The Fermionic canonical commutation relations and the Jordan-Wigner transform

Michael A. Nielsen[1,*]

[1] *School of Physical Sciences, The University of Queensland, Brisbane, Queensland 4072, Australia*

(Dated: July 29, 2005)

I. INTRODUCTION

When you learn undergraduate quantum mechanics, it starts out being all about wavefunctions and Hamiltonians, finding energy eigenvalues and eigenstates, calculating measurement probabilities, and so on.

If your physics education was anything like mine, at some point a mysterious jump occurs. People teaching more advanced subjects, like quantum field theory, condensed matter physics, or quantum optics, start "imposing canonical commutation relations" on various field operators.

Any student quickly realizes that "imposing canonical commutation relations" is extremely important, but, speaking personally, at the time I found it quite mysterious exactly what people meant by this phrase. It's only in the past few years that I've obtained a satisfactory understanding of how this works, and understood why I had such trouble in the first place.

These notes contain two parts. The first part is a short tutorial explaining the Fermionic canonical commutation relations (CCRs) from an elementary point of view: the different meanings they can have, both mathematical and physical, and what mathematical consequences they have. I concentrate more on the mathematical consequences than the physical in these notes, since having a good grasp of the former seems to make it relatively easy to appreciate the latter, but not so much vice versa. I may come back to the physical aspect in some later notes.

The second part of the notes describes a beautiful application of the Fermionic CCRs known as the *Jordan-Wigner transform*. This powerful tool allows us to map a system of interacting qubits onto an *equivalent* system of interacting Fermions, or, vice versa, to map a system of Fermions onto a system of qubits.

Why is this kind of mapping interesting? It's interesting because it means that anything we understand about one type of system (e.g., Fermions) can be immediately applied to learn something about the other type of system (e.g., qubits).

I'll describe an application of this idea, taking what appears to be a very complicated one-dimensional model of interacting spin-$\frac{1}{2}$ particles, and showing that it is equivalent to a simple model of non-interacting Fermions. This enables us to solve for the energy spectrum and eigenstates of the original Hamiltonian. This has, of course, intrinsic importance, since we'd like to understand such spin models — they're important for a whole bundle of reasons, not the least of which is that they're perhaps the simplest systems in which quantum phase transitions occur. But this example is only the tip of a much larger iceberg: the idea that the best way of understanding some physical systems may be to map those systems onto mathematically equivalent but physically quite different systems, whose properties we already understand. Physically, we say that we introduce a *quasiparticle description* of the original system, in order to simplify its understanding. This idea has been of critical importance in much of modern physics, including the understanding of superconductivity and the quantum Hall effect.

Another application of the Jordan-Wigner transform, which I won't describe in detail here, but which might be of interest to quantum computing people, is to the quantum simulation of a system of Fermions. In particular, the Jordan-Wigner transform allows us to take a system of interacting Fermions, and map it onto an equivalent model of interacting spins, which can then, in principle, be simulated using standard techniques on a quantum computer. This enables us to use quantum computers to efficiently simulate systems of interacting Fermions. This is not a trivial problem, as can be seen from the following quote from Feynman, in his famous 1982 paper on quantum computing:

> "[with Feynman's proposed quantum computing device] could we imitate every quantum mechanical system which is discrete and has a finite number of degrees of freedom? I know, almost certainly, that we could do that for any quantum mechanical system which involves Bose particles. I'm not sure whether Fermi particles could be described by such a system. So I leave that open."

It wasn't until 20 years later, and the work by Somma, Ortiz, Gubernatis, Knill and Laflamme (Physical Review A, 2002) that this problem was resolved, by making use of the Jordan-Wigner transform.

II. FERMIONS

A. The canonical commutation relations for Fermions

Suppose we have a set of operators $a_1, \ldots, a_n$ acting on some Hilbert space V. Then we say that these operations

*nielsen@physics.uq.edu.au and www.qinfo.org/people/nielsen

satisfy the *canonical commutation relations (CCRs) for Fermions* if they satisfy the equations

$$\{a_j, a_k^\dagger\} = \delta_{jk} I; \quad \{a_j, a_k\} = 0, \tag{1}$$

where $\{A, B\} \equiv AB + BA$ is the anticommutator. Note that when we take the conjugate of the second of these relations we obtain $\{a_j^\dagger, a_k^\dagger\} = 0$, which is sometimes also referred to as one of the CCRs. It is also frequently useful to set $j = k$, giving $a_j^2 = (a_j^\dagger)^2 = 0$.

How should one understand the CCRs? One way of thinking about the CCRS is in an axiomatic mathematical sense. In this way of thinking they are purely mathematical conditions that can be imposed on a set of matrices: for a given set of matrices, we can simply check and verify whether those matrices satisfy or do not satisfy the CCRs. For example, when the state space V is that of a single qubit, we can easily verify that the operator $a = |0\rangle\langle 1|$ satisfies the Fermionic CCRs. From this axiomatic point of view the question to ask is what consequences about the structure of V and the operators a_j can be deduced from the fact that the CCRs hold.

There's also a more sophisticated (but still entirely mathematical) way of understanding the CCRs, as an instance of the relationship between abstract algebraic objects (such as groups, Lie algebras, or Hopf algebras), and their representations as linear maps on vector spaces. My own knowledge of representation theory is limited to a little representation theory of finite groups and of Lie algberas, and I certainly do not see the full context in the way an expert on representation theory would. However, even with that limited background, one can see that there are common themes and techniques: what may appear to be an isolated technique or trick is often really an instance of a much deeper idea or set of ideas that become obvious once once one has enough broad familiarity with representation theory. I'm not going to pursue this point of view in these notes, but thought it worth mentioning for the sake of giving context and motivation to the study of other topics.

Finally, there's a physical way in which we can understand the CCRs. When we want to describe a system containing Fermions, one way to begin is to start by writing down a set of operators satisfying the CCRs, and then to try to guess what sort of Hamiltonian involving those operators could describe the interactions observed in the system, often motivated by classical considerations, or other rules of thumb. This is, for example, the sort of point of view pursued in the BCS theory of superconductivity, and which people are trying to pursue in understanding high temperature superconductors.

Of course, one can ask why physicists want to use operators satisfying the Fermionic CCRs to describe a system of Fermions, or why anyone, mathematician or physicist, would ever write down the CCRs in the first place. These are good questions, which I'm not going to try to answer here, although one or both questions might make a good subject for some future notes. (It is, of course, a *lot* easier to answer these questions once you understand the material I present here.)

Instead, I'm going to approach the Fermionic CCRs from a purely mathematical point of view, asking the question "What can we deduce from the fact that a set of operators satisfying the CCRs exists?" The surprising answer is that we can deduce quite a lot about the structure of V and the operators a_j simply from the fact that the a_j satisfy the canonical commutation relations!

B. Consequences of the fermionic CCRs

We will assume that the vector space V is finite dimensional, and that there are n operators $a_1, \ldots, a_n$ acting on V and satisfying the Fermionic CCRs. At the end of this paragraph we're going to give a broad outline of the steps we go through. Upon a first read, some of these steps may appear a little mysterious to the reader not familiar with representation theory. In particular, please don't worry if you get a little stuck in your understanding of the outline at some points, as the exposition is very much at the bird's-eye level, and not all detail is visible at that level. Nonetheless, the reason for including this broad outline is the belief that repeated study will pay substantial dividends, if it is read in conjunction with the more detailed exposition to follow, or similar material on, e.g., representations of the Lie algebra $su(2)$. Indeed, the advantage of operating at the bird's-eye level is that it makes it easier to see the connections between these ideas, and the use of similar ideas in other branches of representation theory.

- We'll start by showing that the operators $a_j^\dagger a_j$ are positive Hermitian operators with eigenvalues 0 and 1.

- We'll show that a_j acts as a *lowering operator* for $a_j^\dagger a_j$, in the sense that if $|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 1, then $a_j|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 0. If $|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 0, then $a_j|\psi\rangle$ vanishes.

- Similarly, $a_j^\dagger$ acts as a *raising operator* for $a_j^\dagger a_j$, in the sense that if $|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 0, then $a_j^\dagger|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 1. If $|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 1, then $a_j^\dagger|\psi\rangle$ vanishes.

- We prove that the operators $a_j^\dagger a_j$ form a mutually commuting set of Hermitian matrices, and thus there exists a state $|\psi\rangle$ which is a simultaneous eigenstate of $a_j^\dagger a_j$ for *all* values $j = 1, \ldots, n$.

- By raising and lowering the state $|\psi\rangle$ in all possible combinations, we'll construct a set of 2^n orthonormal states which are simultaneous eigenstates of the $a_j^\dagger a_j$. The corresponding vector of eigenvalues uniquely labels each state in this orthonormal basis.

- Suppose the vector space spanned by these 2^n simultaneous eigenstates is W. At this point, we know that a_j and $a_j^\dagger$ map W into W, and, indeed, we know everything about the action a_j and $a_j^\dagger$ have on W.

- Suppose we define $W_\perp$ to be the orthocomplement of W in V. Then we'll show that the a_j and $a_j^\dagger$ map $W_\perp$ into itself, and their restrictions to $W_\perp$ satisfy the Fermionic CCRs. We can then repeat the above procedure, and identify a 2^n-dimensional subspace of $W_\perp$ on which we know the action of the a_j and $a_j^\dagger$ exactly.

- We iterate this procedure until $W_\perp$ is the trivial vector space, at which point it is no longer possible to continue. At this point we have established an orthonormal basis for the whole of V with respect to which we can explicilty write down the action of both a_j and $a_j^\dagger$.

Let's go through each of these steps in more detail.

The $a_j^\dagger a_j$ are positive Hermitian with eigenvalues 0 and 1: Observe that the $a_j^\dagger a_j$ are positive (and thus Hermitian) matrices. We will show that $(a_j^\dagger a_j)^2 = a_j^\dagger a_j$, and thus the eigenvalues of $a_j^\dagger a_j$ are all 0 or 1.

To see this, observe that $(a_j^\dagger a_j)^2 = a_j^\dagger a_j a_j^\dagger a_j = -(a_j^\dagger)^2 a_j^2 + a_j^\dagger a_j$, where we used the CCR $\{a_j, a_j^\dagger\} = I$. Note also that $a_j^2 = 0$ by the CCR $\{a_j, a_j\} = 0$. It follows that $(a_j^\dagger a_j)^2 = a_j^\dagger a_j$, as claimed.

The a_j are lowering operators: Suppose $|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 1. Then we claim that $a_j|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 0. To see that $a_j|\psi\rangle$ is normalized, note that $\langle\psi|a_j^\dagger a_j|\psi\rangle = \langle\psi|\psi\rangle = 1$, where we used the fact that $|\psi\rangle$ is an eigenstate of $a_j^\dagger a_j$ with eigenvalue 1 to establish the first equality. To see that it has eigenvalue 0, note that $a_j^\dagger a_j a_j|\psi\rangle = 0$, since $\{a_j, a_j\} = 0$.

Exercise: Suppose $|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 0. Show that $a_j|\psi\rangle = 0$.

The a_j are raising operators: Suppose $|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 0. Then we claimed that $a_j^\dagger|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 1.

To see the normalization, we use the CCR $\{a_j, a_j^\dagger\} = I$ to deduce $\langle\psi|a_j a_j^\dagger|\psi\rangle = -\langle\psi|a_j^\dagger a_j|\psi\rangle + \langle\psi|\psi\rangle$. But $a_j^\dagger a_j|\psi\rangle = 0$, by the eigenvalue assumption, and $\langle\psi|\psi\rangle = 1$, whence $\langle\psi|a_j a_j^\dagger|\psi\rangle = 1$, which is the desired normalization condition.

To see that $a_j^\dagger|\psi\rangle$ is an eigenstate with eigenvalue 1, use the CCR $\{a_j, a_j^\dagger\} = I$ to deduce that $a_j^\dagger a_j a_j^\dagger|\psi\rangle = -a_j^\dagger a_j^\dagger a_j|\psi\rangle + a_j^\dagger|\psi\rangle = a_j^\dagger|\psi\rangle$, where the final equality can be deduced either from the assumption that $a_j^\dagger a_j|\psi\rangle = 0$, or from the CCR $\{a_j^\dagger, a_j^\dagger\} = 0$. This is the desired eigenvalue equation for $a_j^\dagger|\psi\rangle$.

Exercise: Suppose $|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 1. Show that $a_j^\dagger|\psi\rangle = 0$.

The $a_j^\dagger a_j$ form a mutually commuting set of observables: To see this, let $j \neq k$, and apply the CCRs repeatedly to obtain $a_j^\dagger a_j a_k^\dagger a_k = a_k^\dagger a_k a_j^\dagger a_j$, which is the desired commutativity.

Existence of a common eigenstate: It is well known that a mutually commuting set of Hermitian operators possesses a common eigenbasis. This fact is usually taught in undergraduate quantum mechanics courses; for completeness, I've included a simple proof in an appendix to these notes. We won't make use of the full power of this result here, but instead simply use the fact that there exists a normalized state $|\psi\rangle$ which is a simultaneous eigenstate of all the $a_j^\dagger a_j$ operators. In particular, for all j we have:

$$a_j^\dagger a_j|\psi\rangle = \alpha_j|\psi\rangle, \tag{2}$$

where for each j either $\alpha_j = 0$ or $\alpha_j = 1$. It will be convenient to assume that $\alpha_j = 0$ for all j. This assumption can be made without loss of generality, by applying lowering operators to the $|\psi\rangle$ for each j such that $\alpha_j = 1$, resulting in a normalized state $|\text{vac}\rangle$ such that $a_j^\dagger a_j|\text{vac}\rangle = 0$ for all j.

Defining an orthonormal basis: For any vector $\alpha = (\alpha_1, \ldots, \alpha_n)$, where each $\alpha_j = 0$ or 1, define a corresponding state:

$$|\alpha\rangle \equiv (a_1^\dagger)^{\alpha_1} \ldots (a_n^\dagger)^{\alpha_n}|\text{vac}\rangle. \tag{3}$$

It is clear that there are 2^n such states $|\alpha\rangle$, and that they form an orthonormal set spanning a subspace of V that we shall call W.

The action of the a_j and $a_j^\dagger$ on W: How do a_j and $a_j^\dagger$ act on W? Stated another way, how do they act on the orthonormal basis we have constructed for W, the states $|\alpha\rangle$? Applying the CCRs and the definition of the states $|\alpha\rangle$ it is easy to verify that the action of a_j is as follows:

- Suppose $\alpha_j = 0$. Then $a_j|\alpha\rangle = 0$.

- Suppose $\alpha_j = 1$. Let α' be that vector which results when the jth entry of α is changed to 0. Then $a_j|\alpha\rangle = -(-1)^{s_\alpha^j}|\alpha'\rangle$, where $s_\alpha^j \equiv \sum_{k=1}^{j-1} \alpha_k$.

The action of $a_j^\dagger$ on W is similar:

- Suppose $\alpha_j = 0$. Let α' be that vector which results when the jth entry of α is changed to 1. Then $a_j^\dagger|\alpha\rangle = -(-1)^{s_\alpha^j}|\alpha'\rangle$, where $s_\alpha^j \equiv \sum_{k=1}^{j-1} \alpha_k$.
- Suppose $\alpha_j = 1$. Then $a_j^\dagger|\alpha\rangle = 0$.

Action of a_j and $a_j^\dagger$ on $W_\perp$: We have described the action of the a_j and the $a_j^\dagger$ on the subspace W. What of the action of these operators on the remainder of V? To answer that question, we first show that a_j and $a_j^\dagger$ map the orthocomplement $W_\perp$ into itself.

To see this, let $|\psi\rangle \in W_\perp$, and consider $a_j|\psi\rangle$. We wish to show that $a_j|\psi\rangle \in W_\perp$ also, i.e., that for any $|\phi\rangle \in W$ we have $\langle\phi|a_j|\psi\rangle = 0$. This follows easily by considering the complex conjugate quantity $\langle\psi|a_j^\dagger|\phi\rangle$, and observing that $a_j^\dagger|\phi\rangle \in W$, since $|\phi\rangle \in W$, and thus $\langle\psi|a_j^\dagger|\phi\rangle = 0$. A similar argument shows that $a_j^\dagger$ maps $W_\perp$ into itself.

Consider now the operators $\tilde{a}_j$ obtained by restricting a_j to $W_\perp$. Provided $W_\perp$ is nontrivial it is clear that these operators satisfy the CCRs on $W_\perp$. Repeating the above argument, we can therefore identify a 2^n-dimensional subspace of $W_\perp$ on which we can compute the action of $\tilde{a}_j$ and $\tilde{a}_j^\dagger$, and thus of a_j and $a_j^\dagger$.

We may iterate this procedure many times, but the fact that V is finite dimensional means that the process must eventually terminate. At the point of termination we will have broken up V as a direct sum of some finite number d of orthonormal 2^n-dimensional vector spaces, $W_1, W_2, \ldots, W_d$, and on each vector space we will have an orthonormal basis with respect to which the action of a_j and $a_j^\dagger$ is known precisely.

Stated another way, we can introduce an orthonormal basis $|\alpha, k\rangle$ for V, where α runs over all n-bit vectors, and $k = 1, \ldots, d$, and such that the action of the a_j and $a_j^\dagger$ is to leave k invariant, and to act on $|\alpha\rangle$ as described above. In this representation it is clear that V can be regarded as a tensor product $C^{2^n} \otimes C^d$, with the action of a_j and $a_j^\dagger$ trivial on the C^d component. We will call this the *occupation number representation* for the Fermi algebra a_j.

It's worth pausing to appreciate what has been achieved here: starting only from the CCRs for $a_1, \ldots, a_n$ we have proved that V can be broken down into a tensor product of a 2^n-dimensional vector space and a d-dimensional vector space, with the a_js acting nontrivially only on the 2^n-dimensional component. Furthermore, the action of the a_js is completely known. I think it's quite remarkable that we can say so much: at the outset it wasn't even obvious that the dimension of V should be a multiple of 2^n!

When $d = 1$ we will call this the *fundamental representation* for the Fermionic CCRs. (This is the terminology I use, but I don't know if it is standard or not.) Up to a change of basis it is clear that all other representations can be obtained by taking a tensor product of the fundamental representation with the trivial action on a d-dimensional vector space.

C. Diagonalizing a Fermi quadratic Hamiltonian

Suppose $a_1, \ldots, a_n$ satisfy the Fermionic CCRs, and we have a system with Hamiltonian

$$H_{\text{free}} = \sum_j \alpha_j a_j^\dagger a_j, \tag{4}$$

where $\alpha_j \geq 0$ for each value of j. In physical terms, this is the Hamiltonian used to describe a system of free, i.e., non-interacting, Fermions.

Such Hamiltonians are used, for example, in the simplest possible quantum mechanical model of a metal, the Drude-Sommerfeld model, which treats the conduction electrons as free Fermions. Such a model may appear pretty simplistic (especially after we solve it, below), but actually there's an amazing amount of physics one can get out of such simple models. I won't dwell on these physical consequences here, but if you're unfamiliar with the Drude-Sommerfeld theory, you could profitably spend a couple of hours looking at the first couple of chapters in a good book on condensed matter physics, like Ashcroft and Mermin's "Solid State Physics", which explains the Drude-Sommerfeld model and its consequences in detail. (Why such a simplistic model does such a great job of describing metals is another long story, which I may come back to in a future post.)

Returning to the abstract Hamiltonian H_{free}, the positivity of the operators $a_j^\dagger a_j$ implies that $\langle\psi|H_{\text{free}}|\psi\rangle \geq 0$ for any state $|\psi\rangle$, and thus the ground state energy of H_{free} is non-negative. However, our earlier construction also shows that we can find at least one state $|\text{vac}\rangle$ such that $a_j^\dagger a_j|\text{vac}\rangle = 0$ for all j, and thus $H_{\text{free}}|\text{vac}\rangle = 0$. It follows that the ground state energy of H_{free} is exactly 0.

This result is easily generalized to the case where the α_j have any sign, with the result that the ground state energy is $\sum_j \min(0, \alpha_j)$, and the ground state $|\psi\rangle$ is obtained from $|\text{vac}\rangle$ by applying the raising operator a_j for all j with $\alpha_j < 0$. More generally, the allowed energies of the excited states of this system correspond to sums over subsets of the α_j.

Exercise: Express the excited states of the system in terms of $|\text{vac}\rangle$.

Just by the way, readers with an interest in computational complexity theory may find it interesting to note a connection between the spectrum of H_{free} and the *Subset-Sum* problem from computer science. The *Subset-Sum* problem is this: given a set of integers $x_1, \ldots, x_n$, with repetition allowed, is there a subset of those integers which adds up to a desired target, t? Obviously, the problem of determining whether H_{free} has a particular energy is equivalent to the *Subset-Sum* problem, at least in the case where the α_j are integers. What is interesting is that the *Subset-Sum* problem is known to be **NP-Complete**, in the language of computational complexity theory, and thus is regarded as computationally intractable. As a consequence, we deduce that the problem of determining whether a particular value for energy is in the spectrum of $H_{\rm free}$ is in general **NP-Hard**, i.e., at least as difficult as the **NP-Complete** problems. Similar results hold for the more general Fermi Hamiltonians considered below. Furthermore, this observation suggests the possibility of an interesting link between the physical problem of estimating the density of states, and classes of problems in computational complexity theory, such as the counting classes (e.g., **#P**), and also to approximation problems.

Let's generalize our results about the spectrum of $H_{\rm free}$. Suppose now that we have the Hamiltonian

$$H = \sum_{jk} \alpha_{jk} a_j^\dagger a_k. \tag{5}$$

Taking the adjoint of this equation we see that in order for H to be hermitian, we must have $\alpha_{jk}^* = \alpha_{kj}$, i.e., the matrix α whose entries are the α_{jk} is itself hermitian.

Suppose we introduce new operators $b_1, \ldots, b_n$ defined by

$$b_j \equiv \sum_{k=1}^n \beta_{jk} a_k, \tag{6}$$

where β_{jk} are complex numbers. We are going to try to choose the β_{jk} so that (1) the operators b_j satisfy the Fermionic CCRs, and (2) when expressed in terms of the b_j, the Hamiltonian H takes on the same form as $H_{\rm free}$, and thus can be diagonalized.

We begin by looking for conditions on the complex numbers β_{jk} such that the b_j operators satisfy Fermionic CCRs. Computing anticommutators we find

$$\{b_j, b_k^\dagger\} = \sum_{lm} \beta_{jl}\beta_{km}^* \{a_l, a_m^\dagger\}. \tag{7}$$

Substituting the CCR $\{a_l, a_m^\dagger\} = \delta_{lm} I$ and writing $\beta_{km}^* = \beta_{mk}^\dagger$ gives

$$\{b_j, b_k^\dagger\} = \sum_{lm} \beta_{jl}\delta_{lm}\beta_{mk}^\dagger I = (\beta\beta^\dagger)_{jk} I \tag{8}$$

where $\beta\beta^\dagger$ denotes the matrix product of the matrix β with entries β_{jl} and its adjoint $\beta^\dagger$. To compute $\{b_j, b_k\}$ we use the linearity of the anticommutator bracket in each term to express $\{b_j, b_k\}$ as a sum over terms of the form $\{a_l, a_m\}$, each of which is 0, by the CCRs. As a result, we have:

$$\{b_j, b_k\} = 0. \tag{9}$$

It follows that provided $\beta\beta^\dagger = I$, i.e., provided β is unitary, the operators b_j satisfy the Fermionic CCRs.

Let's assume that β is unitary, and change our notation, writing $u_{jk} \equiv \beta_{jk}$ in order to emphasize the unitarity of this matrix. We now have

$$b_j = \sum_k u_{jk} a_k. \tag{10}$$

Using the unitarity of u we can invert this equation to obtain

$$a_j = \sum_k u_{jk}^\dagger b_k. \tag{11}$$

Substituting this expression and its adjoint into H and doing some simplification gives us

$$H = \sum_{lm} (u\alpha u^\dagger)_{lm} b_l^\dagger b_m. \tag{12}$$

Since α is hermitian, we can choose u so that $u\alpha u^\dagger$ is diagonal, with entries λ_j, the eigenvalues of α, giving us

$$H = \sum_j \lambda_j b_j^\dagger b_j. \tag{13}$$

This is of the same form as $H_{\rm free}$, and thus the ground state energy and excitation energies may be computed in the same way as we described earlier.

What about the ground state of H? Assuming that all the λ_j are non-negative, it turns out that a state $|\psi\rangle$ satisfies $a_j^\dagger a_j|\psi\rangle = 0$ for all j if and only if $b_j^\dagger b_j|\psi\rangle = 0$ for all j, and so the ground state for the two sets of Fermi operators is the same.

This follows from a more general observation, namely, that $a_j^\dagger a_j|\psi\rangle = 0$ if and only if $a_j|\psi\rangle = 0$. In one direction, this is trivial: just multiply $a_j|\psi\rangle = 0$ on the left by $a_j^\dagger$. In the other direction, we multiply $a_j^\dagger a_j|\psi\rangle = 0$ on the left by a_j to obtain $a_j a_j^\dagger a_j|\psi\rangle = 0$. Substituting the CCR $a_j a_j^\dagger = -a_j^\dagger a_j + I$, we obtain

$$(-a_j^\dagger a_j^2 + a_j)|\psi\rangle = 0. \tag{14}$$

But $a_j^2 = 0$, so this simplifies to $a_j|\psi\rangle = 0$, as desired.

Returning to the question of determining the ground state, supposing $a_j^\dagger a_j|\psi\rangle = 0$ for all j, we immediately have $a_j|\psi\rangle = 0$ for all j, and thus $b_j|\psi\rangle = 0$ for all j, since the b_j are linear functions of the a_j, and thus $b_j^\dagger b_j|\psi\rangle = 0$ for all j. This shows that the ground state for the two sets of Fermi operators, a_j and b_j, is in fact the same. The excitations for H may be obtained by applying raising operators $b_j^\dagger$ to the ground state.

Exercise: Suppose some of the λ_j are negative. Express the ground state of H in terms of the simultaneous eigenstates of the $a_j^\dagger a_j$.

The Hamiltonian $H = \sum_{jk} \alpha_{jk} a_j^\dagger a_k$ we diagonalized earlier can be generalized to any Hamiltonian which is quadratic in Fermi operators, by which we mean it may contain terms of the form $a_j^\dagger a_k, a_j a_k^\dagger, a_j a_k$ and $a_j^\dagger a_k$. We will not allow linear terms like $a_j + a_j^\dagger$. Additive constant terms γI are easily incorporated, since they simply displace all elements of the spectrum by an amount γ. There are several ways one can write such a Hamiltonian, but the following form turns out to be especially convenient for our purposes:

$$H = \sum_{jk} \left(\alpha_{jk} a_j^\dagger a_k - \alpha_{jk}^* a_j a_k^\dagger + \beta_{jk} a_j a_k - \beta_{jk}^* a_j^\dagger a_k^\dagger \right) \quad (15)$$

The reader should spend a little time convincing themselves that for the class of Hamiltonians we have described, it is always possible to write the Hamiltonian in this form, up to an additive constant γI, and with α hermitian and β antisymmetric.

This class of Hamiltonian appears to have first been diagonalized in an appendix to a famous *Annals of Physics* paper by Lieb, Schultz and Mattis, dating to 1961 (volume 16, pages 407-466), and the procedure we follow is inspired by theirs. We begin by defining operators $b_1, \ldots, b_n$:

$$b_j \equiv \sum_k \left(\gamma_{jk} a_k + \mu_{jk} a_k^\dagger \right). \quad (16)$$

We will try to choose the complex numbers γ_{jk} and μ_{jk} to ensure that: (1) the operators b_j satisfy Fermionic CCRs; and (2) when expressed in terms of the b_j, H has the same form as $H_{\rm free}$, and so can be diagonalized.

A calculation shows that the condition $\{b_j, b_k^\dagger\} = \delta_{jk} I$ is equivalent to the condition

$$\gamma\gamma^\dagger + \mu\mu^\dagger = I, \quad (17)$$

while the condition $\{b_j, b_k\} = 0$ is equivalent to the condition

$$\gamma\mu^T + \mu\gamma^T = 0. \quad (18)$$

These are straightforward enough equations, but their meaning is perhaps a little mysterious. More insight into their structure is obtained by rewriting the connection between the a_js and the b_js in an equivalent form using vectors whose individual entries are not numbers, but rather are operators such as a_j and b_j, and using a block matrix with blocks γ, μ, μ^* and γ^*:

$$\left[\begin{array}{c} b_1 \\ \vdots \\ b_n \\ b_1^\dagger \\ \vdots \\ b_n^\dagger \end{array} \right] = \left[\begin{array}{cc} \gamma & \mu \\ \mu^* & \gamma^* \end{array} \right] \left[\begin{array}{c} a_1 \\ \vdots \\ a_n \\ a_1^\dagger \\ \vdots \\ a_n^\dagger \end{array} \right]. \quad (19)$$

The conditions derived above for the b_js to satisfy the CCRs are equivalent to the condition that the transformation matrix

$$T \equiv \left[\begin{array}{cc} \gamma & \mu \\ \mu^* & \gamma^* \end{array} \right] \quad (20)$$

is unitary, which is perhaps a somewhat less mysterious condition than the earlier equations involving γ and μ. One advantage of this representation is that it makes it easy to find an expression for the a_j in terms of the b_j, simply by inverting this unitary transformation, to obtain:

$$\left[\begin{array}{c} a_1 \\ \vdots \\ a_n \\ a_1^\dagger \\ \vdots \\ a_n^\dagger \end{array} \right] = T^\dagger \left[\begin{array}{c} b_1 \\ \vdots \\ b_n \\ b_1^\dagger \\ \vdots \\ b_n^\dagger \end{array} \right]. \quad (21)$$

The next step is to rewrite the Hamiltonian in terms of the b_j operators. To do this, observe that:

$$H = [a_1^\dagger \ldots a_n^\dagger a_1 \ldots a_n] \left[\begin{array}{cc} \alpha & -\beta^* \\ \beta & -\alpha^* \end{array} \right] \left[\begin{array}{c} a_1 \\ \vdots \\ a_n \\ a_1^\dagger \\ \vdots \\ a_n^\dagger \end{array} \right]. \quad (22)$$

It is actually this expression for H which motivated the original special form which we chose for H. The expression is convenient, for it allows us to easily transform back and forth between H expressed in terms of the a_j and H in terms of the b_j. We already have an expression in terms of the b_j operators for the column vector containing the a and $a^\dagger$ terms. With a little algebra this gives rise to a corresponding expression for the row vector containing the $a^\dagger$ and a terms:

$$[a_1^\dagger \ldots a_n^\dagger a_1 \ldots a_n] = [b_1^\dagger \ldots b_n^\dagger b_1 \ldots b_n] T. \quad (23)$$

This allows us to rewrite the Hamiltonian as

$$H = [b^\dagger b] T M T^\dagger \left[\begin{array}{c} b \\ b^\dagger \end{array} \right], \quad (24)$$

where we have used the shorthand $[b^\dagger b]$ to denote the vector with entries $b_1^\dagger, \ldots, b_n^\dagger, b_1, \ldots, b_n$, and

$$M = \left[\begin{array}{cc} \alpha & -\beta^* \\ \beta & -\alpha^* \end{array} \right]. \quad (25)$$

Supposing we can choose T such that $TMT^\dagger$ is diagonal, we see that the Hamiltonian can be expressed in the form of $H_{\rm free}$, and the energy spectrum found, following our earlier methods.

Since α is hermitian and β antisymmetric it follows that M also is hermitian, and so can be diagonalized for some choice of unitary T. However, the fact that the b_js must satisfy the CCRs constrains the class of Ts available to us. We need to show that such a T can be used to do the diagonalization.

We will give a heuristic and somewhat incomplete proof that this is possible, before making some brief remarks about what is required for a rigorous proof. I've omitted the rigorous proof, since the way I understand it is uses a result from linear algebra that, while beautiful, I don't want to explain in full detail here.

Suppose T is any unitary such that

$$TMT^\dagger = \begin{bmatrix} d & 0 \\ 0 & -d \end{bmatrix}, \quad (26)$$

where d is diagonal, and we used the special form of M to deduce that the eigenvalues are real and appear in matched pairs $\pm\lambda$. We'd like to show that T can be chosen to be of the desired special form. To see that this is plausible, consider the map $X \rightarrow SX^*S^\dagger$, where S is a block matrix:

$$S = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}. \quad (27)$$

Applying this map to both sides of the earlier equation we obtain

$$ST^*M^*T^TS^\dagger = \begin{bmatrix} -d & 0 \\ 0 & d \end{bmatrix} = -TMT^\dagger. \quad (28)$$

But $M^* = -S^\dagger MS$, and so we obtain:

$$-ST^*S^\dagger MST^TS^\dagger = -TMT^\dagger. \quad (29)$$

It is at least plausible that we can choose T such that $ST^*S^\dagger = T$, which would imply that T has the required form. What this actually shows is, of course, somewhat weaker, namely that $T^\dagger ST^*S^\dagger$ commutes with M.

One way of obtaining a rigorous proof is to find a T satisfying

$$TMT^\dagger = \begin{bmatrix} d & 0 \\ 0 & -d \end{bmatrix}, \quad (30)$$

and then to apply the cosine-sine (or CS) decomposition from linear algebra, which provides a beautiful way of representing block unitary matrices, and which, in this instance, allows us to obtain a T of the desired form with just a little more work. The CS decomposition may be found, for example, as Theorem VII.1.6 on page 196 of Bhatia's book "Matrix Analysis" (Springer-Verlag, New York, 1997).

Problem: Can we extend these results to allow terms in the Hamiltonian which are *linear* in the Fermi operators?

III. THE JORDAN-WIGNER TRANSFORM

In this section we describe the Jordan-Wigner transform, explaining how it can be used to map a system of qubits (i.e., spin-$\frac{1}{2}$ systems) to a system of Fermions, and vice versa. We also explain a nice applications of these ideas, to solving one-dimensional quantum spin systems.

Suppose we have an n-qubit system, with the usual state space C^{2^n}, and Pauli operators X_j, Y_j, Z_j acting on qubit j. We are going to use these operators to *define* a set of a_j operators acting on C^{2^n}, and satisfying the Fermionic CCRs.

To begin, suppose for the sake of argument that we have found such a set of operators. Then from our earlier discussion the action of the a_j operators in the occupation number representation $|\alpha\rangle = |\alpha_1, \ldots, \alpha_n\rangle$ must be as follows:

- Suppose $\alpha_j = 0$. Then $a_j|\alpha\rangle = 0$.
- Suppose $\alpha_j = 1$. Let α' be that vector which results when the jth entry of α is changed to 0. Then $a_j|\alpha\rangle = -(-1)^{s_\alpha^j}|\alpha'\rangle$, where $s_\alpha^j \equiv \sum_{k=1}^{j-1} \alpha_k$.

If we identify the occupation number state $|\alpha\rangle$ with the corresponding computational basis state $|\alpha\rangle$, then this suggests taking as our *definition*

$$a_j \equiv -\left(\otimes_{k=1}^{j-1} Z_k\right) \otimes \sigma_j, \quad (31)$$

where σ_j is used to denote the matrix $\sigma \equiv |0\rangle\langle 1|$ acting on the jth qubit. It is easily verified that these operators satisfy the Fermionic CCRs. This definition of the a_j is known as the *Jordan-Wigner transform.* It allows us to define a set of operators a_j satisfying the Fermionic CCRs in terms of the usual operators we use to describe qubits, or spin-$\frac{1}{2}$ systems.

The Jordan-Wigner transform can be inverted, allowing us to express the Pauli operators in terms of the Fermionic operators $a_1, \ldots, a_n$. In particular, we have

$$Z_j = a_j a_j^\dagger - a_j^\dagger a_j. \quad (32)$$

This observation may also be used to obtain an expression for X_j by noting that $X_j = \sigma_j + \sigma_j^\dagger$, and thus:

$$X_j = -(Z_1 \ldots Z_{j-1})(a_j + a_j^\dagger). \quad (33)$$

Substituting in the expressions for $Z_1, \ldots, Z_{j-1}$ in terms of the Fermionic operators gives the desired expression for X_j in terms of the Fermionic operators. Similarly, we have

$$Y_j = i(Z_1 \ldots Z_{j-1})(a_j^\dagger - a_j), \quad (34)$$

which, together with the expression for the Z_j operators, enables us to express Y_j solely in terms of the Fermionic operators.

These expressions for X_j and Y_j are rather inconvenient, involving as they do products of large numbers of Fermi operators. Remarkably, however, for certain simple *products* of Pauli operators it is possible to obtain quite simple expressions in terms of the Fermi operators. In particular, with a little algebra we see that:

$$Z_j = a_j a_j^\dagger - a_j^\dagger a_j$$

$$X_j X_{j+1} = (a_j^\dagger - a_j)(a_{j+1} + a_{j+1}^\dagger)$$

$$Y_jY_{j+1} = -(a_j^\dagger + a_j)(a_{j+1}^\dagger - a_{j+1})$$

$$X_jY_{j+1} = i(a_j^\dagger - a_j)(a_{j+1}^\dagger - a_{j+1})$$

$$Y_jX_{j+1} = i(a_j^\dagger + a_j)(a_{j+1}^\dagger + a_{j+1}).$$

Suppose now that we have an n-qubit Hamiltonian H that can be expressed as a sum over operators from the set $Z_j, X_jX_{j+1}, Y_jY_{j+1}, X_jY_{j+1}$ and Y_jX_{j+1}. An example of such a Hamiltonian is the transverse Ising model,

$$H = \alpha \sum_j Z_j + \beta \sum_j X_jX_{j+1}, \tag{35}$$

which describes a system of magnetic spins with nearest neighbour couplings of strength β along their $\hat{x}$ axes, and in an external magnetic field of strength α along the $\hat{z}$ axis.

For any such Hamiltonian, we see that it is possible to re-express the Hamiltonian as a Fermi quadratic Hamiltonian. As we saw in an earlier section , determining the energy levels is then a simple matter of finding the eigenvalues of a $2n \times 2n$ matrix, which can be done very quickly. In particular, finding the ground state energy is simply a matter of finding the smallest eigenvalue of that matrix, which is often particularly easy. In the case of models like the transverse Ising model, it is even possible to do this diagonalization analytically, giving rise to exact expressions for the energy spectrum. Details can be found in the paper by Lieb, Schulz and Mattis mentioned earlier, or books such as the well-known book by Sachdev on quantum phase transitions.

Exercise: What other products of Pauli operators can be expressed as quadratics in Fermi operators?

Problem: I've made some pretty vague statements about finding the spectrum of a matrix being "easy". However, I must admit that I'm speaking empirically here, in the sense that in practice I know this is easily done on a computer, but I don't know a whole lot about the computational complexity of the problem. One obvious observation is that finding the spectrum is equivalent to finding the roots of the characteristic equation, which is easily computed, so the problem may be viewed as being about the computational complexity of root-finding.

IV. APPENDIX ON MUTUALLY COMMUTING OBSERVABLES

Any undergraduate quantum mechanics course covers the fact that a mutually commuting set of Hermitian operators possesses a common eigenbasis. Unfortunately, in my experience this fact is usually proved rather early on, and suffers from being presented in a slightly *too* elementary fashion, with inductive constructions of explicit basis sets and so on. The following proof is still elementary, but from a slightly more sophisticated perspective. It is, I like to imagine, rather more like what would be given in an advanced course in linear algebra, were linear algebraists to actually cover this kind of material. (They don't, so far as I know, having other fish to fry.)

Suppose $H_1, \ldots, H_m$ are commuting Hermitian (indeed, normal suffices) operators with spectral decompositions:

$$H_j = \sum_{jk} E_{jk} P_{jk}, \tag{36}$$

where E_{jk} are the eigenvalues of H_j, and P_{jk} are the corresponding projectors. Since the H_j commute, it is not difficult to verify that for any quadruple j, k, j', k' the operators P_{jk} and $P_{j'k'}$ also commute. For a vector $\vec{k} = (k_1, \ldots, k_m)$ define the operator

$$P_{\vec{k}} \equiv P_{1k_1} P_{2k_2} \ldots P_{mk_m}. \tag{37}$$

Note that the order of the operators on the right-hand side does not matter, since they all commute with one another. The following equations all follow easily by direct computation, the mutual commutativity of the P_{jk} operators, and standard properties of the spectral decomposition:

$$P_{\vec{k}}^\dagger = P_{\vec{k}}; \quad \sum_{\vec{k}} P_{\vec{k}} = I; \quad P_{\vec{k}} P_{\vec{k}'} = \delta_{\vec{k}\vec{k}'} P_{\vec{k}}. \tag{38}$$

Thus, the operators $P_{\vec{k}}$ form a complete set of orthonormal projectors. Furthermore, suppose we have $P_{\vec{k}}|\psi\rangle = |\psi\rangle$. Then we will show that for any j we have $P_{jk_j}|\psi\rangle = |\psi\rangle$, so $|\psi\rangle$ is an eigenstate of H_j with eigenvalue k_j. This shows that the operators $P_{\vec{k}}$ project onto a complete orthonormal set of simultaneous eigenspaces for the H_j, and will complete the proof.

Our goal is to show that if $P_{\vec{k}}|\psi\rangle = |\psi\rangle$ then for any j we have $P_{jk_j}|\psi\rangle = |\psi\rangle$. To see this, simply multiply both sides of $P_{\vec{k}}|\psi\rangle = |\psi\rangle$ by P_{jk_j}, and observe that $P_{jk_j} P_{\vec{k}} = P_{\vec{k}}$. This gives $P_{\vec{k}}|\psi\rangle = P_{jk_j}|\psi\rangle$. But $P_{\vec{k}}|\psi\rangle = |\psi\rangle$, so we obtain $|\psi\rangle = P_{jk_j}|\psi\rangle$, which completes the proof.