\cdot\nabla + \beta \tag{3.2}$$

where $\boldsymbol{\alpha}$ and β are matrices that satisfy the 'Dirac algebra'

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij}, \quad \{\alpha_i, \beta\} = 0. \tag{3.3}$$

To ensure the Hermiticity of H_D Dirac assumed the $\boldsymbol{\alpha}$ and β matrices were Hermitian. In addition he assumed that $\beta^2 = \mathbb{I}$.

To see that (3.2) gives the correct energy eigenvalues, square the Hamiltonian:

$$H_D^2 \psi = E^2 \psi, \tag{3.4}$$

$$(\ (\boldsymbol{\alpha} \cdot \mathbf{p}\ )^2 + \beta^2 m^2 + \ (\boldsymbol{\alpha} \cdot \mathbf{p}\ )\beta m + \beta m\ (\boldsymbol{\alpha} \cdot \mathbf{p}\ )\ )\psi = E^2 \psi. \tag{3.5}$$

Making use of eq(3.3), this reduces to

$$(p^2 + m^2)\psi = E^2 \psi \tag{3.6}$$

or $E = \pm\sqrt{p^2 + m^2}$, which, in the positive case, is the energy of a free relativistic particle of mass m. The negative case of course is what led Dirac to propose the existence of antiparticles, in order to explain how a free particle could have negative energy.

3.1.1 Properties of the Dirac Algebra

Matrices of different dimensionality can satisfy (3.3); we will refer to these as different representations of the Dirac equation.

Pauli Matrices

The simplest representation is 2-dimensional, and there are two of these: $\alpha_i \to \pm\sigma_i$ where σ_i are the Pauli matrices. It is not surprising that the Pauli matrices show

up here given that they form a basis for all traceless Hermitian 2×2 matrices with real coefficients, and with the inclusion of the 2×2 identity matrix $\mathbb{I}_{2\times 2} \equiv \sigma_0$ they form a basis for all 2×2 matrices. The two 2-d representations are referred to as the left-handed ($\alpha_i \to \sigma_i$) and right-handed ($\alpha_i \to -\sigma_i$)Weyl representations. For the 2-d representation, we must have $\beta = 0$, because no 2×2 matrix anti-commutes with all three σ_i:

let A be an arbitrary 2×2 matrix,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{3.7}$$

and recall the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.8}$$

If $\{\sigma_1, A\} = 0$ then

$$\begin{pmatrix} b+c & a+d \\ b+c & a+d \end{pmatrix} = 0 \quad \Rightarrow c = -b, \;\; d = -a \tag{3.9}$$

but if $\{\sigma_2, A\} = 0$ then

$$\begin{pmatrix} 2ib & 0 \\ 0 & 2ib \end{pmatrix} = 0 \quad \Rightarrow b = 0 \tag{3.10}$$

and if $\{\sigma_3, A\} = 0$ then

$$\begin{pmatrix} 2a & 0 \\ 0 & 2a \end{pmatrix} = 0 \quad \Rightarrow a = 0. \tag{3.11}$$

In fact there are a number of properties of the σ_i that we will use heavily in this chapter; we pause now to identify and prove some of them for easy reference later. Similar to the proof that no 2×2 matrix anti-commutes with all three σ_i, one can show:

(a) if a matrix commutes with all of the σ_i then it must be diagonal;

(b) if $A\sigma_i^* = \sigma_i A$, then $A = 0$,

(c) if $A\sigma_i^* = -\sigma_i A$ then: A must have the form $A = be_2$ where b is a complex number and $e_2 = i\sigma_2$,or more explicitly

$$e_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3.12}$$

The left and right handed Weyl represention are the only two choices for 2-d matrices that satisfy the Dirac algebra; any other 2-d representation is unitarily equivalent to a Weyl representation:

if α_i are 2×2 Hermitian matrices and $\{\alpha_i, \alpha_j\} = 2\delta_{ij}$ then $\alpha_i = U\sigma_i U^\dagger$ or $\alpha_i = -U\sigma_i U^\dagger$ where U is a unitary matrix.

Because we are interested in non-Hermitian representations as well, we note that

if α_i are 2×2 matrices and $\{\alpha_i, \alpha_j\} = 2\delta_{ij}$ then $\alpha_i = V\sigma_i V^{-1}$ or $\alpha_i = -V\sigma_i V^{-1}$ where V is an invertible matrix.

A simple way to obtain higher dimensional representations of the Dirac algebra is to construct direct sums of Weyl representations. For example we construct a 4-d rep-

resentation by pairing a left and a right Weyl representation; the most straightforward way to do this is

$$\alpha_i \to \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix}. \tag{3.13}$$

More general 4×4 representations can be obtained by

$$\alpha_i = \begin{pmatrix} V\sigma_i V^\dagger & 0 \\ 0 & -W\sigma_i W^\dagger \end{pmatrix} \quad \text{or} \quad \alpha_i = \begin{pmatrix} V\sigma_i V^{-1} & 0 \\ 0 & -W\sigma_i W^{-1} \end{pmatrix} \tag{3.14}$$

for the Hermitian and non-Hermitian cases respectively. Along the same lines one can show that β follows similar constraints. The 2×2 case requires β=0 , in the 4-d representation all valid choices of β are unitarily equivalent to

$$\beta = m \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.15}$$

The proof of these statements of unitary equivalence are straightforward but we will not pause to show them all here.

Lorentz Invariance

The last aspect of the Dirac algebra that we will explore in this section is the connection to Lorentz invariance. It turns out that in addition to giving the correct energy eigenvalues, the Dirac algebra ensures a Lorentz invariant theory.

Lorentz invariance can be demonstrated in a number of ways; here will take the brute force approach of showing that eigenfunctions of the Hamiltonian that undergo a Lorentz transformation remain eigenfunctions of the same Hamiltonian. Suppose $u \exp(i\mathbf{p} \cdot \mathbf{r})$ is an eigenfunction of H_D with energy E:

$$(\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m)u = Eu. \tag{3.16}$$

Suppose the state u' related to u by a Lorentz boost along the x-axis: $u' = \Lambda^{-1}u$, where $\Lambda = \exp(-iK_x\zeta)$, ζ is the rapidity parameter and we posit[1] that the generators of boosts and rotations are themselves the α_i :

$$K_x = \frac{i}{2}\alpha_x, \quad J_x = -\frac{i}{2}\alpha_y\alpha_z, \quad \text{etc.} \tag{3.17}$$

To see if the theory is Lorentz invariant we must determine whether u' is an eigenfunction of H_D. Plugging u' into eq (3.16),

$$(\boldsymbol{\alpha}\cdot\mathbf{p} + \beta m)\Lambda u' = E\Lambda u'. \tag{3.18}$$

Multiplying by Λ, we have

$$\Lambda(\boldsymbol{\alpha}\cdot\mathbf{p} + \beta m)\Lambda u' = E\Lambda^2 u' \tag{3.19}$$

$$\Lambda(\alpha_x p_x + \alpha_y p_y + \alpha_z p_z)\Lambda u' + \Lambda\beta m\Lambda u' = E\Lambda^2 u' \tag{3.20}$$

Noting that

$$\Lambda = \exp\left(\frac{\zeta}{2}\alpha_x\right) = \cosh\frac{\zeta}{2}\mathbb{I} + \alpha_x \sinh\frac{\zeta}{2} \tag{3.21}$$

and using eq(3.3), we see that

$$\Lambda\alpha_x\Lambda = \alpha_x\cosh\zeta + \sinh\zeta\ \mathbb{I}. \tag{3.22}$$

Because α_x anticommutes with α_y, α_z, and β, those components do not change under

[1] More on this choice in section 3.2.1.

the boost:

$$\begin{aligned} \Lambda\alpha_y\Lambda &= \alpha_y \\ \Lambda\alpha_z\Lambda &= \alpha_z \\ \Lambda\beta\Lambda &= \beta. \end{aligned} \tag{3.23}$$

Finally, it is useful to note that

$$\Lambda^2 = \cosh\zeta\,\mathbb{I} + \alpha_x \sinh\zeta \tag{3.24}$$

Using this to rewrite eq(3.20), we have

$$(\alpha_x(p_x cosh\zeta - E sinh\zeta) + \alpha_y py + \alpha_z p_z)u' + \beta m u' = E(\cosh\zeta\mathbb{I} - p_x \sinh\zeta)u' \tag{3.25}$$

Recognizing that the momentum and energy of the boosted state u' will have the standard form,

$$\begin{aligned} E' &= E\cosh\zeta - p_x \sinh\zeta \\ p'_x &= p_x \cosh\zeta - E\sinh\zeta \\ p'_y &= p_y \\ p'_z &= p_z, \end{aligned} \tag{3.26}$$

and writing eq(3.25) in terms of these, we have

$$\begin{aligned} (\alpha_x p'_x + \alpha_y p' y + \alpha_z p'_z)u' + \beta m u' &= (E\cosh\zeta - p_x\sinh\zeta)u' \\ (\boldsymbol{\alpha}\cdot\mathbf{p}' + \beta m)u' &= E'u' \end{aligned} \tag{3.27}$$

Thus, by virtue of the Dirac algebra, the theory is Lorentz invariant. Note that the proof did not assume anything about the specific form of the $\boldsymbol{\alpha}$ and β matrices, nor their dimensionality, nor that they were Hermitian, only that they obey the Dirac algebra.

3.1.2 Fundamental Representation

Armed with these properties of the $\boldsymbol{\alpha}$, we now construct the standard Hermitian Dirac equation as it provides a basis for the rest of the models we discuss. Typically we think of 'the Dirac equation' as the equations of motion for a Dirac fermion,

$$i\sigma^\mu \partial_\mu \psi_L = m\psi_R \tag{3.28}$$

$$i\bar{\sigma}^\mu \partial_\mu \psi_R = m\psi_L \tag{3.29}$$

where $\bar{\sigma}^\mu = -\sigma^\mu$ for $\mu = 1, 2, 3$ and $\bar{\sigma}^\mu = \sigma^\mu$ for $\mu = 0$. These equations of motion arise from the Lagrangian

$$\mathcal{L}_{\text{Dirac}} = \psi_L^\dagger i\sigma^\mu \partial_\mu \psi_L + \psi_R^\dagger i\bar{\sigma}^\mu \partial_\mu \psi_R + m\psi_L^\dagger \psi_R + m\psi_R^\dagger \psi_L \tag{3.30}$$

by the standard techniques of variational calculus provided we treat ψ and $\psi^\dagger$ as independent quantities. The physical interpretation suggested by (3.29) is that a Dirac fermion is a pair of Weyl spinors coupled by a mass term. Recall the 2-d Weyl representation requires β=0, so a Weyl particle can be thought of as a (fictional) massless 2-component spinor ψ_L or ψ_R.

Although the form of the Dirac equation we use here, eq (3.2), is closer to the way

Dirac originally constructed relativistic quantum mechanics, it bears little notational resemblance to the modern way of writing the Dirac equation (3.29). To see the modern notation arise from (3.2), consider the most straightforward[2] 4×4 representation of α_i and β:

$$\alpha_i \to \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \tag{3.31}$$

and

$$\beta = m \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.32}$$

The Schrödinger equation states

$$H\psi = i\frac{\partial \psi}{\partial t}. \tag{3.33}$$

Assume ψ is comprised of a left- and right-handed Weyl spinor:

$$\psi \to \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$$

Using $H_D = -i\boldsymbol{\alpha} \cdot \nabla + \beta$ and the representations (3.31) and (3.32), we have

$$-i \begin{pmatrix} \sigma \cdot \nabla & 0 \\ 0 & -\sigma \cdot \nabla \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} + \begin{pmatrix} 0 & m \\ m & 0 \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = i\frac{\partial}{\partial t} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{3.34}$$

$$\begin{pmatrix} -i(\sigma \cdot \nabla)\psi_L \\ -i(-\sigma \cdot \nabla)\psi_R \end{pmatrix} + \begin{pmatrix} m\psi_R \\ m\psi_L \end{pmatrix} = i\frac{\partial}{\partial t} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{3.35}$$

[2] As stated in previous section any other valid choice for α_i and β can be unitarily transformed into these.

With a little rearranging,

$$\begin{aligned} i(\sigma \cdot \nabla)\psi_L + i\frac{\partial \psi_L}{\partial t} &= m\psi_L \\ i(-\sigma \cdot \nabla)\psi_R + i\frac{\partial \psi_R}{\partial t} &= m\psi_R \end{aligned} \tag{3.36}$$

or,

$$i\sigma^\mu \partial_\mu \psi_L = m\psi_R \tag{3.37}$$

$$i\bar{\sigma}^\mu \partial_\mu \psi_R = m\psi_L \tag{3.38}$$

Thus, the familiar Dirac equation is resolved from the Hamiltonian (3.2) and the representations (3.31) and (3.32). We call this the 'fundamental' representation because it describes the fermions originally proposed by Dirac, *ie* a pair of Weyl spinors coupled by mass m, with dispersion appropriate to a massive relativistic particle, $E = \pm\sqrt{p^2 + m^2}$. Physically the Dirac fermion is a particle-antiparticle pair, the particle being the spinor with the positive eigenvalue and the antiparticle being the spinor with the negative eigenvalue.

3.1.3 Dirac quartet

In the previous section, we paired up two Weyl representations via direct sum and obtained a far more interesting particle than that described by either of the individual Weyl representations. Naturally one might extend this procedure to higher dimensional representations and ask whether, for example, a pair of Dirac fermions combine to form still more interesting particles. In this section we address this question, assuming the rules of ordinary Hermitian quantum mechanics.

The direct sum of two fundamental representations (3.31) and (3.32) is of course comprised of four Weyl fermions, 2 left-handed and 2 right-handed, so we call this model the Dirac 'quartet'. Keep in mind that this is actually an 8-dimensional representation, as we will often use various 2×2 matrix abbreviations to make the formulae more manageable.

For the 8×8 representation we have

$$\alpha_i \to \begin{bmatrix} \sigma_i & 0 & 0 & 0 \\ 0 & \sigma_i & 0 & 0 \\ 0 & 0 & -\sigma_i & 0 \\ 0 & 0 & 0 & -\sigma_i \end{bmatrix}. \tag{3.39}$$

and the most general choice of the mass matrix is

$$\beta = \begin{pmatrix} 0 & M \\ M^\dagger & 0 \end{pmatrix} \quad \text{where} \quad M = \begin{pmatrix} m_1\sigma_0 & m_2\sigma_0 \\ m_3\sigma_0 & m_4\sigma_0 \end{pmatrix} \tag{3.40}$$

where the m's are arbitrary complex numbers.

In the case of a single Dirac fermion, the particle-antiparticle pair are inextricably linked through the mass matrix β; there is no way to decouple ψ_L from ψ_R unless $\beta = 0$. From the more complicated form of the β matrix in the 8-d case (eq (3.40)) its tempting to assume the particle described by the Dirac quartet is also an inextricable merger of left and right-handed Weyl particles coupled by the mass matrix. However, a suitable unitary transformation shows that the Dirac quartet decouples into two independent Dirac fermions. To see this, we employ the process of singular value decomposition.

Singular value decomposition is essentially the statement that an $n \times n$ matrix M can be written as $M = U\mu V^\dagger$ where μ is the diagonal matrix comprised of the square roots of the eigenvalues of $M^\dagger M$ (or $MM^\dagger$), which are all real and positive:

$$\mu = \begin{pmatrix} \mu_1 & & & \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_n \end{pmatrix} \tag{3.41}$$

and $U^\dagger$ is the eigenmatrix of $MM^\dagger$ and $V^\dagger$ is the eigenmatrix of $M^\dagger M$. The most general form of β is eq(3.40), so in this case $M^\dagger M$ has two-fold degenerate eigenvalues μ_1^2 and μ_2^2. Thus

$$\mu = \begin{pmatrix} \mu_1\sigma_0 & 0 \\ 0 & \mu_2\sigma_0 \end{pmatrix} \tag{3.42}$$

We can write the eigenmatrices U and V as

$$U = \begin{pmatrix} u_1\sigma_0 & u_2\sigma_0 \\ u_3\sigma_0 & u_4\sigma_0 \end{pmatrix} \quad , \quad V = \begin{pmatrix} v_1\sigma_0 & v_2\sigma_0 \\ v_3\sigma_0 & v_4\sigma_0 \end{pmatrix} \tag{3.43}$$

where u_i and v_i are complex numbers. Now

$$\beta = \begin{pmatrix} 0 & M \\ M^\dagger & 0 \end{pmatrix} = \begin{pmatrix} 0 & U\mu V^\dagger \\ V\mu U^\dagger & 0 \end{pmatrix} \tag{3.44}$$

and if we perform a unitary transformation $W^\dagger \beta W$ where

$$W = \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} \tag{3.45}$$

this transforms β in the following way:

$$W^\dagger \beta W = \begin{pmatrix} 0 & U^\dagger M V \\ V^\dagger M^\dagger U & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mu \\ \mu & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & \mu_1\sigma_0 & 0 \\ 0 & 0 & 0 & \mu_2\sigma_0 \\ \mu_1\sigma_0 & 0 & 0 & 0 \\ 0 & \mu_2\sigma_0 & 0 & 0 \end{pmatrix} \tag{3.46}$$

Note that this transformation leaves the α_i unchanged: $W^\dagger \alpha_i W = \alpha_i$. We now introduce a second unitary transformation,

$$Y = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.47}$$

This transformation exchanges a left-handed representation for a right-handed one in the α_i:

$$\begin{aligned} \tilde{\alpha_i} &= Y^\dagger \alpha_i Y \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sigma_i & 0 & 0 & 0 \\ 0 & 0 & \sigma_i & 0 \\ 0 & -\sigma_i & 0 & 0 \\ 0 & 0 & 0 & -\sigma_i \end{pmatrix} = \begin{pmatrix} \sigma_i & 0 & 0 & 0 \\ 0 & -\sigma_i & 0 & 0 \\ 0 & 0 & \sigma_i & 0 \\ 0 & 0 & 0 & -\sigma_i \end{pmatrix} \end{aligned}$$

and rearranges the non-zero entries of the mass matrix:

$$\begin{aligned}\tilde{\beta} &= Y^\dagger \beta Y \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & \mu_1 & 0 & 0 \\ 0 & 0 & 0 & \mu_2 \\ \mu_1 & 0 & 0 & 0 \\ 0 & 0 & \mu_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mu_1 & 0 & 0 \\ \mu_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu_2 \\ 0 & 0 & \mu_2 & 0 \end{pmatrix}\end{aligned}$$

So the action of the transformed Hamiltonian $H = -i\tilde{\boldsymbol{\alpha}} \cdot \nabla + \tilde{\beta} m$ on a quartet eigenstate

$$\begin{pmatrix} \begin{bmatrix} \sigma \cdot \nabla & \mu_1 \\ \mu_1 & -\sigma \cdot \nabla \end{bmatrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \begin{bmatrix} \sigma \cdot \nabla & \mu_2 \\ \mu_2 & -\sigma \cdot \nabla \end{bmatrix} \end{pmatrix} \begin{pmatrix} \psi_{L_1} \\ \psi_{R_1} \\ \psi_{L_2} \\ \psi_{R_2} \end{pmatrix} = -\frac{\partial}{\partial t} \begin{pmatrix} \psi_{L_1} \\ \psi_{R_1} \\ \psi_{L_2} \\ \psi_{R_2} \end{pmatrix} \tag{3.48}$$

Where we have inserted brackets to highlight the partitioning into two independent fermions: the upper left corner of the transformed α and β match up to describe one Dirac fermion of mass μ_1 and the lower right corners match up to form a second Dirac fermion of mass μ_2. We leave it as an exercise to show that the wave function ansatz,

$$\begin{pmatrix} \psi_{L1} \\ \psi_{L2} \\ \psi_{R1} \\ \psi_{R2} \end{pmatrix} \tag{3.49}$$

decouples under the two transformations as

$$Y^{\dagger}\left(W^{\dagger}\psi\right) \sim \begin{pmatrix} \psi_{L1} \\ \psi_{R1} \\ \psi_{L2} \\ \psi_{R2} \end{pmatrix}. \tag{3.50}$$

A similar procedure can be applied to , for example, 12-d and higher representations; thus there are no more 'fundamental' particles contained in higher dimensional representations of the Dirac algebra– only concatenations of independent 4-d fermions. However, while the Dirac quartet does not describe a new type of fundamental fermion, it can be thought of as a toy model for 2 generations of Dirac neutrinos.

3.2 PT Dirac Equation

We now turn to the construction of the non-Hermitian Dirac equation. Ordinary Hermitian quantum mechanics and quantum field theory successfully describe the behavior of elementary particles –will the repercussions of having relaxed Hermiticity propagate through the theory and result in a different behavior for fermions? If so then we could use the precision with which the Hermitian theory is known to constrain any deviations therefrom. This was our original motivation in constructing the PT analogue of the Dirac equation. As we will now demonstrate, the fundamental representation of the Dirac equation is not exclusive to Hermitian quantum mechanics. Constructing the analogous 4-d representation using the principles of PT quantum mechanics one obtains exactly the same theory as the Dirac equation we all know and love. That the Dirac equation emerges from the generalization to PT quantum mechanics completely intact is a remarkable finding, as it suggests there is nothing

special about Hermitian quantum mechanics and that the assumption of Hermiticity is rather arbitrary. A carefully constructed theory that guarantees real eigenvalues, unitary time evolution, and the other criteria discussed in Chapter 1 is sufficient to give the famous Dirac equation. Furthermore, parity and time-reversal symmetry are physically motivated and eventually enforced even in a Hermitian theory. PT quantum mechanics gives these principles, rather than convenient linear algebraic properties, a prominent role in the development of the theory. And in the case of Dirac fermions, this comes at no expense to known experimental results.

3.2.1 Construction and Conditions

The construction of the PT Dirac equation proceeds from a few simple assumptions. We keep the Hamiltonian as eq (3.2) but now allow the α_i and β matrices to be non-Hermitian. We keep the action of time reversal and parity the same as they are in ordinary fermionic theory, namely $T^2 = -1$, $P^2 = 1$, and

$$\begin{aligned} P\psi(\mathbf{r}) &= S\psi(-\mathbf{r}) \\ T\psi(\mathbf{r}) &= Z\psi^*(\mathbf{r}) \end{aligned} \tag{3.51}$$

where S is a $2m \times 2m$ matrix that exchanges the left and right handed components of the 2m-dimensional wave function, and $Z = i\sigma_y \mathbb{I}_{m\times m}$. Many matrices satisfy these general conditions, but rather than choose specific representations at the onset, we keep the forms of $\boldsymbol{\alpha}$, β, S and Z open and allow the principles of special relativity and PT quantum mechanics to fix these matrices.

Determining the specific form of these matrices is tedious and not very enlightening, however it is an extremely important step in the construction of the PT Dirac theory. By letting the principles of special relativity and Lorentz invariance guide us

to the correct representations for these matrices, we are not so much *choosing* what the PT Dirac equation is as we are letting Nature tell us what fermionic Dirac theory would be like in a world where said principles are still upheld, but the assumption of Hermiticity is not. As we will show in this chapter, these conditions are constraining enough so as to keep the non-Hermitian 4-d representation precisely the same as the fundamental representation in Hermitian theory, but are liberal enough so that the higher dimensional representations (specifically, the 8-d representation) allow new physics to creep through.

Conditions

In this section we first enumerate the conditions and then derive them.

(i) $\boldsymbol{\alpha}$ and β are assumed to obey the Dirac algebra eq (3.3) to ensure relativistic dispersion. From the $\boldsymbol{\alpha}$'s we construct the boost and rotation generators according to eq (3.26).

(ii) For parity and time reversal to be compatible with the boosts and rotations we need $Z\alpha_i^* = -\alpha_i Z$ and $\{S, \alpha_i = 0\}$.

(iii) We assume P and T commute, leading to $ZZ^* = -1$, $S^2 = 1$ and $SZ = ZS^*$ respectively.

(iv) We assume the Hamiltonian commutes with the combined operation PT , leading to $\alpha_i SZ = SZ\alpha_i^*$ and $\beta SZ = SZ\beta^*$.

(v) We impose self-duality, to ensure that the eigenvectors of H_D will be orthogonal under the PT inner product, as explained in detail in the previous chapter. This amounts to imposing $\alpha_i = Z^T \alpha_i^T Z^\dagger$ and $\beta = -Z^T \beta^T Z^\dagger$.

(vi) We require that H_D have unbroken PT symmetry, which is essentially the same as requiring that the energy eigenvalues be real, as explained in the previous chapter.

In the derivation of these condition we will assume a 4-d representation for illustration purposes. The generalization to representations of different dimensionality is trivial.

Condition (i)

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij}, \quad \{\alpha_i, \beta\} = 0 \text{ and } K_x = \tfrac{i}{2}\alpha_x, \quad J_x = -\tfrac{i}{2}\alpha_y\alpha_z, \quad \text{etc.}$$

It was Dirac's remarkable discovery that the matrices that make up the kinematic part of the Dirac Hamiltonian $\alpha_x, \alpha_y, \alpha_z$ can also play the role of generators of boosts and rotations, and in doing so, they form a representation of the Lorentz algebra. More specifically, we assume

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij}. \tag{3.52}$$

If we assign

$$K_x \to \frac{i}{2}\alpha_x \quad K_y \to \frac{i}{2}\alpha_y \quad K_z \to \frac{i}{2}\alpha_z \tag{3.53}$$

and

$$J_x \to -\frac{i}{2}\alpha_y\alpha_z \quad J_y \to -\frac{i}{2}\alpha_z\alpha_x \quad J_z \to -\frac{i}{2}\alpha_x\alpha_y. \tag{3.54}$$

as the generators of boosts and rotations, they obey the Lorentz algebra:

$$\begin{aligned} &[J_x, J_y] = iJ_z, && [K_x, K_y] = 0, && [K_x, J_z] = -iK_y \\ &[K_x, K_y] = -iJ_z, && [K_x, J_y] = iK_z, && \ldots \quad \text{etc.} \end{aligned} \tag{3.55}$$

Given the generators, we construct boosts and rotations in the standard way, for example a boost along the x-axis is given by

$$\Lambda_x(\zeta) = \exp(-iK_x\zeta) = \exp\left(\frac{1}{2}\alpha_x\zeta\right) = \cosh\frac{\zeta}{2} + \alpha_x \sinh\frac{\zeta}{2}. \tag{3.56}$$

Explicit evaluation of the exponential is facilitated by use of the anti-commutation relation $\alpha_x^2 = 1$. Similarly a rotation about the z-axis may be written as

$$R_z(\theta) = \exp(-iJ_z\theta) = \exp\left[-i\frac{1}{2}\theta(-i\alpha_x\alpha_y)\right] = \cos\frac{\theta}{2} - \alpha_x\alpha_y \sin\frac{\theta}{2}; \tag{3.57}$$

here, the equality is facilitated by use of $(-i\alpha_x\alpha_y)^2 = 1$ which follows from the anti-commutation relations eq (3.52).

Armed with expressions for boosts and rotations in terms of the α_i let us investigate how they transform under parity and time-reversal. Under parity space co-ordinates are inverted but the time is left unchanged :

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = P \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \quad \text{where} \quad P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{3.58}$$

It is easy to verify that

$$\begin{aligned} P\Lambda_x(\zeta)P &= \Lambda_x(-\zeta), \\ PR_z(\theta)P &= R_z(\theta). \end{aligned} \tag{3.59}$$

Obviously, similar relations hold for boosts along other directions or rotations about other axes. In terms of the generators eq (3.59) may be written as

$$\begin{aligned} PK_xP &= -K_x \\ PJ_zP &= J_z. \end{aligned} \tag{3.60}$$

Evidently similar relations must hold for other boost and rotation generators.

Under time-reversal, $t \to -t$ but the spatial co-ordinates are unchanged:

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = T \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \quad \text{where} \quad T = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.61}$$

It is easy to verify that

$$\begin{aligned} T^{-1}\Lambda_x(\zeta)T &= \Lambda_x(-\zeta) \\ T^{-1}R_z(\theta)T &= R_z(\theta); \end{aligned} \tag{3.62}$$

Similar relations hold for boosts along other directions and rotations about other axes.

Condition (ii)

$$\{S, \alpha_i\} = 0 \text{ and } Z\alpha_i^* = -\alpha_i Z$$

This condition constrains the α_i so that they are compatible with boosts and rotations. Note that the action of P and T on a spinor plane wave is given by

$$\begin{aligned} Pu\exp(i\mathbf{k}\cdot\mathbf{r}) &= Su\exp(-i\mathbf{k}\cdot\mathbf{r}), \\ Tu\exp(i\mathbf{k}\cdot\mathbf{r}) &= Zu^*\exp(-i\mathbf{k}\cdot\mathbf{r}). \end{aligned} \tag{3.63}$$

Since parity is a linear operator and all spinor wave functions can be written as a superposition of plane waves, eq (3.63) fully defines the parity operator.

To derive the condition on S and α_i, recall that according to eq (3.60) parity and boost generators satisfy the relation $PK_xP = -K_x$. By applying both sides of this expression to a spinor plane wave, making use of eq (3.63) and the correspondence eq (3.56), we obtain $S\alpha_xS = -\alpha_x$. Using $S^2 = 1$, we find $S\alpha_x = -\alpha_xS$ or $\{S, \alpha_x\} = 0$. The x direction was chosen arbitrarily and the same analysis would hold for y and z, so we conclude $\{S, \alpha_i\} = 0$. According to eq (3.60), parity and rotation generators satisfy the relation $PJ_zP = J_z$. By a similar analysis of this relation we conclude that S and $\boldsymbol{\alpha}$ must obey the consistency condition

$$[S, \alpha_x\alpha_y] = 0 \tag{3.64}$$

and permutations thereof. However this is not an independent condition. Eq (3.64) follows as a consequence of $\{S, \alpha_i\} = 0$ by use of the identity $[A, BC] = \{A, B\}C - B\{A, C\}$. Turning now to the constraint between Z and α_i. T is anti-linear ($T\psi = Z\psi^*$) and odd ($T^2 = -1$), which implies $Z^* = -Z^{-1}$. So $T^{-1}\psi = -Z\psi^*$. Recall that

$T^{-1}\Lambda_x(\zeta)T = \Lambda(-\zeta)$, and using eq(3.56), we obtain

$$\begin{aligned} -Z\alpha_x^* Z^* &= -\alpha_x \\ Z\alpha_x^* &= -\alpha_x Z \end{aligned} \tag{3.65}$$

The same analysis would hold for boosts along the y and z directions, so we conclude $Z\alpha_i^* = -\alpha_i Z$.

Condition iii

$$S^2 = 1$$

$$ZZ^* = -1$$

$$SZ = ZS^*$$

Time reversal is odd and $P^2 = 1$ so eq (3.63) immediately leads to $S^2 = 1$ and $ZZ^* = -1$. We may sometimes find it convenient to express this latter condition as $Z^T Z^\dagger = -1$. Finally, since we impose that $[P, T] = 0$, it follows from eq (3.63) that $SZ = ZS^*$, which we may also write as $Z^T S^T = S^\dagger Z^T$.

Condition (iv)

$$\alpha_i SZ = SZ\alpha_i^*$$

$$\beta SZ = SZ\beta^*$$

These conditions follow immediately from $[H, PT] = 0$ and the action of P and T on a plane wave eq (3.63) .

Condition (v)

$$\alpha_i SZ = SZ\alpha_i^*$$

$$\beta SZ = SZ\beta^*$$

In Chapter 1 we define the PT inner product for finite dimensional vector spaces as

$$\begin{aligned}(\phi,\psi)_{PT} &\equiv -(PT\phi)^T Z^\dagger \psi, \\ &= -(SZ\phi^*)^T Z^\dagger \psi, \\ &= -\phi^\dagger Z^T S^T Z^\dagger \psi. \end{aligned} \tag{3.66}$$

From (3.63) we see the combined operation of PT on a plane wave is

$$PTu\exp(i\mathbf{k}\cdot\mathbf{r}) = SZu^*\exp(i\mathbf{k}\cdot\mathbf{r}). \tag{3.67}$$

The PT inner product of two plane waves is given by a natural generalization of eq (3.66),

$$\begin{aligned}(u\exp i\mathbf{k}\cdot\mathbf{r}, v\exp i\mathbf{p}\cdot\mathbf{r})_{PT} &= -\int d\mathbf{r}\, u^\dagger Z^T S^T Z^\dagger v \exp i\mathbf{k}\cdot\mathbf{r}\exp i\mathbf{p}\cdot\mathbf{r}, \\ &= -(2\pi)^3\delta(\mathbf{k}+\mathbf{p})u^\dagger Z^T S^T Z^\dagger v. \end{aligned} \tag{3.68}$$

Now $Z^T S^T Z^\dagger = S^\dagger Z^T Z^\dagger = -S^\dagger$ because $Z^T S^T = S^\dagger Z^T$ and $Z^T Z^\dagger = -1$.

Hence we obtain the final simplified expression for the PT inner product of plane-waves in a relativistic quantum theory:

$$(u\exp i\mathbf{k}\cdot\mathbf{r}, v\exp i\mathbf{p}\cdot\mathbf{r})_{PT} = (2\pi)^3\delta(\mathbf{k}+\mathbf{p})u^\dagger S^\dagger v. \tag{3.69}$$

As described in Chapter 1, self-duality is the condition that

$$(Hu\exp i\mathbf{k}\cdot\mathbf{r}, v\exp i\mathbf{p}\cdot\mathbf{r})_{PT} = (u\exp i\mathbf{k}\cdot\mathbf{r}, Hv\exp i\mathbf{p}\cdot\mathbf{r})_{PT} \tag{3.70}$$

Making use of eqs (3.2) and (3.69) it is easy to show that

$$\begin{aligned} S\alpha_i &= -\alpha_i S^\dagger, \\ S\beta &= \beta^\dagger S. \end{aligned} \tag{3.71}$$

Combining this with $[H, PT] = 0$ gives

$$\begin{aligned} \beta &= -Z^T \beta^T Z^\dagger, \\ \alpha_i &= Z^T \alpha_i^T Z^\dagger. \end{aligned} \tag{3.72}$$

This form is more convenient for our purpose than eq (3.71).

Condition (vi)

Condition (vi) is simply that H and PT have unbroken symmetry, which is described in Chapter 1. This condition essentially guarantees that the eigenvalues of H are real and does not place any additional constraints on S, Z, α_i, or β.

This concludes the derivation of the conditions that must be met by the PT Dirac theory in order to be consistent with special relativity and PT quantum mechanics. We note that conditions (i-iii) are also imposed in ordinary Dirac theory, and conditions (iv-vi) here substitute for Hermiticity.

3.3 Model 4

Now we are equipped to construct the analogue of the fundamental (4-d) representation of the Dirac equation in the framework of PT quantum mechanics. We call this theory 'Model 4'. Recall from section 3.1.1 that any set of 2×2 matrices that satisfy $\{\alpha_i, \alpha_j\} = 2\delta_{ij}$ has to be of the form $V\sigma_i V^{-1}$ or $-W\sigma_i W^{-1}$ where V and W are

invertible 2×2 matrices[3] In constructing Model 4 we begin by taking the $\boldsymbol{\alpha}$ matrices to be the direct sum of these two possibilities:

$$\alpha_i = \begin{bmatrix} V\sigma_i V^{-1} & 0 \\ 0 & -W\sigma_i W^{-1} \end{bmatrix}. \tag{3.73}$$

Next we turn to the choice of Z. There are three conditions that relate $\boldsymbol{\alpha}$ and Z, conditions (ii), (iv), and (v); we will use these conditions and the above ansatz for $\boldsymbol{\alpha}$ to specify Z. Let's assume Z is of the form

$$Z = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{3.74}$$

where A, B, C and D are 2×2 matrices. Imposing condition (ii) we find upon multiplying matrices that these 2×2 matrices must satisfy the following relations

$$\begin{aligned} V^{-1}AV^*\sigma_i^* &= -\sigma_i V^{-1}AV^*, \\ W^{-1}DW^*\sigma_i^* &= -\sigma_i W^{-1}DW^*, \\ V^{-1}BW^*\sigma_i^* &= \sigma_i V^{-1}DW^*, \\ W^{-1}CV^*\sigma_i^* &= \sigma_i W^{-1}CV^*. \end{aligned} \tag{3.75}$$

[3]If the matrices are Hermitian then V and W are unitary matrices, ie any 2×2 Hermitian representation of the Dirac algebra is unitarily equivalent to either the left handed or right handed Weyl representation.

Recall from section 3.1.1

$$\begin{aligned} M\sigma_i^* = -\sigma_i M \quad &\Rightarrow \quad M = be_2 \quad \text{where} \quad b \quad \in \mathbb{C} \\ M\sigma_i^* = \sigma_i M \quad &\Rightarrow \quad M = 0. \end{aligned} \tag{3.76}$$

Thus condition (ii) leads to the conclusion that $A = Vae_2V^{-1*}$, $D = Wde_2W^{-1*}$ where a and d are undetermined complex numbers, and $B = C = 0$. In other words

$$Z = \begin{pmatrix} Vae_2V^{-1*} & 0 \\ 0 & Wde_2W^{-1*} \end{pmatrix}. \tag{3.77}$$

Next we impose condition (iv) that $ZZ^* = -1$ using the form of Z above. This leads immediately to $|a|^2 = |d|^2 = 1$, so we can write a and d as

$$a = \exp i\phi_a, \quad d = \exp i\phi_d. \tag{3.78}$$

Finally we impose condition (v) that $\alpha_i = Z^T\alpha_i^T Z^\dagger$. Using $e_2^T = -e_2$ and $e_2\sigma_i^T e_2 = \sigma_i$, we can write

$$\begin{aligned} V^\dagger V\sigma_i &= \sigma_i V^\dagger V, \\ W^\dagger W\sigma_i &= \sigma_i W^\dagger W. \end{aligned} \tag{3.79}$$

From section 3.1.1 we know that if a matrix commutes with all three Pauli matrices then it must be diagonal, so $V^\dagger V$ must be diagonal. One can also argue that $V^\dagger V$ must have positive real eigenvalues, thus the diagonal entries of V must be positive. Hence we conclude

$$V^\dagger V = v^2\sigma_0, \quad W^\dagger W = w^2\sigma_0 \tag{3.80}$$

where v and w are positive real numbers. This implies that

$$V^{-1} = \frac{1}{v^2}V^\dagger, \quad W^{-1} = \frac{1}{w^2}W^\dagger. \tag{3.81}$$

and moreover that

$$\frac{1}{v}V \quad \text{and} \quad \frac{1}{w}W \tag{3.82}$$

are unitary matrices.

In conclusion, after imposition of the three conditions that interelate Z and $\boldsymbol{\alpha}$ we find that these matrices have the form

$$\alpha_i = \begin{pmatrix} v^{-1}V\sigma_i v^{-1}V^\dagger & 0 \\ 0 & -w^{-1}W\sigma_i w^{-1}W^\dagger \end{pmatrix} \tag{3.83}$$

and

$$Z = \begin{pmatrix} v^{-1}V \exp i\phi_a e_2 v^{-1}V^T & 0 \\ 0 & w^{-1}W \exp i\phi_d e_2 w^{-1}W^T \end{pmatrix} \tag{3.84}$$

These results can be summarized more elegantly by defining

$$U = \begin{pmatrix} v^{-1}V \exp i\frac{\phi_a}{2} & 0 \\ 0 & w^{-1}W \exp i\frac{\phi_d}{2} \end{pmatrix}. \tag{3.85}$$

Evidently U is unitary and we may write

$$\alpha_i = U \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} U^\dagger \tag{3.86}$$

and

$$Z = U \begin{pmatrix} e_2 & 0 \\ 0 & e_2 \end{pmatrix} U^T. \tag{3.87}$$

It follows from this representation that α_i *are* Hermitian, $\alpha_i = \alpha_i^\dagger$. Unlike Dirac, however, we did not explicitly impose this Hermiticity, it emerged as a consequence of the conditions of PT quantum mechanics. This already suggests that Model 4 will be equivalent to the fundamental Dirac fermion, but so far we have only looked at the conditions which relate Z and α_i. Now we turn to the specification of S.

There are three conditions that relate α_i and S, namely, conditions (ii), (iii) and (iv). Motivated by the form of α_i and Z, eqs (3.86) and (3.87) we write S in the form $S = U\mathbb{S}U^\dagger$ and rewrite conditions (d)-(g) in terms of $\mathbb{S}$. Condition (ii) simplifies to

$$\mathbb{S}\begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} + \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix}\mathbb{S} = 0. \tag{3.88}$$

Similarly condition (iii) simplifies to

$$\mathbb{S}^2 = 1 \tag{3.89}$$

and

$$\mathbb{S}\begin{pmatrix} e_2 & 0 \\ 0 & e_2 \end{pmatrix} = \begin{pmatrix} e_2 & 0 \\ 0 & e_2 \end{pmatrix}\mathbb{S}^*. \tag{3.90}$$

Finally condition (iv) becomes

$$\mathbb{S}\begin{pmatrix} e_2 & 0 \\ 0 & e_2 \end{pmatrix}\begin{pmatrix} \sigma_i^* & 0 \\ 0 & -\sigma_i^* \end{pmatrix} = \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix}\mathbb{S}\begin{pmatrix} e_2 & 0 \\ 0 & e_2 \end{pmatrix}. \tag{3.91}$$

Making use of $e_2^2 = -\sigma_0$ and $e_2\sigma_i^* e_2 = \sigma_i$ we can show that eq (3.91), is identical to eq (3.88), so the task now is to determine $\mathbb{S}$ by enforcing eqs (3.88), (3.89) and (3.90). To this end we write

$$\mathbb{S} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{3.92}$$

where a, b, c and d are arbitrary 2×2 matrices. Equation (3.88) leads to the conclusion that $\{a, \sigma_i\} = \{d, \sigma_i\} = 0$ while $[b, \sigma_i] = [c, \sigma_i] = 0$. Since there is no 2×2 matrix that anti-commutes with all three Pauli matrices, we must have $a = d = 0$. And a matrix that commutes with all three Pauli matrices must be proportional to the identity, so condition (ii) implies that the parity matrix has the form

$$\mathbb{S} = \begin{pmatrix} 0 & b_0\sigma_0 \\ c_0\sigma_0 & 0 \end{pmatrix} \tag{3.93}$$

where b_0 and c_0 are arbitrary complex numbers. Next we substitute the form eq (3.93) into condition (iii). This leads to the conclusion that $b_0c_0 = 1$. Finally we substitute the form eq (3.93) into condition (iii), which forces $b_0 = b_0^*$ and $c_0 = c_0^*$.

So it finally emerges that $\mathbb{S}$ has the form

$$\mathbb{S} = \begin{pmatrix} 0 & b_0\sigma_0 \\ b_0^{-1}\sigma_0 & 0 \end{pmatrix} \tag{3.94}$$

where b_0 is a non-zero real number, and the parity matrix S is given by $S = U\mathbb{S}U^\dagger$.

Now we must determine the mass matrix for Model 4. Here, the conditions that guide us are (i), (iv), and (v). Again motivated by the form of $\boldsymbol{\alpha}, Z$ and S for Model 4 we posit that

$$\beta = UBU^\dagger. \tag{3.95}$$

In terms of this ansatz condition (i) simplifies to

$$\begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} B + B \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} = 0. \tag{3.96}$$

Condition (i) simplifies to

$$\mathbb{S}\begin{pmatrix} e_2 & 0 \\ 0 & e_2 \end{pmatrix} B^* = B\mathbb{S}\begin{pmatrix} e_2 & 0 \\ 0 & e_2 \end{pmatrix}. \tag{3.97}$$

Finally condition (v) simplifies to

$$B = -\begin{pmatrix} e_2 & 0 \\ 0 & e_2 \end{pmatrix} B^T \begin{pmatrix} e_2 & 0 \\ 0 & e_2 \end{pmatrix}. \tag{3.98}$$

By the same reasoning that applied to condition (i) for the case of parity and the matrix S, we see that condition (i) constrains B to have the form

$$B = \begin{pmatrix} 0 & m_1\sigma_0 \\ m_2\sigma_0 & 0 \end{pmatrix} \tag{3.99}$$

where m_1 and m_2 are arbitrary complex numbers. Condition (ii) then implies that $m_1 = m_2 \equiv m$, and forces

$$b_0 m^* = \frac{1}{b_0} m \quad \text{and} \quad b_0 m = \frac{1}{b_0} m^*. \tag{3.100}$$

We have already shown b_0 is real, so these equations have two possible solutions: either $m = 0$ and b_0 is arbitrary, or m is non-zero and real and $b_0 = \pm 1$. The latter corresponds to the massive Model 4; the former to the massless. In this chapter we shall only concentrate on the massive case.

Now we have completely specified the α_i, β, S and Z matrices; we summarize them here for convenience:

$$
\begin{aligned}
\alpha_i &= U \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} U^\dagger \\
Z &= U \begin{pmatrix} e_2 & 0 \\ 0 & e_2 \end{pmatrix} U^T \\
S &= U \begin{pmatrix} 0 & \pm\sigma_0 \\ \pm\sigma_0 & 0 \end{pmatrix} U^\dagger \\
\beta &= U \begin{pmatrix} 0 & m\sigma_0 \\ m\sigma_0 & 0 \end{pmatrix} U^\dagger
\end{aligned} \tag{3.101}
$$

where m is a real parameter (the mass parameter) and U is the unitary matrix specified in eq (3.85).

Comparing these matrices to the ones that correspond to the fundamental Hermitan Dirac equation in section 3.1.2, we see that the matrices that comprise the Hamiltonian eq (3.2) are identical, so the Hamiltonians themselves are identical. Thus Model 4 emerges as equivalent in every regard to the fundamental representation of the Dirac equation, though not by design– simply by enforcing the Dirac algebra and principles of PT quantum mechanics.

3.4 Model 8

In section 3.1.3 we saw that higher dimensional representations of the Hermitian Dirac equation can be decoupled into independent fermions, by means of singular value decomposition. Given that Model 4, the PT analogue of the fundamental Dirac fermion, is completely equivalent to the Hermitian case, one might naturally expect that higher dimensional representations of the PT Dirac equation are equivalent to

higher dimensional representations of the Hermitian Dirac equation and will similarly decouple. We show in this section that is not the case. It is an extremely interesting property of T_{odd} non-Hermitian quantum mechanics that a new type of particle is described by the higher dimensional representation; the only other time[4] a particle was discovered by means of solving the Dirac equation was over 80 years ago when the great man himself proposed the existence of anti-matter.

3.4.1 Construction

By similar process to Model 4 , we successively apply the conditions (i-vi) to obtain the form of α_i, β, S and Z matrices in the 8-d case. The steps taken are very similar to those used in constructing Model 4, so we will not present them in detail here, but rather refer the reader to the previous section from whence the derivations of the 8-d case can be extrapolated. We choose α_i to be the direct sum of a pair of left-handed and a pair of right-handed non-Hermitian representations:

$$\alpha_i \to \begin{bmatrix} V\sigma_i V^{-1} & 0 & 0 & 0 \\ 0 & V\sigma_i V^{-1} & 0 & 0 \\ 0 & 0 & -W\sigma_i W^{-1} & 0 \\ 0 & 0 & 0 & -W\sigma_i W^{-1} \end{bmatrix}. \tag{3.102}$$

But application of conditions (ii)-(v) leaves us with a model in which α has the form

$$\alpha_i \to \begin{bmatrix} \sigma_i & 0 & 0 & 0 \\ 0 & \sigma_i & 0 & 0 \\ 0 & 0 & -\sigma_i & 0 \\ 0 & 0 & 0 & -\sigma_i \end{bmatrix} \tag{3.103}$$

[4]Shortly after the Dirac equation was set forth, Majorana discovered that it allowed another type of particle, now called a Majorana fermion. But there has not been a definitive observation of a particle of this type.

which is identical to the α_i for the Dirac quartet (eq (3.39)). However the mass matrix β for Model 8 has the general form

$$\beta = \begin{pmatrix} 0 & M \\ M^* & 0 \end{pmatrix} \quad \text{where} \quad M = \begin{bmatrix} (m_0+m_3)\sigma_0 & (m_1-im_2)\sigma_0 \\ (m_1+im_2)\sigma_0 & (m_0-m_3)\sigma_0 \end{bmatrix} \tag{3.104}$$

which is very different from the Dirac quartet model (cf eq (3.40)). This mass matrix is non-Hermitian:

$$\beta = \begin{pmatrix} 0 & 0 & (m_0+m_3)\sigma_0 & (m_1-im_2)\sigma_0 \\ 0 & 0 & (m_1+im_2)\sigma_0 & (m_0-m_3)\sigma_0 \\ (m_0+m_3)\sigma_0 & (m_1+im_2)\sigma_0 & 0 & 0 \\ (m_1-im_2)\sigma_0 & (m_0-m_3)\sigma_0 & 0 & 0 \end{pmatrix}. \tag{3.105}$$

Only in the case $m_2 = 0$ does this β reduce to a special case of eq(3.40). Needless to say, non-Hermitian mass matrices are not allowed within Hermitian quantum mechanics. It is this (arguably) slight difference between the Dirac quartet mass matrix and the Model 8 mass matrix:

$$\beta_{Dirac} = \begin{pmatrix} 0 & M \\ M^\dagger & 0 \end{pmatrix} \qquad \text{vs.} \qquad \beta_{Model8} = \begin{pmatrix} 0 & M \\ M^* & 0 \end{pmatrix} \tag{3.106}$$

that is solely responsible for the new physics contained in Model 8.

In order to explore this new physics we must determine the S and Z matrices that specify the action of the P and T operators, and determine a valid CPT inner

product . Through conditions (i-vi) we also determine the S and Z operators; we find

$$S = \begin{pmatrix} 0 & 0 & \sigma_0 & 0 \\ 0 & 0 & 0 & \sigma_0 \\ \sigma_0 & 0 & 0 & 0 \\ 0 & \sigma_0 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad Z = \begin{pmatrix} i\sigma_y & 0 & 0 & 0 \\ 0 & i\sigma_y & 0 & 0 \\ 0 & 0 & i\sigma_y & 0 \\ 0 & 0 & 0 & i\sigma_y \end{pmatrix}. \tag{3.107}$$

Parenthetically we note these expressions reveal another remarkable qualitative feature of Model 8, namely that the Hamiltonian fails to commute with both P and T although by design it does commute with the combined symmetry PT. Thus Model 8 breaks parity and time reversal even at the non-interacting level[5].

The first step in exploring the new physics contained in Model 8 is to solve the Hamiltonian eq(3.2) using the representations (3.103) and (3.105); we refer to the Hamiltonian so represented as H_8. We attempt plane wave solutions of the form

$$\psi = u \exp(i\mathbf{p} \cdot \mathbf{r}) \tag{3.108}$$

where u is a fixed spinor. Substitution of the ansatz eq (3.108) into the eigenvalue equation $H_8\psi = E\psi$ reduces it to an 8×8 matrix eigenvalue equation,

$$\begin{pmatrix} \boldsymbol{\sigma} \cdot \mathbf{p} & 0 & (m_0 + m_3)\sigma_0 & (m_1 - im_2)\sigma_0 \\ 0 & \boldsymbol{\sigma} \cdot \mathbf{p} & (m_1 + im_2)\sigma_0 & (m_0 - m_3)\sigma_0 \\ (m_0 + m_3)\sigma_0 & (m_1 + im_2)\sigma_0 & -\boldsymbol{\sigma} \cdot \mathbf{p} & 0 \\ (m_1 - im_2)\sigma_0 & (m_0 - m_3)\sigma_0 & 0 & -\boldsymbol{\sigma} \cdot \mathbf{p} \end{pmatrix} u = Eu. \tag{3.109}$$

[5]One might wish to define a different parity operator setting $S = \beta/m_{\text{eff}}$ for the restricted Model 8. This parity operator does exchange left and right handed components and it also commutes with the Hamiltonian. However this parity operator is not an observable because it is not CPT self-adjoint. This criterion for observability is discussed later in the text.

To further simplify the problem we attempt solutions of the form

$$u = \begin{pmatrix} a\,\xi_+(\hat{\mathbf{n}}) \\ b\,\xi_+(\hat{\mathbf{n}}) \\ c\,\xi_+(\hat{\mathbf{n}}) \\ d\,\xi_+(\hat{\mathbf{n}}) \end{pmatrix} \quad \text{or} \quad u = \begin{pmatrix} a\,\xi_-(\hat{\mathbf{n}}) \\ b\,\xi_-(\hat{\mathbf{n}}) \\ c\,\xi_-(\hat{\mathbf{n}}) \\ d\,\xi_-(\hat{\mathbf{n}}) \end{pmatrix}. \tag{3.110}$$

Here $\hat{\mathbf{n}}$ is the unit vector along $\mathbf{p}$. Thus $\mathbf{p} = p\hat{\mathbf{n}}$ where p is the magnitude of $\mathbf{p}$ and

$$\begin{aligned} \hat{\mathbf{n}}_x &= \sin\theta\cos\varphi \\ \hat{\mathbf{n}}_y &= \sin\theta\sin\varphi \\ \hat{\mathbf{n}}_z &= \cos\theta. \end{aligned} \tag{3.111}$$

We denote the eigenspinors of $\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}$ as $\xi_\pm(\hat{\mathbf{n}})$, given by

$$\xi_+(\hat{\mathbf{n}}) = \begin{pmatrix} \cos\theta/2 \\ \exp(i\varphi)\sin\theta/2 \end{pmatrix} \qquad \xi_-(\hat{\mathbf{n}}) = \begin{pmatrix} -\exp(-i\varphi)\sin\theta/2 \\ \cos\theta/2 \end{pmatrix}. \tag{3.112}$$

For brevity, we will write the ansätze, eq (3.110) as

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \otimes \xi_+(\hat{\mathbf{n}}) \quad \text{or} \quad \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \otimes \xi_-(\hat{\mathbf{n}}). \tag{3.113}$$

Substituting eq (3.110) into eq (3.109) reduces it to 4×4 matrix eigenvalue equations;

namely,

$$\begin{pmatrix} p & 0 & m_0+m_3 & m_1-im_2 \\ 0 & p & m_1+im_2 & m_0-m_3 \\ m_0+m_3 & m_1+im_2 & -p & 0 \\ m_1-im_2 & m_0-m_3 & 0 & -p \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = E \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}, \quad (3.114)$$

for the positive helicity case and

$$\begin{pmatrix} -p & 0 & m_0+m_3 & m_1-im_2 \\ 0 & -p & m_1+im_2 & m_0-m_3 \\ m_0+m_3 & m_1+im_2 & p & 0 \\ m_1-im_2 & m_0-m_3 & 0 & p \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = E \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}, \quad (3.115)$$

for the negative helicity case.

For simplicity we now set $m_1 = m_3 = 0$; (from now on we will refer to this restricted case as 'Model 8' even though the full Model 8 is more general). It is convenient to define the 2×2 matrix

$$M = \begin{pmatrix} m_0 & -im_2 \\ im_2 & m_0 \end{pmatrix} \quad (3.116)$$

and to define the two-component column vectors

$$\zeta_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad \zeta_- = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \quad (3.117)$$

The reader may recognize $\zeta_\pm$ as eigenvectors of σ_2 with eigenvalue ± 1 respectively.

As a result one can see

$$M\zeta_+ = (m_0 + m_2)\zeta_+, \quad M\zeta_- = (m_0 - m_2)\zeta_-$$
$$M^*\zeta_+ = (m_0 - m_2)\zeta_+, \quad M^*\zeta_- = (m_0 + m_2)\zeta_- \tag{3.118}$$

We now attempt solutions to eqs (3.114) and (3.115) that are of the form

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} \alpha\,\zeta_+ \\ \beta\,\zeta_+ \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} \alpha\,\zeta_- \\ \beta\,\zeta_- \end{pmatrix}. \tag{3.119}$$

For brevity we will write eq (3.119) in the more compact notation

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_+ \quad \text{or} \quad \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_-. \tag{3.120}$$

These ansätze conveniently reduce the 4×4 matrix equations to 2×2 equations. For example if we substitute the ζ_+ ansatz into the positive helicity eq (3.114), we obtain the 2×2 eigenvalue problem

$$\begin{pmatrix} p & m_0 + m_2 \\ m_0 - m_2 & -p \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{3.121}$$

Proceeding in this way our set of four ansätze, $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_\pm \otimes \xi_\pm(\hat{\mathbf{n}})$, reduce solution of the 8×8 matrix eigenvalue problem, eq (3.109), to a set of four 2×2 matrix

eigenvalue problems summarized below:

$$\begin{array}{lll} \textbf{Ansatz} & & \textbf{Eigenproblem} \\ u = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_+ \otimes \xi_+(\hat{\mathbf{n}}) & \Rightarrow & \begin{pmatrix} p & m_0+m_2 \\ m_0-m_2 & -p \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\ u = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_- \otimes \xi_+(\hat{\mathbf{n}}) & \Rightarrow & \begin{pmatrix} p & m_0-m_2 \\ m_0+m_2 & -p \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\ u = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_+ \otimes \xi_-(\hat{\mathbf{n}}) & \Rightarrow & \begin{pmatrix} -p & m_0+m_2 \\ m_0-m_2 & p \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\ u = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_- \otimes \xi_-(\hat{\mathbf{n}}) & \Rightarrow & \begin{pmatrix} -p & m_0-m_2 \\ m_0+m_2 & p \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \end{array} \tag{3.122}$$

Note that the 2×2 matrices are all manifestly non-Hermitian.

Let us now solve the eigenvalue eq (3.121). Let us denote the matrix in that equation as H_1. Evidently H_1 is traceless and has determinant $-p^2 - m_0^2 + m_2^2$. Right away, this tells us that the eigenvalues are $\pm\varepsilon$ where

$$\varepsilon = \sqrt{p^2 + m_0^2 - m_2^2}. \tag{3.123}$$

To determine the eigenvectors it is convenient to first introduce the parameters

$0 \leq \alpha \leq \pi$ and $-\infty < \gamma < \infty$ in terms of which

$$\begin{aligned} p &= \varepsilon \cos\alpha \\ m_0 &= \varepsilon \sin\alpha \cosh\gamma \\ m_2 &= \varepsilon \sin\alpha \sinh\gamma \end{aligned} \tag{3.124}$$

In terms of this parametrization we may write

$$H_1 = \begin{pmatrix} p & m_0 + m_2 \\ m_0 - m_2 & -p \end{pmatrix} \to \varepsilon \begin{pmatrix} \cos\alpha & \sin\alpha \exp\gamma \\ \sin\alpha \exp(-\gamma) & -\cos\alpha \end{pmatrix} \equiv \varepsilon h_1. \tag{3.125}$$

Evidently the matrix h_1 has eigenvalues ± 1; hence its eigenvectors may be constructed by the projection operator method. Let us denote the eigenvectors $\eta_{1\pm}$.

We define the projectors

$$\begin{aligned} P_{1+} = \frac{1}{2}(1 + h_1) &= \begin{pmatrix} 1 + \cos\alpha & \sin\alpha \exp\gamma \\ \sin\alpha \exp(-\gamma) & 1 - \cos\alpha \end{pmatrix} \\ P_{1-} = \frac{1}{2}(1 - h_1) &= \begin{pmatrix} 1 - \cos\alpha & -\sin\alpha \exp\gamma \\ -\sin\alpha \exp(-\gamma) & 1 + \cos\alpha \end{pmatrix}. \end{aligned} \tag{3.126}$$

In exactly the same way as we constructed the eigenspinors of $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$, we now find the eigenvectors of h_1 to be

$$\eta_{1+} = \begin{pmatrix} \cos(\alpha/2) \exp(\gamma/2) \\ \sin(\alpha/2) \exp(-\gamma/2) \end{pmatrix} \quad \eta_{1-} = \begin{pmatrix} -\sin(\alpha/2) \exp(\gamma/2) \\ \cos\alpha/2 \exp(-\gamma/2) \end{pmatrix}. \tag{3.127}$$

Proceeding in this way we can construct the eigenvectors of all four eqs (3.122).

The eigenstates of the restricted Model 8 are of the form $\psi = u\exp(i\mathbf{p}\cdot\mathbf{r})$. For a given wave-vector $\mathbf{p}$ there are four positive energy solutions and four negative energy. The energy is given by $\pm\varepsilon$ where $\varepsilon = (p^2+m_0^2-m_2^2)^{1/2}$. We denote the positive energy eigenspinors $u_1(\mathbf{p}), u_2(\mathbf{p}), u_3(\mathbf{p})$ and $u_4(\mathbf{p})$; we denote negative energy eigen-spinors $v_1(\mathbf{p})$, $v_2(\mathbf{p})$, $v_3(\mathbf{p})$ and $v_4(\mathbf{p})$. The eigenspinors $u_1(\mathbf{p})$ and $v_1(\mathbf{p})$ have the form

$$\begin{aligned} u_1 &= \eta_{1,\text{pos}}\otimes\zeta_+\otimes\xi_+(\hat{\mathbf{n}}) & v_1 &= \eta_{1,\text{neg}}\otimes\zeta_+\otimes\xi_+(\hat{\mathbf{n}}) \\ \eta_{1,\text{pos}} &= \begin{bmatrix} \cos(\alpha/2)\exp(\gamma/2) \\ \sin(\alpha/2)\exp(-\gamma/2) \end{bmatrix} & \eta_{1,\text{neg}} &= \begin{bmatrix} -\sin(\alpha/2)\exp(\gamma/2) \\ \cos(\alpha/2)\exp(-\gamma/2) \end{bmatrix}. \end{aligned} \tag{3.128}$$

The eigenspinors $u_2(\mathbf{p})$ and $v_2(\mathbf{p})$ have the form

$$\begin{aligned} u_2 &= \eta_{2,\text{pos}}\otimes\zeta_-\otimes\xi_+(\hat{\mathbf{n}}) & v_2 &= \eta_{2,\text{neg}}\otimes\zeta_-\otimes\xi_+(\hat{\mathbf{n}}) \\ \eta_{2,\text{pos}} &= \begin{bmatrix} \cos(\alpha/2)\exp(\gamma/2) \\ \sin(\alpha/2)\exp(-\gamma/2) \end{bmatrix} & \eta_{2,\text{neg}} &= \begin{bmatrix} -\sin(\alpha/2)\exp(\gamma/2) \\ \cos(\alpha/2)\exp(-\gamma/2) \end{bmatrix}. \end{aligned} \tag{3.129}$$

The eigenspinors $u_3(\mathbf{p})$ and $v_3(\mathbf{p})$ have the form

$$\begin{aligned} u_3 &= \eta_{3,\text{pos}}\otimes\zeta_+\otimes\xi_-(\hat{\mathbf{n}}) & v_3 &= \eta_{3,\text{neg}}\otimes\zeta_+\otimes\xi_-(\hat{\mathbf{n}}) \\ \eta_{3,\text{pos}} &= \begin{bmatrix} \sin(\alpha/2)\exp(\gamma/2) \\ \cos(\alpha/2)\exp(-\gamma/2) \end{bmatrix} & \eta_{3,\text{neg}} &= \begin{bmatrix} \cos(\alpha/2)\exp(\gamma/2) \\ -\sin(\alpha/2)\exp(-\gamma/2) \end{bmatrix}. \end{aligned} \tag{3.130}$$

The eigenspinors $u_4(\mathbf{p})$ and $v_4(\mathbf{p})$ have the form

$$u_4 = \eta_{4,\text{pos}} \otimes \zeta_- \otimes \xi_-(\hat{\mathbf{n}}) \qquad v_4 = \eta_{4,\text{neg}} \otimes \zeta_- \otimes \xi_-(\hat{\mathbf{n}})$$
$$\eta_{4,\text{pos}} = \begin{bmatrix} \sin(\alpha/2)\exp(-\gamma/2) \\ \cos(\alpha/2)\exp(\gamma/2) \end{bmatrix} \qquad \eta_{4,\text{neg}} = \begin{bmatrix} \cos(\alpha/2)\exp(-\gamma/2) \\ -\sin(\alpha/2)\exp(\gamma/2) \end{bmatrix}. \tag{3.131}$$

The parameters α and γ are defined in eq (3.124), the spinors $\zeta_\pm$ are defined in eq (3.117) and $\xi_\pm(\hat{\mathbf{n}})$ are defined by eq (3.112).

In short, we find that for a given momentum $\mathbf{p}$ there are four positive energy eigenvectors and four with negative energy. The energy momentum dispersion is $E = \pm\sqrt{p^2 + m_{\text{eff}}^2}$ corresponding to a relativistic particle of mass $m_{\text{eff}} = \sqrt{m_0^2 - m_2^2}$. For the energy eigenvalues to be real we need to impose $m_0^2 \geq m_2^2$. This leads right away to a remarkable phenomenon that is impossible in conventional Hermitian quantum mechanics: $m_{\text{eff}} = 0$ when $m_0^2 = m_2^2$, so Model 8 describes a particle with an effective mass of zero but a non-zero mass matrix.

At the very least, therefore, Model 8 corresponds to a new type of particle, not allowed within the Standard Model. In Hermitian quantum mechanics and quantum field theory, the only irreducible solutions to the Hamiltonian eq(3.2) are 4-component Dirac fermions, and the only massless solutions are 2-component Weyl spinors. Model 8 is an irreducible 8-component wavefunction , that is massless in a particular regime. Before going into the proof that Model 8 does not reduce to independent fermions like the quartet, we comment on the possiblity that the particle described by Model 8 is not a new particle *per se*, but rather a new description of a particle that has been known for over 50 years , the neutrino.

Prior to 1999 , the neutrino was thought to be a massless left-handed Weyl fermion. Then flavor oscillations were observed and theorists were forced to come up with a

different description. The simplest way to describe flavor oscillations was to assume that the neutrino did in fact have a small but non-zero mass, thereby facilitating the need for a right-handed Weyl fermion to which the left-handed one was coupled. If we assume the mass matrix that couples the two species and the Hamiltonian[6] cannot be simultaneously diagonalized , this would give rise to oscillations (and hence a mass difference) between the two different neutrino species.

The Dirac quartet is a toy model of precisely this type of neutrino. Assume the Hamiltonian [eqs (3.2), (3.39) and (3.40)] is written in an interaction basis so that the first and third pair of components of the eight component wave function correspond to the first generation neutrino and the second and fourth pair of components to the second generation neutrino. In principle these two neutrino states are mixed by the mass matrix leading to neutrino oscillations. However if both neutrino masses vanish ($\mu_1 = \mu_2 = 0$), that corresponds to a vanishing mass matrix and hence no neutrino oscillations. Indeed it is for this reason that the observation of neutrino oscillations is generally considered evidence that neutrinos have mass [19], but it is important to point out that experiments do not measure the bare mass of the neutrino. Nowadays it is ubiquitously acknowledged and stated that neutrinos have mass, when the more accurate statement is that neutrinos flavor oscillate.

Contrast this with two generations of neutrinos coupled by the mass matrix of Model 8. In this case the mass matrix does not need to vanish even if the neutrino masses vanish ($m_{\text{eff}} = 0$ does not require m_0 and m_2 to vanish). Within the conventional Dirac or Majorana framework the observation of neutrino oscillations inexorably implies that neutrinos have mass. PT quantum mechanics clearly allows for a quite distinct phenomenology for neutrino oscillations than conventional models. In order to determine whether PT quantum mechanics can describe a massless neutrino that still allows flavor oscillations, much work needs to be done, specifically,

[6]The relevant Hamiltonian here would be the fully interacting electroweak portion of the Standard Model Hamiltonian, not just H_8.

one must develop Model 8 as a fully interacting quantum field theory. This is the topic of work in progress. However, even if the particle described by Model 8 is ruled out by existing neutrino data, it is an extremely interesting new type of particle and certainly worthy of further investigation.

3.4.2 CPT Inner Product

To complete the formulation of Model 8 as a PT quantum theory we must now specify the dynamically determined CPT inner product. In this case, as in general, the PT inner product, eq (**??**), is not satisfactory because it is not positive definite: although $(\psi, \psi)_{PT}$ is real, its sign may be positive or negative and $(\psi, \psi)_{PT} = 0$ is possible for non-trivial states ψ. In general the CPT inner product is constructed by taking degenerate multiplets of the Hamiltonian and organizing them into states that are mutually orthogonal under the PT inner product. For these special basis states the CPT inner product is defined as the absolute value of their PT inner product. The CPT inner product for other states can be determined by expanding in this special basis and using the assumed bilinearity of the CPT inner product.

Before we immerse ourselves in a detailed analysis it is helpful to give the final result which is simpler than its derivation. We find that

$$\begin{aligned}
(u_i(\mathbf{p}) \exp i\mathbf{p}\cdot\mathbf{r}, u_j(\mathbf{k}) \exp i\mathbf{k}\cdot\mathbf{r})_{\mathrm{CPT}} &= \mathrm{vol}\delta_{\mathbf{p},\mathbf{k}}\delta_{ij} \\
(v_i(\mathbf{p}) \exp i\mathbf{p}\cdot\mathbf{r}, v_j(\mathbf{k}) \exp i\mathbf{k}\cdot\mathbf{r})_{\mathrm{CPT}} &= \mathrm{vol}\delta_{\mathbf{p},\mathbf{k}}\delta_{ij} \\
(u_i(\mathbf{p}) \exp i\mathbf{p}\cdot\mathbf{r}, v_j(\mathbf{k}) \exp i\mathbf{k}\cdot\mathbf{r})_{\mathrm{CPT}} &= 0. \qquad (3.132)
\end{aligned}$$

Here $u_i(\mathbf{p}) \exp i\mathbf{p}\cdot\mathbf{r}$ and $v_i(\mathbf{p}) \exp i\mathbf{p}\cdot\mathbf{r}$ (with $i = 1, \ldots, 4$) denote the plane wave positive and negative energy states of Model 8 with momentum $\mathbf{p}$ that were derived in the previous section. Needless to say, if we simply computed the naive inner product of these states ($\int d\mathbf{r}\ \psi^\dagger(\mathbf{r})\phi(\mathbf{r})$) we would not find that these Model 8 eigenstates are

orthonormal.

Now we turn to the derivation of these results. We start by evaluating the PT inner product of the Model 8 eigenstates. The PT inner product of a pair of plane waves may be written as

$$(u \exp i\mathbf{k}\cdot\mathbf{r}, v \exp i\mathbf{p}\cdot\mathbf{r})_{\rm PT} = \text{vol}\ \delta_{\mathbf{k}+\mathbf{p},0}\ u^\dagger S v. \tag{3.133}$$

The momentum constraint reveals that the PT inner product of each plane wave eigenstate with itself is zero! We will find that positive energy states of momentum $\mathbf{k}$ have non-zero PT inner product with states that have momentum $-\mathbf{k}$ by virtue of the delta function in eq (3.133).

Following the notation of the preceding section we write the spinors u and v as $v = \eta\otimes\zeta\otimes\xi$ and $u = \eta'\otimes\zeta'\otimes\xi'$. One can check that $u^\dagger v = (\eta'^\dagger\eta)(\zeta'^\dagger\zeta)(\xi'^\dagger\xi)$. Similarly from the form of S given in eq (2-106) one can check that $Sv = (\sigma_x\eta)\otimes\zeta\otimes\xi$. Thus $u^\dagger S v = (\eta'^\dagger\sigma_x\eta)(\zeta'^\dagger\zeta)(\xi'^\dagger\xi)$.

Armed with these results and the explicit forms of the eigenspinors computed in the last section we may now calculate the spinor elements $u^\dagger S v$ where u is one of the four spinors $u_i(-\mathbf{p})$ and v is one of the four spinors $u_i(\mathbf{p})$. We will find that $u_1(-\mathbf{p})$ only overlaps with $u_3(\mathbf{p})$; $u_2(-\mathbf{p})$ only with $u_4(\mathbf{p})$; $u_3(-\mathbf{p})$ only with $u_1(\mathbf{p})$; and $u_4(-\mathbf{p})$ only with $u_2(\mathbf{p})$. In particular for the first pair we find $u_1^\dagger(-\mathbf{p})Su_3(\mathbf{p}) = -\exp(-i\varphi)$ where φ is the azimuthal angle corresponding to the momentum $\mathbf{p}$.

This gives us license to define

$$\begin{aligned}
\psi^+_{13\mathbf{p}} &= \frac{1}{\sqrt{2}}\left[e^{-i\varphi/2}u_1(-\mathbf{p})\exp(-i\mathbf{p}\cdot\mathbf{r}) - e^{i\varphi/2}u_3(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{r})\right].\\
\psi^-_{13\mathbf{p}} &= \frac{1}{\sqrt{2}}\left[e^{-i\varphi/2}u_1(-\mathbf{p})\exp(-i\mathbf{p}\cdot\mathbf{r}) + e^{i\varphi/2}u_3(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{r})\right].
\end{aligned} \tag{3.134}$$

By design these states are mutually orthogonal under the PT inner product. Unlike the plane wave eigenstates they are not self-orthogonal under the PT inner product. Their inner products may be summarized as

$$(\psi^{s}_{13\mathbf{p}}(\mathbf{r}), \psi^{s'}_{13\mathbf{p}}(\mathbf{r}))_{\mathrm{PT}} = s\delta_{ss'} \text{ vol.} \tag{3.135}$$

Here $s = +$ or $-$ and s' likewise. Having found a set of states that are PT orthogonal we now define their CPT inner product as the norm of the PT inner product:

$$(\psi^{s}_{13\mathbf{p}}(\mathbf{r}), \psi^{s'}_{13\mathbf{p}}(\mathbf{r}))_{\mathrm{CPT}} = \delta_{ss'} \text{ vol.} \tag{3.136}$$

Next we may invert eq (3.134) to obtain

$$u_1(-\mathbf{p})\exp(-i\mathbf{p}\cdot\mathbf{r}) \;=\; \frac{1}{\sqrt{2}}e^{i\varphi/2}\psi^{+}_{13\mathbf{p}}(\mathbf{r}) + \frac{1}{\sqrt{2}}e^{i\varphi/2}\psi^{-}_{13\mathbf{r}}(\mathbf{r}) \tag{3.137}$$

Making use of eq (3.136), (3.137) and the bilinearity of the CPT inner product we obtain

$$\begin{aligned}
\{u_1(-\mathbf{p})\exp(-i\mathbf{p}\cdot\mathbf{r}), u_1(-\mathbf{p})\exp(-i\mathbf{p}\cdot\mathbf{r})\}_{\mathrm{CPT}} &= \text{vol} \\
\{u_3(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{r}), u_3(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{r})\}_{\mathrm{CPT}} &= \text{vol} \\
\{u_1(-\mathbf{p})\exp(-i\mathbf{p}\cdot\mathbf{r}), u_3(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{r})\}_{\mathrm{CPT}} &= 0 \\
\{u_3(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{r}), u_1(-\mathbf{p})\exp(-i\mathbf{p}\cdot\mathbf{r})]_{\mathrm{CPT}} &= 0.
\end{aligned} \tag{3.138}$$

The other pairs of states linked by the PT inner product [namely, $u_2(-\mathbf{p})\exp(-i\mathbf{p}\cdot\mathbf{r})$ and $u_4(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{r})$; $u_3(-\mathbf{p})\exp(-i\mathbf{p}\cdot\mathbf{r})$ and $u_1(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{r})$; and $u_4(-\mathbf{p})\exp(-i\mathbf{p}\cdot\mathbf{r})$ and $u_2(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{r})$] may be analyzed to give results similar to eq (3.138). Combining these results and extending the analysis

to negative energy states finally leads to the verification of eq (3.132).

Armed with the CPT inner product we can now determine the observables of the theory. In PT quantum mechanics the operators corresponding to observables must be CPT self-adjoint [4, 17]. The CPT adjoint of an operator A is defined as the operator $A^\star$ for which $(A^\star\phi, \psi)_{CPT} = (\phi, A\psi)_{CPT}$; an operator is self-adjoint if $A = A^\star$.

Now we would like to dispatch any concern that perhaps the restricted Model 8 is merely an elaborate way to rewrite a trivial Hermitian model, namely a pair of 4×4 Dirac Hamiltonians, each of mass m_{eff}, assembled into an 8×8 block. Certainly such a Dirac pair model also has an energy momentum dispersion $E = \pm\sqrt{p^2 + m_{\text{eff}}^2}$ with four positive and four negative energy eigenfunctions for a given momentum that we denote $u_i^{\text{Dirac}}(\mathbf{p}) \exp(i\mathbf{p} \cdot \mathbf{r})$. From the correspondence between the eigenfunctions of the Dirac pair model and the eigenfunctions of Model 8, $u_i(\mathbf{p}) \exp(i\mathbf{p} \cdot \mathbf{r})$, we can indeed construct a transformation that maps the Model 8 Hamiltonian to the Dirac pair Hamiltonian. That transformation maps Model 8 wave functions $\psi_8(\mathbf{r})$ to Dirac pair wave functions $\psi_{\text{Dirac}}(\mathbf{r})$ via the convolution

$$\psi_{\text{Dirac}}(\mathbf{r}) = \int d\mathbf{r}'\, L(\mathbf{r} - \mathbf{r}')\psi_8(\mathbf{r}'). \tag{3.139}$$

The kernel L has a range set by the non-hermiticity parameter m_2; an explicit formula will be given elsewhere [18]. That the transformation eq (3.139) is non-local shows clearly that Model 8 and the Dirac pair model have different physics: if we coupled them to the same gauge or scalar field we would get different outcomes. This argument, together with the broken parity and time-reversal symmetry and the prospect of massless oscillations all clearly reveal the trans-Dirac character of Model 8.

We now construct Lorentz covariant bilinears to facilitate the study of interactions. To this end we write the 8-component wave function as a column of four

two-component spinors

$$\psi = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \eta_1 \\ \eta_2 \end{pmatrix}. \tag{3.140}$$

From the form of $\boldsymbol{\alpha}$ for Model 8 we see that ξ_1 and ξ_2 transform like left handed spinors under boosts and rotations and η_1 and η_2 like right-handed. Furthermore parity exchanges ξ_1 with η_1 and ξ_2 with η_2. Thus, just as in Dirac theory, $\xi_i^\dagger \eta_j$ and $\eta_i^\dagger \xi_j$ are all scalars under boosts and rotations. Furthermore the symmetric combination $\xi_1^\dagger \eta_1 + \eta_1^\dagger \xi_1$ is a true scalar being invariant under parity whereas the antisymmetric combination $-i(\xi_1^\dagger \eta_1 + \eta_1^\dagger \xi_1)$ is a pseudo-scalar as it changes sign under parity. Similarly the currents $(\xi_i^\dagger \xi_j, \xi_i^\dagger \boldsymbol{\sigma} \xi_j)$ and $(\eta_i^\dagger \eta_j, -\eta_i^\dagger \boldsymbol{\sigma} \eta_j)$ are four-vectors under boosts and rotations. By making appropriate symmetric and anti-symmetric combinations we can construct currents that are true vectors or axial vectors under parity. Interactions can now be studied by Yukawa coupling the scalar bilinears to a scalar field or the vector currents to a gauge field [18].

Finally we reformulate the restricted Model 8 as a quantum field theory. To this end it is helpful to introduce a new notation for the eigenfunctions of Model 8. We denote the eigenfunctions as $u_i(\mathbf{p}) \exp i\mathbf{p}\cdot\mathbf{r}$ where $i = 1, \ldots, 4$ corresponds to positive energy and $i = 5, \ldots, 8$ to negative energy. Since the u's are not eigenspinors of a Hermitian matrix they are not orthonormal. It is convenient to introduce a set of dual spinors $\tilde{u}_i(\mathbf{p})$ that satisfy $\tilde{u}_i^\dagger(\mathbf{p}) u_j(\mathbf{p}) = \delta_{ij}$. These orthonormality conditions are sufficient to specify the $\tilde{u}$'s but we also note that they are eigenfunctions of $H_D^\dagger$ or equivalently of the restricted Model 8 with $m_2 \to -m_2$.

We now introduce particle and anti-particle creation and annihilation operators $c_i(\mathbf{p}), c_i^\star(\mathbf{p}), d_i(\mathbf{p})$ and $d_i^\star(\mathbf{p})$ where $i = 1 \ldots 4$. These operators obey the fermionic

anti-commutation relations

$$[c_i(\mathbf{p}), c_j^\star(\mathbf{k})]_+ = [d_i(\mathbf{p}), d_j^\star(\mathbf{k})]_+ = \delta(\mathbf{k}-\mathbf{p})\delta_{ij}, \tag{3.141}$$

and all other pairs of operators anti-commute. In terms of these creation operators we may write the Model 8 Hamiltonian as

$$H = \int d\mathbf{p} \sum_{i=1}^{4} \sqrt{p^2 + m_{\text{eff}}^2}\, [c_i^\star(\mathbf{p}) c_i(\mathbf{p}) + d_i^\star(\mathbf{p}) d_i(\mathbf{p})]\,. \tag{3.142}$$

Similar expressions can be written for the momentum, parity and other operators. Next we introduce local field operators

$$\begin{aligned} \hat{\psi}(\mathbf{r}) &= \sum_{i=1}^{4} \int d\mathbf{p} \left[c_i(\mathbf{p}) u_i(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{r}} + d_i^\star(\mathbf{p}) v_i(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{r}} \right] \\ \hat{\psi}^\star(\mathbf{r}) &= \sum_{i=1}^{4} \int d\mathbf{p} \left[c_i^\star(\mathbf{p}) \tilde{u}_i^\dagger(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{r}} + d_i(\mathbf{p}) \tilde{v}_i^\dagger(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{r}} \right]. \end{aligned} \tag{3.143}$$

The novel twist here is the appearance of $\tilde{u}$ and $\tilde{v}$ in the definition of $\psi^\star$. These operators obey the canonical anti-commutation relation

$$[\psi_a(\mathbf{r}), \psi_b^\star(\mathbf{r})]_+ = \delta_{ab}\delta(\mathbf{r}-\mathbf{r}'). \tag{3.144}$$

Thus far we have simply re-written the non-interacting Model 8 in the language of canonical field theory. But we may now consistently treat interactions by coupling appropriate bilinears of the field operators to external scalar and gauge fields. A subtlety that would arise in the perturbative analysis of interactions is that the dynamically determined inner product would also need to be recomputed perturbatively [18].

Chapter 4

Potential Applications within Condensed Matter

A major goal in condensed matter physics is to represent the low-energy physics of strongly interacting quantum many-body systems in terms of weakly interacting quasiparticles that are either bosonic or fermionic [20]. In a seminal paper Dyson [7] showed that a Heisenberg ferromagnet could be represented as a theory of weakly interacting bosons called magnons or spin waves; this representation allowed thermodynamic calculations of unprecedented accuracy.

Dyson's formulation had the unorthodox feature that the bosons were governed by a Hamiltonian that was superficially non-Hermitian. More precisely there were two inner products at work in Dyson's representation of a ferromagnet. First, there was what we will call the "kinematic inner product" with regard to which the boson creation and annihilation operators were adjoints of each other. In other words, this was the inner product with regard to which the quasiparticles were bosons. Second there was the "dynamical inner product" with regard to which the Hamiltonian was self-adjoint. Conversely, however, the quasiparticles were not bosonic with respect to the dynamical inner product and the Hamiltonian was not self-adjoint with respect

to the kinematic inner product.

By contrast the conventional approach is far more restrictive in that there is only a single inner product with regard to which the quasiparticles are defined and with regard to which the Hamiltonian and all other physical operators must be self-adjoint. The purpose of this chapter is to explore whether Dyson's more flexible concept of non-Hermitian quasiparticles can be more broadly applied, particularly to problems that have so far resisted conventional Hermitian analysis.

The $t-J$ model is believed to capture the essential physics of the cuprate superconductors, which represent one of the grand unsolved puzzles of theoretical physics [8]. In this chapter we apply non-Hermitian quantum mechanics to this model and obtain a representation of its low energy physics in terms of a Dyson boson and a Dyson fermion. By design these quasiparticles are defined with respect to a kinematic inner product; the Hamiltonian that governs them is not self-adjoint with respect to the kinematic inner product but with respect to the dynamical inner product.

First we review Dyson's work on ferromagnets, highlighting the role of the two inner products. We then adapt the analysis to antiferromagnets, a useful prelude to the study of the $t-J$ model. In the following section we describe a spin S generalization of the $t-J$ Hamiltonian (the physical case relevant to the cuprates is $S=1/2$). A natural and convenient way to write the $t-J$ Hamiltonian is to use a super-algebra that is a super-symmetric generalization of the su(2) angular momentum algebra. After presenting this supersymmetric formulation of the $t-J$ model we finally write the problem in terms of non-Hermitian quantum mechanics.

4.1 Magnets

4.1.1 Single spin

A single spin has $2s+1$ basic states $|s,m>$ where s is the total spin and m is its z-component. s is the same for all states of the multiplet and $m=-s,\ldots,s$. These states are assumed to be orthonormal

$$< s,m|s,m' >= \delta_{mm'}. \tag{4.1}$$

The spin-operators S_z, S_+ and S_- obey the angular momentum algebra

$$[S_+,S_-]=2S_z,\ \ [S_+,S_z]=-S_+,\ \ [S_-,S_z]=S_-, \tag{4.2}$$

where, as usual, the spin-raising operator $S_+ = S_x+iS_y$ and the spin-lowering operator $S_- = S_x-iS_y$. As shown in textbooks, the effect of these operators on the basis states $|s,m>$ is

$$\begin{aligned} S_+|s,m> &= (s-m)^{1/2}(s+m+1)^{1/2}|s,m+1>, \\ S_-|s,m> &= (s-m+1)^{1/2}(s+m)^{1/2}|s,m-1>, \\ S_z|s,m> &= m|s,m>\ . \end{aligned} \tag{4.3}$$

Dyson introduced an alternative set of basis states

$$|u>= F_u|s,-s+u> \tag{4.4}$$

where $u=0,\ldots,2s$. The state $|0>$ corresponds to having the z-component of the spin maximally down; the states $|1>,|2>,|3>,\ldots$ correspond to raising the

z-component by increments of one. These states are orthogonal but not normalized

$$< u|v >= F_u^2 \delta_{u,v}. \tag{4.5}$$

The normalization factors $F_0 = 1$ and

$$F_u = \left(1 \left[1 - \frac{1}{2s}\right] \left[1 - \frac{2}{2s}\right] \ldots \left[1 - \frac{u-1}{2s}\right]\right)^{1/2} \tag{4.6}$$

for $u = 1, 2, \ldots 2s$. F_u is judiciously chosen to map the spin-raising operator S_+ to the bose creation operator $b^\dagger$, as will be seen below.

Making use of eqs (4.3), (4.4) and (4.6) it is not difficult to show

$$\begin{aligned} S_+|u> &= \sqrt{2s}\sqrt{u+1}|u+1>, \\ S_-|u> &= \sqrt{2s}\left[1 - \frac{u-1}{2s}\right]\sqrt{u}|u-1>, \\ S_z|u> &= (-s+u)|u> . \end{aligned} \tag{4.7}$$

Now consider a different Hilbert space with two operators b and $b^\dagger$ that are the adjoints of each other under a certain inner product, the "kinematic inner product". These operators are assumed to satisfy the bose commutation relations

$$[b, b^\dagger] = 1. \tag{4.8}$$

Provided the kinematic inner product is positive definite it follows inexorably by standard textbook arguments that the basic states in this Hilbert space form an infinite ladder $|u)$ with $u = 0, 1, 2, \ldots$ The state $|0)$ has the defining characteristic

$$b|0) = 0; \tag{4.9}$$

we say this is a state with zero bosons. The state

$$|u) = \frac{1}{\sqrt{u!}}(b^\dagger)^u|0) \tag{4.10}$$

is said to contain u bosons. These states are orthonormal under the kinematic inner product

$$(u|v)_{\rm kin} = \delta_{u,v} \tag{4.11}$$

and the effect of the bose creation and annihilation operators on these states is

$$\begin{aligned} b^\dagger|u) &= \sqrt{u+1}|u+1), \\ b|u) &= \sqrt{u}|u-1), \\ b^\dagger b|u) &= u|u). \end{aligned} \tag{4.12}$$

Following Dyson, we now establish a mapping between the space of spins and the bose oscillator space by identifying the spin state $|u>$ with the boson state $|u)$. Thus

$$|u> \rightarrow |u) \tag{4.13}$$

for $u = 0, \ldots, 2s$. States with more than $2s$ bosons have no spin space counterpart.

Dyson's mapping allows us to export the inner product of the spin space to the bose space. We call this induced inner product the dynamical inner product. Explicitly

$$(u|v)_{\rm dyn} = F_u^2 \delta_{u,v} \tag{4.14}$$

for $u = 0, \ldots, 2s$. We take $F_u = 0$ for $u > 2s$. Thus states with more than $2s$ bosons are "weightless".

Dyson's mapping eq (4.13) also allows us to establish the following correspondence

between spin and bose operators

$$\begin{aligned} S_+ &\rightarrow \sqrt{2s}b^\dagger, \\ S_- &\rightarrow \sqrt{2s}\left[1-\frac{b^\dagger b}{2s}\right]b, \\ S_z &\rightarrow -s+b^\dagger b. \end{aligned} \tag{4.15}$$

This correspondence follows from comparison of eq (4.7) and (4.12).

b and $b^\dagger$ are not the adjoints of each other under the dynamical inner product. Since we are denoting the adjoint with respect to the kinematic inner product as †, let us signify the adjoint with respect to the dynamical inner product by *. We can then see for example that

$$(b^\dagger)^\star = \left[1-\frac{b^\dagger b}{2s}\right]b \tag{4.16}$$

and

$$(b^\dagger b)^\star = b^\dagger b. \tag{4.17}$$

4.1.2 Heisenberg Ferromagnet

We now consider a two-dimensional Heisenberg ferromagnet in which the spins occupy the sites of a square lattice. Thus the lattice sites (m,n) have position vector $\mathbf{r}_{mn} = ma\hat{\mathbf{e}}_x + na\hat{\mathbf{e}}_y$ where $\hat{\mathbf{e}}_x$ and $\hat{\mathbf{e}}_y$ are unit vectors along the x and y axes, m and n are integers, and a is the lattice constant. Each site has four nearest neighbors. The site (m,n) has neighbors located at $\mathbf{r}_{mn}+\boldsymbol{\delta}$ where $\boldsymbol{\delta} = a\hat{\mathbf{e}}_x$, $a\hat{\mathbf{e}}_y$, $-a\hat{\mathbf{e}}_x$ and $-a\hat{\mathbf{e}}_y$ respectively for the four neighbors. We denote the spin operator at position $\mathbf{r}$ as $S_+(\mathbf{r}), S_-(\mathbf{r})$ and $S_z(\mathbf{r})$. Operators at a given site are assumed to obey the angular momentum algebra eq (4.2); spin-operators at different sites are assumed to commute. We consider a spin s ferromagnet so the basic states at each site are a spin multiplet

of $2s+1$ states. The Hamiltonian for a Heisenberg ferromagnet is

$$\begin{aligned} H_F &= -\frac{J}{2}\sum_{\mathbf{r}}\sum_{\boldsymbol{\delta}}[S_z(\mathbf{r})S_z(\mathbf{r}+\boldsymbol{\delta}) + \\ &\quad \frac{1}{2}S_+(\mathbf{r})S_-(\mathbf{r}+\boldsymbol{\delta}) + \frac{1}{2}S_+(\mathbf{r}+\boldsymbol{\delta})S_-(\mathbf{r})]. \end{aligned} \tag{4.18}$$

Thus each spin is coupled to its nearest neighbors. We assume the exchange constant $J > 0$.

Now consider a system of bosons $b(\mathbf{r})$ and $b^\dagger(\mathbf{r})$ that live on a square lattice in two dimensions (lattice constant $= a$). The operators $b(\mathbf{r})$ and $b^\dagger(\mathbf{r})$ are assumed to be adjoints of each other under the kinematic inner product. They are assumed to obey the bosonic commutation relation

$$[b(\mathbf{r}), b^\dagger(\mathbf{r}')] = \delta_{\mathbf{r},\mathbf{r}'}. \tag{4.19}$$

Thus $b^\dagger(\mathbf{r})$ creates bosons at site $\mathbf{r}$; $b(\mathbf{r})$ annihilates them. We may now represent the ferromagnetic Heisenberg Hamiltonian eq (4.18) in terms of bosonic quasiparticles by using Dyson's mapping. From the correspondence eq (4.15) between spin and bose operators we obtain the bosonic form of the Heisenberg Hamiltonian

$$\begin{aligned} \mathcal{H}_F =& \frac{Js}{2}\sum_{\mathbf{r},\boldsymbol{\delta}}[2b^\dagger(\mathbf{r})b(\mathbf{r}) - b^\dagger(\mathbf{r})b(\mathbf{r}+\boldsymbol{\delta}) - b^\dagger(\mathbf{r}+\boldsymbol{\delta})b(\mathbf{r})] \\ &+\frac{J}{4}\sum_{\mathbf{r},\boldsymbol{\delta}}[b^\dagger(\mathbf{r})b^\dagger(\mathbf{r}+\boldsymbol{\delta})b^2(\mathbf{r}+\boldsymbol{\delta}) + b^\dagger(\mathbf{r})b^\dagger(\mathbf{r}+\boldsymbol{\delta})b^2(\mathbf{r})] \\ &-\frac{J}{2}\sum_{\mathbf{r},\boldsymbol{\delta}} b^\dagger(\mathbf{r})b(\mathbf{r})b^\dagger(\mathbf{r}+\boldsymbol{\delta})b(\mathbf{r}+\boldsymbol{\delta}). \end{aligned} \tag{4.20}$$

Note that the boson Hamiltonian $\mathcal{H}_F$ is not self-adjoint under the kinematic inner product ($\mathcal{H}_F^\dagger \neq \mathcal{H}_F$) due to the terms in the second line of eq (4.20). However it is

self-adjoint under the dynamical inner product ($\mathcal{H}_F^\star = \mathcal{H}_F$).

4.1.3 Heisenberg Anti-ferromagnet

A Heisenberg anti-ferromagnet is simply a ferromagnet with $J < 0$. An equivalent but more convenient description of the Heisenberg anti-ferromagnet on a square lattice is the following: Imagine two interpenetrating square lattices, the site labelled (m, n) on the first lattice is located at $\mathbf{r}_1(m, n) = ma\hat{\mathbf{e}}_x + na\hat{\mathbf{e}}_y$. Here m and n are integers. The sites of the second square lattice are displaced from those of the first by $(a/2)\hat{\mathbf{e}}_x + (a/2)\hat{\mathbf{e}}_y$. Thus the site labelled (m, n) on the second lattice is located at $\mathbf{r}_2 = (m + 1/2)a\hat{\mathbf{e}}_x + (n + 1/2)a\hat{\mathbf{e}}_y$. Regardless of the lattice on which it sits, each site has four nearest neighbors. The displacements from a given site to its four nearest neighbor sites are $\boldsymbol{\delta}_1 = (a/2)\hat{\mathbf{e}}_x + (a/2)\hat{\mathbf{e}}_y$, $\boldsymbol{\delta}_2 = (a/2)\hat{\mathbf{e}}_x - (a/2)\hat{\mathbf{e}}_y$, $\boldsymbol{\delta}_3 = -(a/2)\hat{\mathbf{e}}_x + (a/2)\hat{\mathbf{e}}_y$, and $\boldsymbol{\delta}_4 = -(a/2)\hat{\mathbf{e}}_x - (a/2)\hat{\mathbf{e}}_y$. We imagine there is a spin at each site and that the spin at each site is antiferromagnetically coupled to its nearest neighbors. Thus the Hamiltonian for a Heisenberg anti-ferromagnet is

$$\begin{aligned} H_A &= J\sum_{\mathbf{r},\boldsymbol{\delta}}[S_z^{(1)}(\mathbf{r})S_z^{(2)}(\mathbf{r}+\boldsymbol{\delta}) + \\ &\quad \frac{1}{2}S_+^{(1)}(\mathbf{r})S_-^{(2)}(\mathbf{r}+\boldsymbol{\delta}) + \frac{1}{2}S_+^{(2)}(\mathbf{r}+\boldsymbol{\delta})S_-^{(1)}(\mathbf{r})]. \end{aligned} \tag{4.21}$$

The sum over $\mathbf{r}$ in eq (4.21) extends over the sites of the first lattice; the sum over $\boldsymbol{\delta}$ extends over the four nearest neighbor displacements enumerated above. The superscripts $^{(1)}$ and $^{(2)}$ over the spin operators serve to remind us that the spin is on lattice one or lattice two respectively.

For the Heisenberg ferromagnet the exact ground state is that all the spins point maximally down along the z-axis[1]. In Dyson's boson representation the ferromagnetic

[1]Or along any other direction. The ground state of a ferromagnet spontaneously breaks rotational

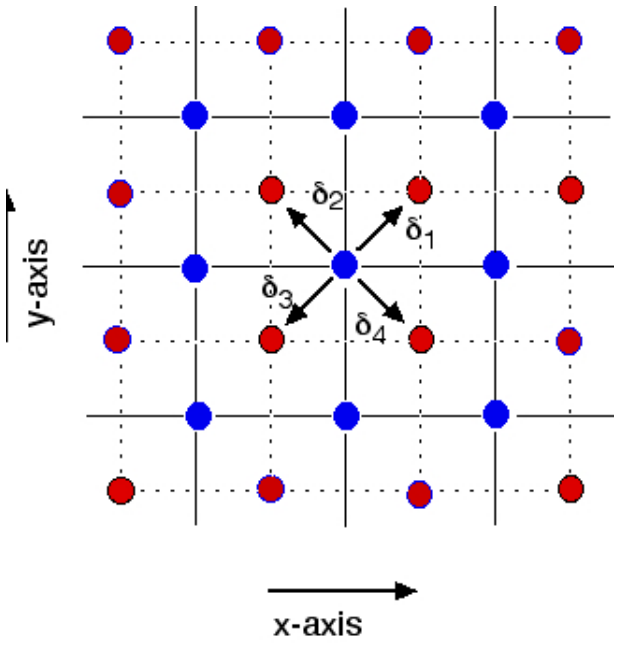

Figure 4.1: Two interpenetrating square lattices. Each site on the first square lattice (shown in blue) has four nearest neighbours that are on the second lattice (shown in red). The displacements to these four neighbors, marked $\delta_1, \delta_2, \delta_3$ and δ_4 are shown. In the cuprates there are copper atoms at both red and blue sites. At half-filling effectively there is a spin-half at each site; these spins are anti-ferromagnetically coupled. In the doped compound some fraction of the sites are occupied by holes.

ground state is the state in which no bosons are present. Anti-ferromagnets present an altogether more formidable problem. The exact ground state for an anti-ferromagnet is not known except in one dimension for the case of spin $s = 1/2$. The ideal 'Néel state' is one in which the spins on the first lattice are maximally down along the z-axis and the spins on the second lattice are maximally up along the z-axis. The Néel state is not the exact ground state of the anti-ferromagnet but it is believed to be qualitatively similar[2] and therefore a good starting point from which to obtain a more accurate picture of the ground and excited states of a Heisenberg anti-ferromagnet. Thus in representing an anti-ferromagnet in terms of Dyson bosons we shall take the Néel state to be the one with no bosons present.

symmetry and thus there is a manifold of equivalent ground states.

[2]There are many circumstances where it is known the Néel state is not even qualitatively right: in one dimension, on a triangular lattice in two dimensions or even on a square lattice in two dimensions if next nearest neighbor interactions act to frustrate Néel ordering.

To this end we establish a second mapping between a single spin and a single bose oscillator. In this second "anti-Dyson" mapping a state with spin maximally up is to be identified with the state of zero bosons. Thus we introduce the anti-Dyson basis for a spin multiplet

$$|u, A >= G_u|s, s-u > \tag{4.22}$$

where $u = 0, \ldots, 2s$. The normalization constant $G_0 = 1$ and

$$G_u = \left(1\left[1-\frac{1}{2s}\right]\left[1-\frac{2}{2s}\right]\ldots\left[1-\frac{u-1}{2s}\right]\right)^{-1/2} \tag{4.23}$$

where $u = 1, 2, \ldots, 2s$. G_u has been judiciously chosen to ensure that the spin raising operator S_+ maps to the bose annihilation operator b, as will be seen below.

Making use of eq (4.3), eq (4.22) and (4.23) it is not difficult to show

$$\begin{aligned} S_+|u; A > &= \sqrt{2s}\sqrt{u}|u-1; A >, \\ S_-|u; A > &= \sqrt{2s}\sqrt{u+1}\left[1-\frac{u}{2s}\right]|u+1; A >, \\ S_z|u; A > &= (s-u)|u; A > . \end{aligned} \tag{4.24}$$

We may now establish an anti-Dyson mapping between spins and bose oscillators by identifying the spin state $|u; A >$ with the bose oscillator state $|u)$. Thus

$$|u; A > \rightarrow |u) \tag{4.25}$$

for $u = 0, \ldots, 2s$. States with more than $2s$ bosons have no spin space counterpart. The anti-Dyson mapping allows us to export a dynamical inner product to the bose space as before. The remarks made earlier about this dynamical inner product apply mutatis mutandis [see the paragraph surrounding eq (4.14)].

The anti-Dyson mapping eq (4.25) also allows us to establish a second correspon-

dence between spin and bose operators

$$\begin{aligned} S_+ &\rightarrow \sqrt{2s}b, \\ S_- &\rightarrow \sqrt{2s}b^\dagger \left[1 - \frac{b^\dagger b}{2s}\right], \\ S_z &\rightarrow s - b^\dagger b. \end{aligned} \tag{4.26}$$

This correspondence follows from comparison of eqs (4.24) and (4.12).

Equipped with the second Dyson mapping we now return to the Heisenberg antiferromagnet. We consider two interpenetrating square lattices as above and assume that there are two kinds of lattice bosons. One kind lives on the sites of the first lattice: $b_1^\dagger(\mathbf{r}_1)$ creates this kind of boson at site $\mathbf{r}_1$; $b_1(\mathbf{r}_1)$ annihilates it. The other kind live on the second lattice and are created and annihilated by $b_2^\dagger(\mathbf{r}_2)$ and $b(\mathbf{r}_2)$ respectively. These creation and annihilation operators are adjoints of each other with respect to the kinematical inner product and are assumed to obey bosonic commutation relations

$$[b_i(\mathbf{r}), b_j(\mathbf{r}')] = \delta_{\mathbf{r},\mathbf{r}'}\delta_{ij} \tag{4.27}$$

where i and j equal 1 or 2.

We may now represent the Hamiltonian for the Heisenberg Hamiltonian eq (4.21) in terms of bosonic quasi-particles using Dyson's mapping between spins and bosons, eq (4.15) on the sites of the first lattice and using the anti-Dyson mapping eq (4.26) on the sites of the second lattice. This strategy ensures that the Néel state corresponds

to the boson vacuum and yields a bosonic form of the Heisenberg Hamiltonian

$$\begin{aligned}
\mathcal{H}_A =& Js \sum_{\mathbf{r},\boldsymbol{\delta}} [b_1^\dagger(\mathbf{r}) b_1(\mathbf{r}) + b_2^\dagger(\mathbf{r}+\boldsymbol{\delta}) b_2(\mathbf{r}+\boldsymbol{\delta})] \\
&+ Js \sum_{\mathbf{r},\boldsymbol{\delta}} [b_1^\dagger(\mathbf{r}) b_2^\dagger(\mathbf{r}+\boldsymbol{\delta}) + b_2(\mathbf{r}+\boldsymbol{\delta}) b_1(\mathbf{r})] \\
&- J \sum_{\mathbf{r},\boldsymbol{\delta}} b_1^\dagger(\mathbf{r}) b_1(\mathbf{r}) b_2^\dagger(\mathbf{r}+\boldsymbol{\delta}) b_2(\mathbf{r}+\boldsymbol{\delta}) \\
&- \frac{J}{2} \sum_{\mathbf{r},\boldsymbol{\delta}} b_1^\dagger(\mathbf{r}) b_2^\dagger(\mathbf{r}+\boldsymbol{\delta}) b_2^\dagger(\mathbf{r}+\boldsymbol{\delta}) b_2(\mathbf{r}+\boldsymbol{\delta}) \\
&- \frac{J}{2} \sum_{\mathbf{r},\boldsymbol{\delta}} b_1^\dagger(\mathbf{r}) b_1(\mathbf{r}) b_1(\mathbf{r}) b_2(\mathbf{r}+\boldsymbol{\delta}). \qquad (4.28)
\end{aligned}$$

Note that the boson Hamiltonian $\mathcal{H}_A$ is not self-adjoint under the kinematic inner product ($\mathcal{H}_A^\dagger \neq \mathcal{H}_A$) due to the terms in the last two lines of eq (4.28). However it is self-adjoint under the dynamical inner product ($\mathcal{H}_A^\star = \mathcal{H}_A$).

4.2 Doped Magnets

A typical cuprate such as $La_{2-x}Sr_xCuO_4$ consists of stacked planes of Cu atoms. Within a plane the Cu atoms are arranged in a square lattice. In the pure compound La_2CuO_4 there is one electron available per Cu atom. If electron-electron interactions were weak the electrons could hop from atom to atom via tunneling. However in the cuprates the electron-electron repulsion is strong, forbidding double occupancy of the Cu sites. Each site is therefore occupied by a single electron. The electrons are locked in place and immobile. A material like this is called a 'Mott insulator'. The only degree of freedom is the electron spin that can point up or down at each site. The decidedly unequal competition between hopping and electron-electron repulsion tends to make the spins align anti-ferromagnetically. The undoped cuprates may therefore described by the antiferromagnetic Heisenberg Hamiltonian.

In the doped compound $La_{2-x}Sr_xCuO_4$ there are only $1-x$ electrons per site and therefore a fraction x of the sites are unoccupied. The absence of electrons ("holons") can hop and when the density of holons is sufficiently high the materials are observed to exhibit strange metallic and then superconducting behavior. The competition between hopping and electron-electron repulsion for the doped compounds is described by the $t-J$ Hamiltonian. In the next section the $t-J$ Hamiltonian is formulated in a way that is particularly well suited to our present purpose.

4.2.1 Supersymmetric formulation of $t-J$ Model

In the parent compound there are two possible states for each site: spin up or spin down. In the doped material each site has three possible states: spin up, spin down or missing electron. The missing electron state corresponds to zero spin and a positive charge $+e$ on the site. In the following it will be useful to consider a spin-s generalization wherein there are $4s+1$ states per site. The site may either be in one of the $2s+1$ states $|s,m>$ with $m=-s,\ldots,s$ or in one of the $2s$ states $|s-1/2,m>$ with $m=-(s-1/2),\ldots,s-1/2$. If the site is in a spin s state, $|s,m>$, the total spin is s, the z-component of the spin is m and the site is assumed to have no charge. On the other hand if it is in a spin $s-1/2$ state, $|s-1/2,m>$, the total spin is $s-1/2$, its z-component is m and the site has a positive charge $+e$ due to the lack of one electron. In summary, whereas the basic states per site of a spin s magnet are a single spin s multiplet $|s,m>$, the basic states per site for our $t-J$ model are a "super-multiplet": a pair of multiplets with spin s and spin $s-1/2$. The physically relevant case is $s=1/2$.

Having specified the basic states at each site we must now describe the basic operators out of which the $t-J$ Hamiltonian will be built. For a magnet these operators are S_+, S_- and S_z. They satisfy the su(2) angular momentum algebra eq (4.2) and their action on the states $|s,m>$ of a spin s multiplet is well-known eq

(4.3). Now it turns out there is a super-algebra that is a natural generalization of the su(2) algebra and the $t-J$ model can be written (super)naturally in terms of the elements of this algebra.

The super-algebra has eight elements. Six of them are raising and lowering operators (also known as Weyl elements): S_+, S_-, R_+, R_-, T_+ and T_-. The remaining two are the Cartan elements A and S_z. Since this is a super-algebra the elements may also be grouped differently into commuting elements (S_+, S_-, S_z, A) and anti-commuting elements (R_+, R_-, T_+, T_-). Just as the su(2) algebra is defined by the commutation relations of its elements eq (4.2), so the super-algebra is defined by the commutation or anti-commutation relations amongst all pairs of its elements. First, there are the diagonal Weyl element relations:

$$\begin{aligned} [S_+, S_-] &= 2S_z, \\ \{R_+, R_-\} &= A + S_z \\ \{T_+, T_-\} &= A - S_z \end{aligned} \tag{4.29}$$

As usual square brackets denote commutators; curly brackets, anti-commutators. Next there are the off-diagonal Weyl commutation relations

$$\begin{aligned} &[S_+, R_+] = -T_+, && [S_-, R_+] = 0, \\ &[S_+, R_-] = 0, && [S_-, R_-] = T_-, \\ &[S_+, T_+] = 0, && [S_-, T_+] = -R_+, \\ &[S_+, T_-] = R_-, && [S_-, T_-] = 0, \end{aligned} \tag{4.30}$$

and the off-diagonal Weyl anti-commutation relations

$$\begin{aligned} \{R_+, T_+\} &= 0, & \{R_-, T_+\} &= S_+, \\ \{R_+, T_-\} &= S_+, & \{R_-, T_-\} &= 0. \end{aligned} \tag{4.31}$$

The Cartan elements A and S_z commute with each other; $[A, S_z] = 0$. The final set of defining relations are the commutators of the Weyl and Cartan elements:

$$\begin{aligned} [S_+, S_z] &= -S_+, & [S_+, A] &= 0, \\ [S_-, S_z] &= S_-, & [S_-, A] &= 0, \\ [R_+, S_z] &= \frac{1}{2}R_+, & [R_+, A] &= -\frac{1}{2}R_+, \\ [R_-, S_z] &= -\frac{1}{2}R_-, & [R_-, A] &= \frac{1}{2}R_-, \\ [T_+, S_z] &= -\frac{1}{2}T_+, & [T_+, A] &= -\frac{1}{2}T_+, \\ [T_-, S_z] &= \frac{1}{2}T_- & [T_-, A] &= \frac{1}{2}T_-. \end{aligned} \tag{4.32}$$

These relations serve to define the algebra.

Now let us describe the action of the algebra elements on the states of a supermultiplet. S_+ and S_- simply raise and lower the z-component of the spin in either multiplet:

$$\begin{aligned} S_+|s, m> &= (s-m)^{1/2}(s+m+1)^{1/2}|s, m+1>, \\ S_+|s-1/2, m> &= (s-1/2-m)^{1/2}(s+1/2+m)^{1/2} \\ &\quad |s-1/2, m+1>, \\ S_-|s, m> &= (s-m+1)^{1/2}(s+m)^{1/2}|s, m-1>, \\ S_-|s-1/2, m> &= (s+1/2-m)^{1/2}(s-1/2+m)^{1/2} \\ &\quad |s-1/2, m-1>. \end{aligned} \tag{4.33}$$

R_+ and R_- switch states between multiplets

$$\begin{aligned} R_+|s,m> &= (s+m)^{1/2}|s-1/2,m-1/2>, \\ R_+|s-1/2,m> &= 0, \\ R_-|s,m> &= 0, \\ R_-|s-1/2,m> &= (s+1/2+m)^{1/2}|s,m+1/2> . \end{aligned} \tag{4.34}$$

Note that R_+ lowers the z-component of spin by half when it changes from spin s to spin $s-1/2$. T_+ and T_- also switch states between multiplets

$$\begin{aligned} T_+|s,m> &= (s-m)^{1/2}|s-1/2,m+1/2>, \\ T_+|s-1/2,m> &= 0, \\ T_-|s,m> &= 0, \\ T_-|s-1/2,m> &= (s+1/2-m)^{1/2}|s,m-1/2>, \end{aligned} \tag{4.35}$$

but whereas R_+ lowers the z-component by half, T_+ raises it. Finally the states of the super-multiplet are eigenstates of A and S_z

$$\begin{aligned} A|s,m> &= s|s,m>, \\ A|s-1/2,m> &= (s+1/2)|s-1/2,m>, \\ S_z|s,m> &= m|s,m>, \\ S_z|s-1/2,m> &= m|s-1/2,m> . \end{aligned} \tag{4.36}$$

Thus the A value distinguishes the multiplets; the S_z value specifies the state within the multiplet. Eqs (4.33), (4.34), (4.35) and (4.36) fully describe the action of the

super-algebra elements on the states of the super-multiplet. The normalization factors in these equations follow inexorably from the commutation and anti-commutation relations that define the super-algebra. Note that the action of S_+, S_- and S_z is exactly as one would expect from the textbook theory of angular momentum; this is because these operators constitute an su(2) subalgebra of our super-algebra.

We can now write the $t-J$ Hamiltonian in supersymmetric form

$$\begin{aligned} H_{t-J} = -\tau \sum_{\mathbf{r},\boldsymbol{\delta}} [R_+(\mathbf{r}+\boldsymbol{\delta})R_-(\mathbf{r}) + R_+(\mathbf{r})R_-(\mathbf{r}+\boldsymbol{\delta}) \\ + T_+(\mathbf{r}+\boldsymbol{\delta})T_-(\mathbf{r}) + T_+(\mathbf{r})T_-(\mathbf{r}+\boldsymbol{\delta})] \\ + J \sum_{\mathbf{r},\boldsymbol{\delta}} [S_z(\mathbf{r})S_z(\mathbf{r}+\boldsymbol{\delta}) - \{A(\mathbf{r}) - 2s\}\{A(\mathbf{r}+\boldsymbol{\delta}) - 2s\} \\ + \frac{1}{2}S_+(\mathbf{r}+\boldsymbol{\delta})S_-(\mathbf{r}) + \frac{1}{2}S_+(\mathbf{r})S_-(\mathbf{r}+\boldsymbol{\delta})]. \end{aligned} \tag{4.37}$$

We assume the super-spins occupy the sites of a square lattice in a plane. The lattice position vectors are $\mathbf{r} = ma\hat{\mathbf{e}}_x + na\hat{\mathbf{e}}_y$ where m and n are integers and the sum over $\mathbf{r}$ in eq (4.37) is over m and n. $\boldsymbol{\delta}$ denotes the four nearest neighbor displacements $\pm a\hat{\mathbf{e}}_x$ and $\pm a\hat{\mathbf{e}}_y$; the sum over $\boldsymbol{\delta}$ in eq (4.37) is over these four values. The super-spin operators at different sites are assumed to commute and at a given site they are assumed to obey the super-algebra defined by eqs (4.29), (4.30), (4.31) and (4.32). Thus the $t-J$ Hamiltonian couples super-spins at neighboring sites.

Finally a word about the symmetry of the Hamiltonian, H_{t-J}. The Heisenberg Hamiltonian H_F eq (4.18) has rotational symmetry. Formally this is demonstrated by defining the total spin operators

$$S_+^{\text{tot}} = \sum_{\mathbf{r}} S_+(\mathbf{r}) \tag{4.38}$$

(and $S^{\rm tot}_-$ and $S^{\rm tot}_z$ similarly) and verifying that $[H_F, S^{\rm tot}_+] = 0$ (as well as $[H_F, S^{\rm tot}_-] = 0$ and $[H_F, S^{\rm tot}_z] = 0$). In the same way we can define the total super-spin operator

$$R^{\rm tot}_+ = \sum_{\mathbf{r}} R_+(\mathbf{r}), \tag{4.39}$$

and similarly for all other elements of the super-algebra. For the $t-J$ Hamiltonian to be supersymmetric it would have to satisfy $[H_{t-J}, R^{\rm tot}_+] = 0$, $[H_{t-J}, S^{\rm tot}_+] = 0$ and so on for all eight elements of the super-algebra. This condition is not met except for special values of the parameters t and J. The $t-J$ Hamiltonian is certainly not supersymmetric for the experimentally relevant values. Thus although the Hamiltonian is built out of supersymmetric algebra elements it is not generally supersymmetric. In this respect it is similar to SUSY extensions of the standard model for which also super-symmetry is broken.

4.2.2 Dysonization of the $t-J$ Hamiltonian

Dyson's key insight was to define magnons as bosonic with respect to a non-standard inner product. For the $t-J$ model we wish to take that scheme one step further and define a 'Dyson fermion' in addition to the Dyson bosons we have already alluded to.

In order to represent the $t-J$ Hamiltonian in terms of Dyson bosons and fermions first let us consider a single super-multiplet corresponding to the states at a single site. The basis states for a super-multiplet that we have so far adopted are the $4s+1$ states $|s, m>$ and $|s-1/2, \mu>$ where $m = -s, \ldots, +s$ and $\mu = -(s-1/2), \ldots, s-1/2$.

Following Dyson we now introduce the alternative basis states

$$|u, 0\rangle = F_{u,0}|s, -s+u\rangle, \quad |u, 1\rangle = F_{u,1}|s-1/2, -(s-1/2)+u\rangle \tag{4.40}$$

where $u = 0, \ldots, 2s$ for the $|u, 0\rangle$ states and $u = 0, \ldots, 2-1$ for the $|u, 1\rangle$ states. Thus $|0, 0\rangle$ corresponds to having a spin s at the site that is maximally down; $|u, 0\rangle$

corresponds to raising the spin u times. Similarly $|0,1\rangle$ corresponds to having a spin $s-1/2$ at the site that is maximally down; $|u,1\rangle$ corresponds to raising that spin u times. The states $|u,0\rangle$ are neutral; the states $|u,1\rangle$ correspond to having having a net charge $+e$ on the site. Usually these sites are described as holons; in light of the supersymmetry discussion above, it seems natural to associate the charge with the presence of a non-Hermitian 'Dyson fermion'.

The states in this basis are orthogonal to each other but not normalized:

$$\langle u,a|v,b\rangle = F_{u,a}^2 \delta_{ab}\delta_{uv}. \tag{4.41}$$

The normalization factors $F_{u,a}$ are chosen judiciously:

$$S^+|u,a\rangle = \sqrt{2s}\sqrt{u+1}|u+1,a\rangle; \tag{4.42}$$

so as to maintain the action of S^+ as a bosonic raising operator. This is accomplished by defining

$$|u,a\rangle = \frac{1}{\sqrt{2s^u}}\frac{1}{\sqrt{u!}}(S^+)^u|0,a\rangle, \tag{4.43}$$

which corresponds to the choice

$$F_{u,a} = \left(1-\frac{1}{2s}\right)^{1/2}\left(1-\frac{2}{2s}\right)^{1/2}\ldots\left(1-\frac{u-1+a}{2s}\right)^{1/2}. \tag{4.44}$$

The $|u,a\rangle$ basis is fully specified by eqs (4.40) and (4.44) or equivalently by eq (4.43).

We may now determine the action of all the super-spin operators in this basis.

The results are

$$\begin{aligned} S^+|u,a\rangle &= \sqrt{2s}\sqrt{u+1}|u+1,a\rangle \\ S^-|u,a\rangle &= \sqrt{2s}\left[1-\frac{u-1+a}{2s}\right]u^{1/2}|u-1,a\rangle \\ S_z|u,a\rangle &= \left(-s+u+\frac{a}{2}\right)|u,a\rangle \\ A|u,a\rangle &= a|u,a\rangle, \end{aligned} \tag{4.45}$$

for the commuting elements of the super-algebra, and

$$\begin{aligned} T^+|u,0\rangle &= \sqrt{2s}|u,1\rangle \\ T^+|u,1\rangle &= 0 \\ T^-|u,0\rangle &= 0 \\ T^-|u,1\rangle &= \left[1-\frac{u}{2s}\right]\sqrt{2s}|u,0\rangle \\ R^+|u,0\rangle &= u^{1/2}|u-1,1\rangle \\ R^+|u,1\rangle &= 0 \\ R^-|u,0\rangle &= 0 \\ R^-|u,1\rangle &= (u+1)^{1/2}|u+1,0\rangle, \end{aligned} \tag{4.46}$$

for the anti-commuting elements.

Now consider a different Hilbert space inhabited by a single Bose creation and annihilation operator pair $(b, b^\dagger)$ and a Fermi pair $(a, a^\dagger)$ that satisfy the canonical commutation relations

$$\begin{aligned} [b,b^\dagger] &= 1, \\ \{a,a^\dagger\} &= 1,\ a^2 = a^{\dagger 2} = 0. \end{aligned} \tag{4.47}$$

We also suppose $[a,b] = [a^\dagger,b] = [a,b^\dagger] = [a^\dagger,b^\dagger] = 0$. The creation and annihilation

operators are adjoints of each other under the kinematical inner product. in this Hilbert space. One can show inexorably from these assumptions that the basic states of this Hilbert space are $|u,0)$ and $|u,1)$ where $u = 0, 1, 2, \ldots$ The state $|0,0)$ has the defining characteristic

$$b|0,0) = a|0,0) = 0; \tag{4.48}$$

it contains neither a b boson not an a fermion. The state

$$|u,0) = \frac{1}{\sqrt{u!}}(b^\dagger)^u|0,0) \tag{4.49}$$

contains u bosons and no fermions. The state

$$|u,1) = \frac{1}{\sqrt{u!}}(b^\dagger)^u a^\dagger|0,0) \tag{4.50}$$

contains u bosons and one fermion. These states are orthonormal under the kinematic inner product

$$(u,a|v,b)_{\rm kin} = \delta_{u,v}\delta_{a,b}. \tag{4.51}$$

We now establish the following mapping between the states of a supermultiplet and the bose-fermi Hilbert space discussed above. The mapping is

$$|u,a\rangle \to |u,a) \tag{4.52}$$

Here $u = 0, \ldots, 2s$ for $a = 0$ and $u = 0, 2s-1$ for $a = 1$. States with more bosons have no counter-part in the super-spin space.

As before this correspondence exports a dynamical inner product to the Bose-Fermi Hilbert space

$$(u,a|v,b)_{\rm dyn} = F^2_{ua}\delta_{uv}\delta_{ab}. \tag{4.53}$$

We assume $F_{u,0} = 0$ for $u > 2s$ and $F_{u,1} = 0$ for $u > 2s-1$. Thus states with a

greater number of bosons are weightless.

The mapping eq (4.52) also allows us to establish a correspondence between super-spin and bose and fermi operators. The correspondence follows from eqs (4.45) and eq (4.46) and is as follows

$$\begin{aligned}
&S^+ \to \sqrt{2s}b^\dagger, && S^- \to \left[1 - \frac{b^\dagger b + a^\dagger a}{2s}\right]\sqrt{2s}b \\
&S_z \to (-s + b^\dagger b + \frac{1}{2}a^\dagger a), && A \to a^\dagger a \\
&T^+ \to \sqrt{2s}a^\dagger && T^- \to \left[1 - \frac{b^\dagger b + a^\dagger a}{2s}\right]\sqrt{2s}a \\
&R^+ \to ba^\dagger && R^- \to ab^\dagger.
\end{aligned} \tag{4.54}$$

Ferromagnetic t-J model

Now let us consider the $t - J$ model eq (3-37). For the cuprates we are interested in anti-ferromagnetic coupling ($J > 0$) but it is instructive to first consider the case of ferromagnetic coupling, $J < 0$.

We introduce a single boson $b(\mathbf{r}), b^\dagger(\mathbf{r})$ and a single fermion $a(\mathbf{r}), a^\dagger(\mathbf{r})$ at each site of the lattice. Bose and Fermi creation and annihilation operators at the same site are taken to be adjoints of each other under the kinematic inner product. Using the correspondence between super-spin operators and bose and fermi operators, eq (4.54), we may write the $t - J$ Hamiltonian as

$$\begin{aligned}
H_{t-J} &= -2\tau s \sum_{\mathbf{r},\boldsymbol{\delta}} [a^\dagger(\mathbf{r}+\boldsymbol{\delta})a(\mathbf{r}) + a^\dagger(\mathbf{r})a(\mathbf{r}+\boldsymbol{\delta})] \\
&\quad + \frac{1}{2}Js \sum_{\mathbf{r},\boldsymbol{\delta}} a^\dagger(\mathbf{r})a(\mathbf{r}) \\
&\quad + \frac{1}{2}Js \sum_{\mathbf{r},\boldsymbol{\delta}} [\, b^\dagger(\mathbf{r})b(\mathbf{r}) - b^\dagger(\mathbf{r}+\boldsymbol{\delta})b(\mathbf{r})] + \ldots
\end{aligned} \tag{4.55}$$

In eq (4.55) we have written out the leading quadratic term in the Dyson represen-

tation of the ferromagnetic $t-J$ Hamiltonian. At this level, it is a theory of non-interacting bosonic spin-waves ("magnons") and fermions with charge $+e$ ("magninos").

The interaction terms that were omitted in eq (4.55) and are presumably small in this representation, are given by

$$\begin{aligned} H_{\rm int} &= -\tau \sum_{\mathbf{r},\boldsymbol{\delta}} b^\dagger(\mathbf{r}) b(\mathbf{r}) a(\mathbf{r}) a^\dagger(\mathbf{r}+\boldsymbol{\delta}) \\ &\quad -\tau \sum_{\mathbf{r},\boldsymbol{\delta}} [b(\mathbf{r}+\boldsymbol{\delta}) b^\dagger(\mathbf{r}) a^\dagger(\mathbf{r}+\boldsymbol{\delta}) a(\mathbf{r}) + b^\dagger(\mathbf{r}) b(\mathbf{r}+\boldsymbol{\delta}) a^\dagger(\mathbf{r}) a(\mathbf{r}+\boldsymbol{\delta})] \\ &\quad -\frac{J}{2} \sum_{\mathbf{r},\boldsymbol{\delta}} [b^\dagger(\mathbf{r}) b(\mathbf{r}) + \frac{1}{2} a^\dagger(\mathbf{r}) a(\mathbf{r})][b^\dagger(\mathbf{r}+\boldsymbol{\delta}) b(\mathbf{r}+\boldsymbol{\delta}) + \frac{1}{2} a^\dagger(\mathbf{r}+\boldsymbol{\delta}) a(\mathbf{r}+\boldsymbol{\delta})] \\ &\quad +\frac{J}{4} \sum_{\mathbf{r},\boldsymbol{\delta}} [a^\dagger(\mathbf{r}) a(\mathbf{r}) + b^\dagger(\mathbf{r}) b(\mathbf{r})][b^\dagger(\mathbf{r}+\boldsymbol{\delta}) b(\mathbf{r})]. \end{aligned} \tag{4.56}$$

The full $t-J$ Hamiltonian, H_{t-J} is not self-adjoint under the kinematic inner product ($H_{t-J}^\dagger \neq H_{t-J}$); however it is self-adjoint under the dynamical inner product, $H_{t-J}^\star = H_{t-J}$.

Anti-ferromagnetic t-J model

For the anti-ferromagnetic $t-J$ model, as for the Heisenberg anti-ferromagnet, it is convenient to imagine a pair of interpenetrating square lattices. The $t-J$ Hamiltonian

may then be re-written

$$\begin{aligned} H_{t-J} &= -\tau \sum_{\mathbf{r},\boldsymbol{\delta}} [R_+^{(2)}(\mathbf{r}+\boldsymbol{\delta}) R_-^{(1)}(\mathbf{r}) + R_+^{(1)}(\mathbf{r}) R_-^{(2)}(\mathbf{r}+\boldsymbol{\delta})] \\ &\quad -\tau \sum_{\mathbf{r},\boldsymbol{\delta}} [T_+^{(2)}(\mathbf{r}+\boldsymbol{\delta}) T_-^{(1)}(\mathbf{r}) + T_+^{(1)}(\mathbf{r}) T_-^{(2)}(\mathbf{r}+\boldsymbol{\delta})] \\ &\quad +J \sum_{\mathbf{r},\boldsymbol{\delta}} [S_z^{(1)}(\mathbf{r}) S_z^{(2)}(\mathbf{r}+\boldsymbol{\delta}) - \{A^{(1)}(\mathbf{r}) - 2s\}\{A^{(2)}(\mathbf{r}+\boldsymbol{\delta}) - 2s\}] \\ &\quad +\frac{J}{2} \sum_{\mathbf{r},\boldsymbol{\delta}} [S_+^{(2)}(\mathbf{r}+\boldsymbol{\delta}) S_-^{(1)}(\mathbf{r}) + S_+^{(1)}(\mathbf{r}) S_-^{(2)}(\mathbf{r}+\boldsymbol{\delta})]. \end{aligned} \tag{4.57}$$

The sum over $\mathbf{r}$ in eq (4.57) extends over the sites of the first lattice; the sum over $\boldsymbol{\delta}$ extends over the four nearest neighbors of each site. The superscripts $^{(1)}$ and $^{(2)}$ over the super-spin operators serve to remind us that the spin is on lattice one or lattice two respectively.

At least for light doping it makes sense to assume that the Néel state is a good starting point for the ground state of the $t-J$ model. In the Néel state the spin is maximally down at each site of the first lattice; it is maximally up at each site of the second lattice. The magnitude of the spin is $s-1/2$ at sites occupied by a hole. It is s at all other sites. In representing the Néel state in terms of Dyson bosons and fermions we shall take the Néel state to have zero bosons and to have Dyson fermions at all the sites with holes.

To this end we establish a second mapping between the states of a single super-spin and the Hilbert space of a single boson and fermion. In this mapping we identify the states with spin maximally up as the state with zero bosons whereas before we had assigned this part to spin maximally down. Thus we introduce the basis

$$\begin{aligned} |u,0\rangle &= F_{u,0}|s, s-u\rangle \\ |u,1\rangle &= F_{u,1}|s-1/2, s-1/2-u\rangle \end{aligned} \tag{4.58}$$

in place of eq (4.40). This time we choose

$$F_{u,a} = \left(1 - \frac{1}{2s}\right)^{-1/2} \left(1 - \frac{2}{2s}\right)^{-1/2} \ldots \left(1 - \frac{u-1+a}{2s}\right)^{-1/2}. \tag{4.59}$$

As before we then establish a mapping $|u, a\rangle$ between the states of the super-spin and the states $|u, a)$ of a bose-fermi system. By virtue of this correspondence we obtain a second mapping between super-spin and bose and fermi operators:

$$\begin{aligned}
&S^+ \to \sqrt{2s}b && S^- \to \sqrt{2s}b^\dagger \left(1 - \frac{b^\dagger b + a^\dagger a}{2s}\right), \\
&S_z \to (s - b^\dagger b - \frac{1}{2}a^\dagger a), && A \to a^\dagger a, \\
&R^+ \to \sqrt{2s}\left(1 - \frac{b^\dagger b}{2s}\right), && R^- \to \sqrt{2s}a, \\
&T^+ \to ba^\dagger, && T^- \to b^\dagger a.
\end{aligned} \tag{4.60}$$

In order to write the $t-J$ Hamiltonian in terms of bosons and fermions we use the first correspondence eq (4.54) on the first lattice and the second correspondence eq (4.60) on the second lattice. Keeping the leading terms up to cubic order we obtain a novel representation of the $t-J$ Hamiltonian in terms of bosons and fermions:

$$\begin{aligned}
h_{t-J} \;=\; & Js \sum_{\mathbf{r},\boldsymbol{\delta}} [b_1^\dagger(\mathbf{r}) b_1(\mathbf{r}) + b_2^\dagger(\mathbf{r}+\boldsymbol{\delta}) b_2(\mathbf{r}+\boldsymbol{\delta})] \\
& + Js \sum_{\mathbf{r},\boldsymbol{\delta}} [a_1^\dagger(\mathbf{r}) a_1(\mathbf{r}) + a_2^\dagger(\mathbf{r}+\boldsymbol{\delta}) a_2(\mathbf{r}+\boldsymbol{\delta})] \\
& + Js \sum_{\mathbf{r},\boldsymbol{\delta}} [b_1(\mathbf{r}) b_2(\mathbf{r}+\boldsymbol{\delta}) + b_2^\dagger(\mathbf{r}+\boldsymbol{\delta}) b_1^\dagger(\mathbf{r})] \\
& - \tau\sqrt{2s} \sum_{\mathbf{r},\boldsymbol{\delta}} [a_1^\dagger(\mathbf{r}) a_2(\mathbf{r}+\boldsymbol{\delta}) b_1(\mathbf{r}) + a_2^\dagger(\mathbf{r}+\boldsymbol{\delta}) a_1(\mathbf{r}) b_1^\dagger(\mathbf{r})] \\
& - \tau\sqrt{2s} \sum_{\mathbf{r},\boldsymbol{\delta}} [a_1^\dagger(\mathbf{r}) a_2(\mathbf{r}+\boldsymbol{\delta}) b_2^\dagger(\mathbf{r}+\boldsymbol{\delta}) + a_2^\dagger(\mathbf{r}+\boldsymbol{\delta}) a_1(\mathbf{r}) b_2(\mathbf{r}+\boldsymbol{\delta})] + \ldots
\end{aligned} \tag{4.61}$$

The remaining interaction terms which are quartic and quintic are presumably small in this representation.

The essential physics of the $t - J$ model in this regime is thus revealed to be that of charged fermions hopping in a background of spin-waves. This represents a novel formulation of the $t - J$ model; it is a tantalizing possibility that the non-Hermitian quasi-particles here defined may illuminate the underlying physics of high T_c materials.

Chapter 5

Conclusion

Though it is one of the fundamental assumptions of quantum mechanics, the requirement that operators be Hermitian is not crucial to the construction of the theory. We have seen that all of the virtues of Hermiticity are retained by adopting a different set of assumptions, in which the parity and time-reversal operators figure largely.

It may seem as though there is no compelling *a priori* reason to do this, given that Hermitian quantum mechanics has been serving physicists well since the early 20th century. The philosophically- minded or adventurous physicist may pursue such modifications to fundamental quantum mechanics in the interest of exploiting the totalitarian principle, but even taking a conservative view one can see *a posteriori* that the ends justify the means in this case. By pursuing non-Hermitian quantum mechanics we open up for analysis entire classes of Hamiltonians that were previously off-limits. The extension to T_{odd} allows fermionic systems to be analyzed in the PT framework constructed by Bender *et al.* Certainly one of the most interesting fermionic systems to study is the fundamental particle itself, governed by the Dirac equation. Dirac of course assumed that operators were Hermitian but remarkably that assumption turns out to be moot. The fundamental Dirac fermion emerges from the PT theory exactly as it does in the Hermitian theory, identical in every

aspect. Hermiticity, therefore, is not the only type of symmetry that can give rise to the world we see around us, to the extent that the world we see around us is described by the Dirac equation. The world world us may contain fermions that are not described by the Dirac equation; for example it is not clear whether neutrinos are Dirac or Majorana. The observation of flavor oscillations precipitated a sea change in the Standard Model description of the neutrino; by flavor oscillating the neutrino acquired mass because the only way to have mass differences in neutrino species is if there is a non-zero mass matrix.

A most remarkable feature of non-Hermitian quantum mechanics is that a new type of particle is described by the higher dimensional representations of the PT Dirac equation. This particle can have zero effective mass even when the mass matrix is non-zero, which suggests that it may be capable of describing massless neutrino oscillations. In order to determine whether Model 8 is a viable candidate for the neutrino, the theory must be generalized to include three species, among other things. This is the topic of future work. Regardless of whether Model 8 describes the neutrino it is a breakthrough in its own right. An independent particle comprised of a quartet of spinors is a new beast, unlike any known fermion, and as such warrants further investigation.

The applications of non-Hermitian quantum mechanics may extend beyond the realm of fundamental physics into the emergent world of condensed matter. Dyson was unwittingly discovered non-Hermitian quantum mechanics in 1956 when he found that high precision calculations of interacting spin waves in a ferromagnet were facilitated by use of a non-Hermitian Hamiltonian. Dyson's technique of defining quasiparticles with respect to a non-standard inner product allows for a novel way of writing the t-J Hamiltonian; this new form of the t-J Hamiltonian may prove more wieldy to calculations and even shed some light on the physics that underlies high temperature superconductivity, arguably the most outstanding problem in theoretical

condensed matter physics [8].

Bibliography

[1] P.A.M. Dirac, *Principles of Quantum Mechanics* (Cambridge University Press, Cambridge, 1930).

[2] S. Weinberg, *Precision Tests of Quantum Mechanics*, Phys Rev Letters **62**, 485 (1989).

[3] F. Wilczek, *Quantum Mechanics of Fractional-Spin Particle*, Phys. Rev. Lett. **49**, 957 (1982).

[4] C.M. Bender and S. Boettcher, *Real Spectra in Non-Hermitian Hamiltonians Having PT Symmetry*, Phys Rev Lett **80**, 5243 (1998); C.M. Bender, *Making Sense of Non-Hermitian Hamiltonians*, Rept. Prog. Phys. **70**, 947 (2007).

[5] White, T.H. *The Once and Future King*, (Putnam, New York, 1958).

[6] Johnson, G. *Strange Beauty: Murray Gell-Mann and the Revolution in Twentieth-Century Physics* (Random House, New York 1999).

[7] F.J. Dyson, *General Theory of Spin-Wave Interactions*, Phys Rev **102**, 1217 (1956).

[8] Leggett,A.J. *What DO we know about high T_c?*, Nature Physics, **2**, 134 (2006).

[9] Smith, L. *Linear Algebra*, Springer-Verlag, (1978) p. 268.

[10] Smith, L. *Linear Algebra*, Springer-Verlag, (1978) p. 283.

[11] Weinberg, *The Quantum Theory of Fields*, (Cambridge University Press,1996).

[12] Townsend, J. *A Modern Approach to Quantum Mechanics* (University Science Books, California, 2000).

[13] Mehta, M.L. *Random Matrix Theory*(Academic Press, New York, 1991).

[14] C.M. Bender, M.V. Berry and A. Mandilara, J. Phys. **A35**, L467 (2002).

[15] Morse,P.M and Feshbach,H. *Methods of Mathematical Physics.*(McGraw-Hill,1953)

[16] P.A.M. Dirac, *The Quantum Theory of the Electron*, Proc. Roy. Soc **A117**, 610 (1928).

[17] K. Jones-Smith and H. Mathur, *A New Class of non-Hermitian Quantum Hamiltonians with PT symmetry.*

[18] K. Jones-Smith and H. Mathur (in progress).

[19] For a review see, for example, M.C. Gonzalez-Garcia and Yosef Nir, *Neutrino masses and mixing: evidence and implications*, Rev. Mod. Phys. **75**, 345 (2003).

[20] P.W. Anderson, *Basic Notions of Condensed Matter Physics* (Westview Press, 2nd edition, 1997)

[21] R.P. Feynman, *Statistical Mechanics*,(Westview Press, 2nd edition, 1998).